# Table of Contents

Deori tonal contrasts: A generational account  
*Prarthana Acharyya and Shakuntala Mahanta*

Does non-finiteness predict modality in Meeteilon?  
*Padmabati Achom*

Uses of ha / ha vai / ha sma vai with or without the Narrative Perfect and Language Layers in the Old Yajurveda-Samhita Texts  
*Kyoko Amano*

New database resources for the Indo-European lexicon  
*Cormac Anderson*

On Aktionsart issues relating to dative-experiencer predicates in Indic  
*Tista Bagchi*

Minimalist analysis of beneficiary/ goal datives in Assamese  
*Chandrika Baruah*

A Semantic Perspective on Modified Numerals in Bangla  
*Nandini Bhattacharya*

Agreement and Pronominalisation: Two different ways to Agree  
*Tanmoy Bhattacharya*

Meiteilon DPs and comparative syntax  
*Tanmoy Bhattacharya and Tor Anders Åfarli*

How to write Koêc (Kiranti, Tibeto-Burman): Linguistic and social aspects in the development and implementation of a writing system  
*Dörte Borchers*

"Historical semiosis of Vedic words: The cases of vajra and ētman.”  
*Jens Braarvig*

The lexicon of the "act of accepting (pratigraha)": an approach to the multilayered Vedic culture  
*Maria Piera Candotti and Tiziana Pontillo*

The typology of pragmatic case marking in Tibeto-Burman  
*Alexander Coupe*

Experiencer subject verbs with object alternation in Vedic  
*Eystein Dahl*
Word-minimality vis-à-vis morphophonemics of genitives in NKB
Shyamal Das

Translation and bilingual lexicography: the case of French to Urdu and Urdu to French, analyzing some specialized bilingual glossaries
Alain Desoulières

Dissociation of Case and Agreement in Santali
Mayuri Dilip and Rajesh Kumar

Bagēi ergativity: attempt of an in depth study
Andrea Drocco

A Linguistic Analyzer for Building a Database of Old Marathi
Oliver Hellwig and Ashwini Deo

Multi-layer annotation of the Rigveda – Methods and Scholarly Applications
Oliver Hellwig and Sven Sellmer

Variations in Meeteilon by Dr Hidam Gaurashyam Singh (Research Fellow), University of Delhi
Gaurashyam Singh Hidam

Out of India? A new look at the linguistic arguments and evidence
Hans Henrich Hock

The Reflexes of Retroflex Approximant (ensus) in Kanna Language
Shashikala Honnappa

Variation of Lexical Items and the Changing Need of Identity Expressions
Thapasya J and Rajesh Kumar

On-line Hindi and Sanskrit dictionary for student community in Croatia and elsewhere
Marijana Janjić, Sara Librenjak, Kristina Kocijan and Lana Orešić

Recognition and multi-layered analysis of converbs in early NIA
Rafał Jaworski and Krzysztof Stronski

The Sky in Ceremony: Textual Variants in Sūryē’s Bridal Hymn (RV 10.85, AVP 18.1-14, AVŚ 14)
Kristen de Joseph

Polysemy and cognitive linguistics. A case of vána
Joanna Jurewicz

Is English supply’ the only route to livelihoods? Language and work in a multilingual economy
Anuradha Kanniganti
Syntactic Variation: A comparison between Meeteilon and Non-Meeteilon speakers  
_Alfina Khaidem_

Resultative Constructions in Hindi-Urdu and Other Indo-Aryan Languages. Synchronic and diachronic perspectives  
_Rainer Kimmig_

Quotative Clauses in Telugu  
_Madelyn Kissock_

Translating Cuisine Verbs from Telugu to Bangla: A Feature Interactive Model (FIM)  
_Rajyarama Koppaka and Abhijit Debnath_

Word order, agreement and case in Hindi from a Dynamic Syntax perspective  
_Andriana Koumbarou_

Labile verbs in Old Indo-Aryan: their origin, decline and fall (A typological perspective)  
_Leonid Kulikov_

The translatability of kavïkratu and the lexical diversity of rigvedic poetological lexicon  
_Frank Köhler_

Northwestern Indo-Aryan and the rise of diversity in the Hindu Kush-Karakoram  
_Henrik Liljegren_

Some aspects of noun morphology in Toto  
_Patrycja Ava Markus_

Discourse particles in Hindi and the relation between grammar, enunciation and "modal" meanings  
_Annie Montaut_

Does denaa GIVE as a V2 render benefactive meaning in Hindi?: A corpus-based comparative analysis in Hindi and Japanese  
_Miki Nishioka_

Exploring the language layer of the Rigvedic dánastuti genre  
_Georges-Jean Pinault_

The Early Vedic pipáya- type  
_Roland Pooth_

The Role of Pitch in Gurung Tone  
_Danielle Ronkos_

áhir budhnyà and dundubhì: the Serpent of the Deep and the earth-drum: hypothesis of etymological and/or cultural connections  
_Paola Maria Rossi_
Diachronic, diatopic and diastratic developments of inherited Indo-Iranian lexicon and phraseology in the history of the Vedic text corpus: historical (Rigveda – Khilas – Atharvaveda – Yajurveda) and comparative (Veda vs. Avesta) perspectives
Velizar Sadovski

A comparative study of the Anaphors in Meiteilon, Nyishi and Toto
Atanu Saha

Atharvaveda Paippalēda, Book 17, Anuvēka 6. The draft-ox vrata, a missing link between Vṛetya culture and Pēṣupatism.
Umberto Selva

A pragmatic account of definite and indefinite descriptions in Hindi - the case of 'hai' and 'hotē hai' forms
Ghanshyam Sharma

Sociocultural Variables in Language Maintenance: Thanjavur Marathi
Kamal Sridhar

Breaking the language barriers: Language mixing as a civilizational strategy in South Asia
S.N. Sridhar

Labelling Finiteness as a Semantic Feature
Shubham Srivastava and Sarada Biswas

Converbs in light verb constructions in New Indo-Aryan
Joanna Tokaj

Saartje Verbeke

Attitudes Towards English in a Remote North Indian Town
Emma Walters

Echo-formation in Kati as a neighbouring language of Kalasha and Khowar
Noboru Yoshioka

Role clusters in Indo-Aryan (diachronic and synchronic aspects)
Boris Zakharyin and Liudmila Khokhlova

A preliminary study of the Tamil comparative postpositions
Wojciech Zeyland
Deori tonal contrasts: A generational account

Shakuntala Mahanta and Prarthana Acharyya
Indian Institute of Technology Guwahati

Deori is a Tibeto-Burman language spoken in the eastern part of Assam and Arunachal Pradesh has been listed as an ‘endangered language’ in Encyclopaedia of the world’s endangered languages (2007a). UNESCO has explicitly listed Deori as a ‘definitely endangered’ language. For the Deori communities Assamese has become the lingua franca for both intra and inter group communication. Deori is also less frequently by the younger generation as they prefer Assamese over their mother tongue thereby making the language significantly endangered. In the existing literature on Deori, there have been many observations about tone. Brown (1895) noted that the lexical items in Deori are distinguished either by tone or by nasalization (apart from other segmentally different minimal pairs). Goswami (1994) clearly introduced the presence of three tones in Deori which changes the lexical meaning of the word. Jacquesson (2005) states, that there is a tonal distinction in Deori. At the same time Jacquesson also mentions that there is a gradual loss of tonal opposition in the language nowadays and it’s a moribund feature in the language. Saikia (2013) noted the presence of three tones in the language – high, level and low tone. Mahanta et.al (forthcoming) states the tonal contrast, high and low, is prevalent in the language though marginal it is. These generalizations of tone in Deori prompted us to look into the tonal distinction in the speech of the younger generation. Five native speakers (18-30 years) participated in the production test. We prepared a word list to record possible tonal contrasts in a carefully constructed production experiment. In the experiment the target words were uttered within a predetermined carrier sentence. The words were uttered in a fix sentence frame (I X said) where X is the target word. All the speakers were bilingual; they were equally fluent in both Assamese and Deuri. Recordings were done in a TASCAM recorder with a Shure headworn microphone. Praat was used to label the target syllable and yield F0 values. A normalized f0 script was used to get the measurements such as duration, intensity, f0 mean, f0 vowel mid, maximum pitch minimum pitch, pitch range, standard deviation, non-normalized values each at 10% interval of time, and z score values also at 10% interval of time. The non-normalized values each at 10% interval of time was transformed into z score via Lobanov’s (1971) z-score metric (z= (F-μ)/σ). In order to confirm whether differences between the tonal categories are significantly different or not one way repeated measure ANOVA test was conducted for all the words separately to get the within subject significant effect on tone for both the vowels.. In the repeated measures ANOVA tone (high and low) was considered as the fixed factor, speaker as the random factor and z score normalized f0 at 10% interval of time as the dependent variable and. The results of the statistical analysis of the younger generation were compared with those of the older generation (Mahanta et al forthcoming) for a better understanding of the tonal representation. The results show a tonal distinction in few homophonous words in both younger and older generation but not consistent throughout. For monosyllabic words the tonal distinction is maintained for the set of words examined except ʃi ‘blood/to make’ which shows no significant difference of tone (p>0.05 [(F (1, 54) = 2.40) p = .127] (Figure 1a-b). Whereas, in disyllabic words, words like ʃitu ‘old/wife of younger brother’ (p<0.05 [(F (1,54) =20.30) p = .000] (vowel 1) ; (p<0.05 [(F (1,54) =.045) p = .002] (vowel 2) and kiri ‘poor/to furnish with heddles’ (p<0.05 [(F (1,54) =72.01) p = .000] (vowel 1) ; (p<0.05 [(F (1,54) =3.56) p = .000] (vowel 2) (Figure 2a-b) showed a significant effect of tone. Whereas words like nini
‘drink/hold’, *tiri* ‘banana/hang’, *tfija* showed no significant effect of tone unlike older generation (Figure 3a-b). And there are words like *uzu* ‘navel/bamboo tube’ (p<0.05 [(F (1, 54) =175.27) *p* = .000] (vowel 1) and *aku* ‘ear/upland’ (p<0.05 [(F (1, 54) =72.29) *p* = .000] (vowel 2) (Figure 4a-d) where the tonal distinction is maintained in either edge of the word, left or right, unlike older generation. Compared to the results cited in Mahanta et al (forthcoming) the tonal distinction in the younger generation has comparatively lessened. However, some tonal pairs with increased information load still maintained the distinction.
References


Does non-finiteness predict modality in Meeteilon?
Padmabati Achom
CASL, University of Delhi
padma.achom@gmail.com

Meeteilon is a Tibeto-Burman language which is spoken in the North-Eastern part of India. Meeteilon neither has an auxiliary verb nor a modal verb unlike English. However, there are four roots which can express modality in Meeteilon. I will thus refer to these roots as ‘modal roots’ in Meeteilon. The modal roots in Meeteilon are as follows:

(i) oi- “possibility” (Epistemic/Deontic)
(ii) ya- “possibility” (Epistemic/Deontic)
(iii) ta- “necessity” (Epistemic/Deontic)
(iv) ŋǝ “can” (Ability)

Modality is concerned with the status of the proposition that describes the event (Palmer, 2001); among other things, modality is contributed by modal particles/ affixes. In Meeteilon, modality is expressed through:

(i) Modal roots
   e.g. John catid-pa ya-i
        John go-INF possible-SAsp
       ‘It is possible for John to go.’
       Epistemic possibility
(ii) Modal affix
     e.g. John catid-l₇am-(gǝ)ni
          John go-Deic-Fut
       ‘John might have gone.’
(iii) Double Negation
     e.g. naŋ catid-ta-ba ya-de/-roi
          you go-Neg-NMZ agree-Neg
       ‘You must go./You must have to go.’
       Deontic Necessity

The modal roots ya- and ta- always take a non-finite embedded clause as in (1)

(1) John catid-pa ya-i
    John go-INF possible-SAsp
    ‘It is possible for John to go.’

In this paper, I will focus on the two modal roots i.e. ya- and ta- which always take a non-finite embedded clause. As shown in (1) the modal root ya- takes a non-finite embedded clause giving the epistemic modality interpretation. However, when the modal root ya- takes a finite embedded clause, the sentence becomes ungrammatical as shown in (2).

*(2) John catid-le ya-i
    John go-Perf possible-SAsp
    ‘It is possible John have gone.’

However, if another non-finite clause is inserted in between the two clauses in (2), then the sentence is grammatical as in (3).

(3) John catid-le hay-ba ya-i
    John go-Perf say-INF possible-SAsp
    ‘It is possible to say John have gone.’
So, the modal root *ya-* can only select a non-finite embedded clause as shown in the examples (1) and (3). Now, let us see how the modal root *ta-* behaves when it selects a non-finite embedded clause in (4) and a finite embedded clause in (5).

(4) mahak gari manuŋ-da tum-ba ta-i
   he/she.3P vehicle inside-LOC sleep-INF necessary-SAsp
M1: ‘It is necessary for him to sleep inside the vehicle.’ (as he has no place to sleep.)
M2: (I) hear the sound of his sleeping inside the vehicle.
Here, *ta-* is ambiguous as it expresses the deontic modality as well as the stative verb ‘hear’ when the embedded clause is non-finite. Now, when the modal root *ta-* takes a finite embedded clause, the modality interpretation is lost and the root *ta-* only interprets the stative verb ‘hear’ as shown below:

(5) mahak gari manuŋ-da tum-me (hayna) ta-i
   he/she.3P vehicle inside-LOC sleep-Perf that.COMP hear-SAsp
   ‘I heard that he slept inside the vehicle.’
Based on the behaviour of the root *ta-* in (4) and (5), it can be concluded to some extend that if the embedded clause structure is non-finite, a modal interpretation becomes possible. However, if the embedded clause is finite, the modality interpretations disappear. This shows that the syntactic structure can predict the function of the modality in Meeteilon. Though, the ambiguity of the root *ta-* in (4) can be explained as the root with a different categoriser (Embick, 2015), the syntactic structure of finite embedded clause in (5) which is not possible to have the modality interpretation while a non-finite embedded clause can have the modality interpretation will be further studied in the paper.

COMP Complementizer LOC Locative
DEIC Deictic Perf Perfect
FUT Future SAsp Simple Aspect
INF Infinitive

References:
Uses of *ha*/*ha vái*/*ha sma vái* with or without the Narrative Perfect and Language Layers in the Old Yajurveda-Saṅhitā Texts

Kyoko Amano (Research Fellow of Osaka University)

It is well-known that the perfect with the particle *ha* is used in the narrative sense in the younger Vedic prose. In the older Vedic prose, Maitrāyaṇī Saṅhitā (MS), Kāṭhaka-Saṅhitā (KS) and Taittirīya-Saṅhitā (TS), there is a distribution for past tense categories: the imperfect for the gods myths and the perfect for the period of predecessors. It is supposed that the latter use of the perfect was extended to the area of the former use of the imperfect.

In the present paper, uses of the particle *ha* in the three Yajurveda-Saṅhitā texts will be examined. The result suggests that the different uses of *ha* characterize different language layers in these texts. The following points are of special interest:

1) *ha* and *ha vái* with the present verb characterize a logical consequence derived from the context; so it means ‘namely, in conclusion.’ Many examples of this use are found in the MS, but less in KS and TS.

2) *ha sma* (vá) with the present indicative indicates a repeated and habitual action in the past. In MS, it is used almost always with āha (functionally present) and indicates a ritual opinion of predecessors; ‘(A predecessor, i. e. Aruṇa Aupaveśi or Keśin Satyakāmi, used to say.’ KS and TS have examples with other verbs than āha.

3) *ha* with the perfect hardly appears in MS, but KS and TS have a lot of examples.

4) *ha vái ... uvāca* in KS corresponds to *ha sma* (vá) āha in MS in the parallel passages. It can be the origin of the narrative use of the perfect.

As a result, it can be said that the language of KS and that of TS are close to each other and that the language of MS has different features from them, although it is generally supposed that MS and KS belonged to the same branch but TS to another. And linguistic innovations occurred not always gradually, but at certain innovative authors. It can provide a new perspective for clarifying the relation among three texts and the composing process of them.

References
New database resources for the Indo-Iranian lexicon

Cormac Anderson
MPI-SHH Jena

This contribution discusses ongoing work on two massive cross-linguistic database resources, which will be of use for linguists working on the lexicon and phonology of Indo-Iranian languages. CoBL (Cognacy in Basic Lexicon) is a database framework for the exploration of cognacy relationships in basic lexicon, while Sound Comparisons aims to document phonetic diversity.

The CoBL project, hosted at the Max Planck Institute for the Science of Human History in Jena, aims to develop a database framework amenable to qualitative and quantitative exploration of cognacy relationships in basic lexicon. The first instantiation deals with the Indo-European language family and will be completed within 2017. Data is collected and entered by a consortium of language experts affiliated to the project, working according to strict guidelines for data entry. Lexical items entered include phonetic and phonemic transcriptions, as well as transliterations where necessary, and are coded for their cognacy relationships. The handling of loanwords is particularly relevant for Indo-Iranian languages, given the prevalence of borrowings in the basic vocabulary of many Indo-Iranian languages.

The contribution also briefly discusses Sound Comparisons, a database Explorer resource for phonetic diversity, which is not yet implemented for Indo-Iranian languages, but which could easily be extended to them. As work on both CoBL and Sound Comparisons is ongoing, and as there are remaining gaps in the coverage of Indic languages in CoBL, expressions of interest in collaboration from workshop participants are particularly welcome.
On Aktionsart issues relating to dative-experiencer predicates in Indic

Tista Bagchi
UGC-Centre for Advanced Studies in Linguistics, University of Delhi

Dative experiencers are attested in the Germanic branch of Indo-European (Hock 1990) notably including the Scandinavian language group (Barðdal et al. 2014; Schätzle, Butt, and Kotcheva 2015), Italic (Giusti and Iovino 2015/2016), and the Iranian (Windfuhr 1990) and Indic (Montaut 2013) subgroups of the Indo-Iranian branch, as well as in the Dravidian family of languages in South Asia (e.g., in the Dravidian language Malayalam: Mohanan and Mohanan 1990). The discussion on dative experiencers has thus far been focused primarily on the issue of whether or not they count as subjects, with differing positions emerging on the issue (e.g., Butt et al. 2006 versus Barðdal et al. 2014). A poorly explored question thus far is the characterization of Aktionsart or “inner aspect” (Vendler 1957/1967; Smith 1991; called “situation type” by Saeed 2002: 106-114) of the classes of predicates that take dative-experiencer arguments. This is all the more surprising since a number of observations have been recorded in the literature about other kinds of semantic characteristics that such predicates are grouped by (on the basis of abundant data noted from both introspective-observational and corpus sources), such as physiological and psychological state expressions and, occasionally, possession or its achievement (cf., for example, Hindi mujhe ek khabar milaa “I-dat one news-item came-into-possession” ‘I received a piece of news’). In this paper, it is noted that the broad kinds of “inner aspect” or Aktionsart classes that one can identify amongst such predicates in the modern Indic languages that have been examined thus far are the following: (i) states (as in mujhe yeh aččhaa lǝg-rǝhaa hai “to-me this good striking is” ‘I like this’, (ii) semelfactives (as in use čo lǝgii “to-him/her injury struck” ‘s/he got injured’) and (iii) in very limited cases, achievements (as in use kučh rihaayǝt milii “to-him/her some relief was-obtained” ‘s/he got some relief’). The paper proceeds to take a brief overview of the manifestation of these Aktionsart classes in dative-experiencer predicates in Sanskrit and certain selected Middle Indic forms (following especially Masica’s 1991 survey of dative-experiencer predicates from a diachronic perspective). It is then argued that these three Aktionsart classes do not exhaust the inductively known totality of dative-experiencer predicates in the Indic languages studied for these at the major recognized historical stages for Indic (Haldar & Bagchi 2000/2007), but that these may include at least one more Aktionsart type, involving a determinate starting-point followed by a stative phase, i.e., a “mirror image” of sorts to the “accomplishment” Aktionsart, that researchers on Aktionsart in South Asian languages may also need to take into account.

References


Giusti, Giuliana, and Rossella Iovino. 2015 / revised draft, 2016. A comparative analysis of psychological verbs in Latin and Italian. Talk delivered at the 45th Linguistic Symposium on Romance Languages (LSRL), Campinas, May 7-10, 2015. [Under submission for publication.]


MINIMALIST ANALYSIS OF BENEFICIARY/GOAL DATIVES IN ASSAMESE

Chandrika Baruah
M.Phil, University of Delhi

This paper aims to provide a minimalist analysis of the goal/beneficiary dative constructions in the Eastern dialect of Assamese. Assamese is a modern Indo-Aryan (IA) language spoken in the state of Assam, in north-eastern part of India. Assamese, like all other IA languages, has descended through several intermediate stages (Middle Indo-Aryan), from Sanskrit.

Dative case is marked in Eastern Assamese (henceforth EA) using case markers, loi or ḍk.

Goal dative is expressed using the case marker loi attached to the indirect argument (as shown in 1 and 2).

1) ram-ɛ tuma-loi kitap-khɔn pɔtha-l-e (Goal)
   Ram-ERG you-DAT book-CL send-PAST-3
   ‘Ram sent the book to you.’

2) mɔi deuta-loi sithi e-khɔn likh-il-u. (Goal)
   I father-DAT letter one-CL write-PAST-1
   ‘I wrote a letter to my father.’

In addition, loi also acts as the benefactive marker in EA. It is attached to the indirect argument of the ditransitive verbs (as shown in 3 and 4). As can be observed, there is no overt difference between (1) and (4). The interpretation of the semantic role of the indirect argument, goal or benefactive, in such situations, is dependent on the context and the intention of the speaker. The paper puts forward a syntactic argument to account for this ambiguity.

3) ram-ɛ tuma-loi kitap-khɔn di-l-e (Benefactive)
   Ram-ERG you-DAT book-CL give-PAST-3
   ‘Ram gave the book for you.’

4) ram-ɛ tuma-loi kitap-khɔn pɔtha-l-e (Benefactive)
   Ram-ERG you-DAT book-CL send-PAST-3
   ‘Ram sent the book to you.

It is also observed that EA invariably uses the marker ḍk, only, to express goal dative when the indirect object is the argument of the verb ‘give’ (as shown in 5).

5) ram-ɛ tuma-k kitap-khɔn di-l-e (Goal)
   Ram-ERG you-DAT book-CL give-PAST-3
   ‘Ram gave the book for you.’

The paper also makes an attempt to explain the beneficiary reading that we get with loi by tracing the historical development of this case marker. According to Grierson (1927),
Magadhi Apabhramsa, which corresponded to the old Eastern Prakrit, was the common source of all eastern IA languages (Assamese, Bangla, Oriya and so on). Kakaty (1941), pointed out that *loi* in Assamese has come down from –*ka+lagi (lai)* from Magadhi, where *ka* was the genitive marker in Magadhi. In Magadhi genitive marker was used in dative constructions (in conformity with Old Indo Aryan use) to indicate direction or purpose (as shown in 6). This marker –*ka + la*gi is still evident in the speech of the Kamrupi (Western dialect of Assamese) speakers (as shown in 7).

6) krsnara pasa-ka lagi (direction)  
Krishna side-GEN DAT  
‘to Krishna’s side’

7) mu-r ma-ɛ mu-k legi upɔhɔr poth-e di-s-il.  
1S-GEN mother-ERG 1Sg-GEN to/for gift send-NF give-PERF-PAST  
‘My mother sent a gift to/for me.’

In modern EA, *ka* is dropped before *lai* (Kakaty, 1941). Further, it can be assumed that *loi*, a postposition, has transformed from *lagi* (‘to’ or ‘for’) into a dative case marker.

For the analysis of the dative constructions Adger (2002), Marrantz (1993) and Pylkkänen (2002) have been followed. Based on Pylkkänen (2002), it is observed that Assamese has both high applicative (between vP and VP) and low applicative (between VP and theme). The high applicative head denotes benefactive relation between the event and the indirect DP. The low applicative on the other hand indicates goal relation. The goal/benefactive DPs are in the [Spec, AppP]. In both the cases, appl head contains [ucase: dative] which checks and values [ucase: dative] on the goal/benefactive at [spec, ApplP] and the v head has [ucase:accusative] that checks and values [ucase: accusative] in the the theme DP via Agree. The paper proposes the raising of theme DP to the specifier of the verbal projection for case reasons.

The paper is not only able to clearly account for the structure of dative constructions in Assamese but also in providing a syntactic argument to resolve the case of ambiguity observed between sentences like (1) and (4).

REFERENCES


A Semantic Perspective on Modified Numerals in Bangla

Nandini Bhattacharya
University of Delhi

Modified numerals in natural language semantics are defined as denoting cardinal numbers that are modified by a quantifier or degree modifiers (+ adposition) and quantifies over the variable sets of the number. In Bangla (a.k.a. Bengali), there are a range of modified numeral expressions that occur as a quantificational phrase. The cardinal numbers in Bangla can be modified by quantifiers, adjective quantity words, degree modifiers (i.e. superlative quantifiers or comparative quantifiers) and lexical counting words. The modified numeral phrases can be observed in the example (1):

1. Ram porikhae ašir opor(e) nɔmbor peyəche
   Ram exam -loc. 80 -gen. over -loc. number got.pres.3rd
   ‘Ram got over 80 marks on the exam.’

Moreover, a numeral in Bangla can be simultaneously modified by two modifiers that further define the cardinality of the numerals. This phenomenon is illustrated by the following example (2):

2. kichu loker ɔnt̪ot̪o ʒutor beši gaɾi acʰe
   some people –gen. at least two –cls. –gen. more car be.pres.hab.
   ‘Some people have more than at least two cars.’

The modified numerals also occur embedded under the scope of negation. This is illustrated by the examples (3 & 4):

3. boi ʒutor ɡam ekšor kom noe
   book two -cls. –gen. price hundred -gen less neg.
   ‘These two books do not cost less than 100.’

4. Sita źinter beši boi pɔreni
   Sita three -gen. more book read.perf. neg.
   ‘Sita did not read more than three books.’

Interestingly, the modifiers occur both proceeding to and postpositionally with the numerals. In contrast to languages such as English and French, which have overt comparative particles (e.g. than, que etc.), Bangla has an optional comparative postposition occurring with the numeral before the degree adjective. This phenomenon can be observed in the examples, such as, / ɡešer çe beši/ ‘more than ten’, / źinter ɡeke kom/ ‘less than three’ etc.

In this paper, I present an overview of the semantic distribution of the adjectival modifiers that modify the numerals in Bangla. In addition, I examine the occurrences of these composite modified numeral phrases in different types of sentences in Bangla and analyse their syntactic ordering (Szabolcsi, 2010). Along with this, I also put forward an account of
the scope interaction of these modified numerals with negation following (Nouwen, 2008, 2010). In conclusion, this paper aims to provide a formal analysis of the modified numerals as scalar expressions in Bangla.

References:


Krifka, Manfred. 2007. “More on the difference between more than two and at least three”.


From grammar to meaning: The spontaneous logicality of language, 139–171. Cambridge, UK: Cambridge University Press.


The fact that the “Central Māgadhan Prākrit” (CMP) languages (Bhattacharya 2016) like Maithili, Magahi, Angika, etc. all have what is known as “multiple” agreement, may indicate a possible contact situation where the substratum Munda languages influence these languages in terms of multiple agreement (as conjectured in Chatterji 1926). Given that some of languages spoken in areas adjoining the CMP geographical region are in fact Munda languages (like Mundari, Santali, Korwa, etc.), this is a reasonable thesis. Chatterji further considers development of the agreement system in CMP languages as a later development as the language of Vidyapati (14th C) had, according to him, a much simpler verb-system, “...with its freedom from the ramifications of pronominal infixes and affixes.”

Comparing this impression on the so-called complex verbal agreement system with the overall framework of studying these languages in the much earlier study of Grierson (1887, 1903), one cannot help but notice a much more refined analysis of the same phenomenon, who does not furthermore identify the complexity of the agreement pattern to Munda languages. However, Grierson (1909) himself in LSI, III, conjectures that various traits possessed by the many Himalayan languages are proto-Munda, and pronominalisation is one such trait, deriving thereby the conclusion that Mundas once lived in the Himalayas.

Thus, there seems to be a contradiction in the two views: whereas for Chaterji, multiple agreement in CMP languages is due to “southern” Munda influence, for Grierson, Munda languages basically arrived from the north. This northern origin of the Munda languages is an older theme to be traced at least in Max Muller (1854) and Forbes (1881) in terms of the so-called Turanian origin of a pre-Aryan race in India.

Add to this the fact that there are many Tibeto-Burman languages that show agreement/pronominalisation phenomenon, are we then to suppose that all those too also have come under the influence of Munda in some way or other? I will instead suggest that the apparent similarity of a certain phenomenon across unrelated languages is not a sufficient condition for/ evidence of migration through a certain landmass, and that Pronominalisation may have developed in Tibeto-Burman languages independently.

Furthermore, the paper will investigate whether the phenomenon shown in Munda languages is to be considered agreement at all. I will run various diagnostics to argue that the concord phenomenon in Munda languages is different from multiple agreement seen in CMP languages. This will be argued on the basis of at least the following types of evidence:

(i) Clitics as forms of pronominals
(ii) Optionality of agreement marking
(iii) Bi-personal verb forms as not being the norm
(iv) Pro-clitic split
(v) Presence of applicative suffixes

Furthermore a crucial difference between CMP languages and Munda languages in terms of the sequence of the ‘agreement’ morphemes postverbally will be shown to accentuate the
difference between the two phenomena, and more importantly, it highlights a syntactic (rather than prosody-based) analysis of the theory of clitic-placement in Munda languages. A syntactic analysis of clitics was missed by earlier scholars because the crucial feature of cyclicity was not taken into consideration.

References

Müller, Max. 1854. Letter to Chevalier Bunsen on the Classification of the Turanian Languages.
Meiteilon DPs and comparative syntax

Tanmoy Bhattacharya  
Department of Linguistics  
University of Delhi  
tanmoy1@gmail.com

Tor A. Åfarli  
Norwegian University of Science and Technology, Trondheim  
tor.aafarli@hf.ntnu.no

One challenge that any study of observed linguistic patterns faces is whether the patterns can be derived from general underlying properties of grammar. In this paper we turn our attention to the phenomenon of the so-called “double” determiners/definiteness in two very different languages, which we compare in the spirit of comparative biology where two languages/varieties (or species) – hitherto unrelated either genetically or areally – are compared with the hunch that this type of comparison might tell us something valuable about common underlying properties where they are least expected.

The apparent surface-like similarity that we are interested in, namely the phenomenon of double determiners/definiteness, is found in Meiteilon, a language of the Kuki-Chin subgroup of the Tibeto-Burman group of languages spoken in the North-Eastern state of Manipur, and in the Halsa dialect, a coastal dialect of Norwegian (Indo-European) spoken in a small community in the northern part of Western Norway. Examples to be compared include the ones in (1) (Meiteilon) and (2) (the Halsa dialect) where in both languages a free determiner agrees with an affixal determiner on the noun.

(1)  
a. əsi  
thisDEF+PROX  
čiŋ-si  
mountain-DEF+PROX  
‘This mountain.’
b. aðu  
thatDEF+DIST  
kar-du  
car-DEF+DIST  
‘That car.’

(2)  
a. te  
thisDEF+PROX  
fjell-et  
mountain-DEF  
‘This mountain.’
b. na  
thatDEF+DIST  
bil-en  
car-DEF  
‘That car.’

Notice that agreement is with both definiteness and the proximal/distal property in Meiteilon, but only with definiteness (and gender and number; not shown here) in the Halsa dialect. We argue as a first approximation that Meiteilon and the Halsa dialect show the same basic structure, as indicated in (3), where the free demonstrative determiner is a D element in the upper functional domain of the DP, and where the determiner affix is in an F head in the lower functional domain, with the N head raising to F to acquire the affix in both languages.

(3)  
[ D … [ F [N] ]]  

Notice that F is the position of any inflectional features that are situated in the lower functional domain of a DP in a language, e.g. gender, number, definiteness, proximal/distal. (3) has been argued to be the right structure for the analysis of DPs in Norwegian and other Indo-European languages (see Grimstad & al. 2016). Meiteilon seems to follow this same general pattern, the only difference being the type of inflectional properties being involved.
However, there are cases in Meiteilon where the upper and lower proximal/distal features in the DP need not match like they do in (1) above, see (4a) (based on Achom et al 2013). Moreover, use of a noun in addition to the demonstrative as in (4b) indicates the selection of different nouns by the two sets of proximal/distal features.

(4)  

a. \textit{əsi(-gi) mə-ča-du}  
\textit{this(-Gen) 3-child-that}  
‘That child of this(3P)’  
b. \textit{cawbi-si(-gi) mə-ča-du}  
\textit{Chaobi-thisDEF+PROX(-Gen) 3-child-DEF+DIST}  
‘That child of Chaobi’

Since the upper demonstrative and the lower affixal demonstrative play different syntactic and semantic roles, they cannot be considered to agree or to be copies of each other.

We propose an analysis of this construction where the initial noun+demonstrative is an embedded independent DP in <Spec, DP> of the superordinate noun phrase, and where the genitive affix is situated in D of the that phrase, cf. the partial representation of (6b) given in (7):

(7)  

\[
[\text{DP1} [\text{DP2 cawbi-si}] [\text{D1 gi}] \ldots [\text{F} [\text{N}]]]
\]

We will support this analysis by showing that the Halsa dialect has a strikingly similar structure, except that the Halsa dialect does not have an affixal proximal/distal demonstrative, but only show upper and lower agreement with regard to definiteness, gender, and number. Still, the Halsa dialect can signal the proximal/distal property by employing an adverb (\textit{der} ‘there’, \textit{her} ‘here’), cf. the example (8a) and its partial structure in (8b):

(8)  

a. \textit{ne mannen sine ungar (der)}  
\textit{thisDEF+PROX man+DEF+SG POSS+PL child+PL thereDIST}  
‘Those children of this man.’  
b. \textit{[DP1 [DP2 ne mannen] [D1 sine] \ldots [F [N]]}}

The addition of the optional distal adverb in (8a) shows that the proximal/distal properties of the DP in the specifier position and the superordinate DP may be different, like in Meiteilon. Comparing (7) and (8b), we argue that even with regard to this construction, Meiteilon and the Halsa dialect are remarkably similar structurally.

Last, we want to discuss Meiteilon DPs that are similar to the simple examples in (1), but where a Genitive affix is (optionally) present on the upper demonstrative, like in (9) (see also Bhattacharya & Thangjam 2003):

(9)  

\textit{əsi(-gi) čiŋ-si}  
\textit{thisDEF+PROX(-Gen) mountain-DEF+PROX}  
‘This mountain.’

We argue tentatively that this type should be assimilated to the type in (6a).
We conclude that languages that are totally unrelated both genetically and contact-wise may still share basic observational traits that can be understood on the basis of underlying common structures. We briefly discuss some fruitful methodological consequences of this insight.

References
Koīc (English: Sunwar, Sunuvar, Sunuwar, Mukhia; Nepali: सुनुवार Sunuvār) has almost 38,000 speakers (Central Bureau of Statistics 2012: 164) and is one of the larger of the 123 languages of Nepal that are mentioned in the Census of 2011 (Central Bureau of Statistics 2012). Most speakers of Koīc are bilingual with Nepal's official language Nepali and those who are literate learned reading and writing Nepali using Devanagari in school. Koīc is traditionally not written but since the advent of democracy at the beginning of the 1990s some Koīc authors wrote and published poems and prose in their mother tongue. Writing and reading Koīc in Devanagari poses no problem for any literate speaker of the language but the standardisation of the spelling system is a recurrent topic discussed among members of the Koīc community in Kathmandu. The discussion centres about signs for those sounds of Koīc phonology that differ from any sound of Nepali phonology, for example the syllable final glottal stop occurring in Koīc but not in Nepali.

Koīc is not only written in Devanagari but has its own writing system, the Jeticha script. The Koīc community in Kathmandu hardly ever uses Jeticha. In neighbouring Sikkim (India), where English is the official language and Koīc (in Sikkim called Mukhia) one of ten additional official languages (Commissioner for Linguistic Minorities 2011: 143), Koīc school books and newspaper articles are published in the Jeticha script but not in Devanagari. In Sikkim, the Jeticha script seems to be preferred to Devanagari not because it is better adapted to Koīc phonology but due to local political developments. This presentation will show how Devanagari is adapted to the phonology of Koīc by different authors and discuss the sound and sign correspondences offered by the Jeticha script. A summary of the political developments in Sikkim that led to favouring Jeticha over Devanagari concludes the presentation.

References
Mukhia, R.B. 1990. „Manuscript“. Gangtok, Sikkim, India
Historical semiosis of Vedic words: The cases of *vajra* and *ātman*

Jens Braarvig  
University of Oslo

The connotations of the Vedic words *vajra* and *ātman* are numerous in the traditions growing out of the Veda, be they Hindu, Buddhist or others. This exemplifies how the meaning of the words in older Vedic contexts are blurred with later interpretations. The talk is an attempt to investigate the connotations of these words on the background of early contextual and comparative materials.
The lexicon of the “act of accepting (pratigraha)”: 
an approach to the multilayered Vedic culture
Maria Piera Candotti, University of Lausanne - Tiziana Pontillo, University of Cagliari

The present joint work finds its ground (and at the same time its scope) in a variation-oriented reading of the Vedic sources here applied to those which mention the act of the so-called “<gift-> acceptance” (pratigraha) within the mechanisms targeted on attaining and distributing the “goods of life” among all the members of a community. The most ancient occurrences are read and contrasted against the subsequent socio-ritual context where the well-known homonymous privilege and peculiar means of livelihood for Brāhmaṇs is depicted. The tentative interpretation of the relevant passages and the consequent reconstruction of the several layers of the Vedic lexicon (and corresponding texts) revolving around the verbal base prati-grah- might contribute in better assessing the existence of a specific Indo-Aryan cultural matrix, that might have pre-existed (and co-existed with) the mainstream Vedic world, and in better understanding how later knowledge systems succeeded in creating a new all-encompassing Śrauta balance. Ideally, what is proposed here is the prosecution of a lexical trend of research that both authors have pursued in the past years (sometimes also in collaboration with others such as dr. Moreno Dore, University of Cagliari, dr. Chiara Neri, University of Rome) tackling lexical issues such as brahmabhūta, yogakṣema and the Vrātyastoma terminology. The present paper shall focus, among others, on the following three crucial passages: AVŚ 3.29.7-8; JB 3.139; BŚŚ 18.24; 47.
The typology of pragmatic case marking in Tibeto-Burman
Alexander R. Coupe
arcoupe@ntu.edu.sg
Nanyang Technological University, Singapore

This paper provides a typological overview of a variable, non-paradigmatic case-marking pattern that is increasingly becoming recognized in modern descriptions of many Tibeto-Burman languages. Pragmatically-motivated core case marking is not a new development, but may have been overlooked in some older work due to the nature of the methodology employed to gather data: it is now known that analyses based on directly elicited data often fail to create the specific pragmatic contexts that motivate the use of core case marking in these languages. Elicited data may consequently produce regular paradigms that are not actually attested in narrated texts, or the contact language used to elicit translation equivalents might exert an adverse effect on the structure of the elicited data, leading to misconstrued interpretations of the case-marking pattern in the language of investigation.

Careful text-based analyses reveal that an increasing number of Tibeto-Burman languages lack a recognizable alignment conforming to an ergative-absolutive, nominative-accusative, active-stative or split-ergative marking pattern (e.g. Chelliah 1997; Coupe 2007, 2011; Hyslop 2011; Lidz 2011; Tournadre 1991 among others). Rather, case marking (or its absence) on a core argument can instead be motivated by a range of pragmatic factors that will be outlined in the paper. The fact that agentive marking variably appears on both the S and A arguments of monovalent and bivalent clauses demonstrates that it cannot be motivated by factors relating to transitivity, as might be expected of a language that has grammaticalized an alignment pattern based on clausal valency.

Discrimination via core case marking is most likely to appear in bivalent clauses with two participants when either one may fulfill the role of the agent. Under these circumstances, some grammatical means of distinguishing an A argument from an O argument may be required to clarify meaning. Conversely, if semantic roles can be unambiguously assigned to core arguments because of the semantic nature of their referents, then there may be no overt relational marking whatsoever. In some Tibeto-Burman languages we observe metaphorical extensions of this basic discriminatory function. Agentive marking is additionally used to contrast one referent from another, to encode increased agency or volitionality, to signal a shift in perspective, or to encode that a referent is behaving in an atypical or unexpected manner.

The recognition of pragmatically-determined case marking in emerging work has potentially important ramifications for syntactic theory, because such innovative uses of relational morphology could provide an initial pathway for the subsequent development of syntactic alignments and other aspects of grammatical complexity.

References


This paper explores the properties of oblique objects associated with experiencer subject verbs in Vedic. Recent research has shown that experiential arguments strongly tend to select the experiencer as their subject argument and that experiencer arguments tend to receive nominative case marking in Vedic (cf. e.g. Dahl 2014). Verbs with experiencer objects are comparatively rare in the language, and the majority of such predicates are built on morphologically derived causative verbs. Vedic experiential predicates thus tend to follow what we along the lines of Bossong (1998) may label a generalized constructional pattern, where the experiencer receives the same case marking as Agentive subjects, i.e. nominative case marking in accusative languages like Vedic. The morphosyntactic realization of stimulus object arguments, on the other hand, is much less uniform, as this argument type can receive accusative, dative, genitive, instrumental, locative or ablative case marking. Even though many psychological verbs consistently select one of these argument realization options, a number of verbs belonging to this group show alternation between two or more object case-marking patterns. An extreme case of object alternation is given in (1), where the verb MAD- ‘revel in, enjoy’ selects an object argument in the accusative (1a), genitive (1b), instrumental (1c) and locative (1d).

(1) a. sāman nú rāyé nidhimán nú ánnaṃ kárāmahe
   goods-ACC now wealth-DAT abundant-ACC now food-ACC prepare-1PL.PRS
   sū purudhā śrāvāṃsi
   well manifold-ACC.PL loud.praises-ACC
   tā no viśvāni jaritā mamattu
   these-ACC we-GEN all-ACC singer-NOM enjoy-3SG.PRF.IMP
   ‘We are now preparing goods, abundant food and manifold loud praises to attain wealth. Our singer shall enjoy all these things!’ (Rigveda X 59.2)

b. ayāṃ svādūr ihá mádiṣṭha āsa
   this-NOM sweet.potion-NOM here most.intoxicating-NOM be.3SG.PRF
   yāsya índro vrtrahātye mamāda
   which-GEN.SG Indra-NOM Vṛtra.slaying-LOC enjoy-3SG.PRF
   ‘This sweet potion here has been the most intoxicating which Indra has enjoyed at the Vṛtra-slaying’ (Rigveda VI 47.2)

c. dádo vāsūni mamādaś ca sómais
   give-2SG.PRS.SBJ goods-ACC enjoy-2SG.PRF.SBJ and soma-INS.PL
   ‘You will give (us) goods and enjoy soma’ (Rigveda VII 24.1)

d. sumnéṣu íd vo ántamā madema
   favours-LOC.PL indeed you-GEN.PL intimate.friend- NOM.PL enjoy-1PL.PRS.OPT
   ‘May we indeed enjoy your favors as intimate friends’ (Rigveda VI 52.14)

In current linguistic research, it is uncontroversial that argument alternations are usually associated with some semantic or pragmatic distinctions (cf. e.g. Beavers 2006, 2010; Dahl 2009, 2014b). Verhoeven (2010) notes that experiencer object verbs are heterogeneous with regard to agentivity and stativity and it is tempting to suggest that experiencer subject verbs may be expected to show analogous properties across languages. This paper pursues this idea, exploring whether alternations like those illustrated in (1) indeed reflect semantic or pragmatic distinctions arising from the lexical semantic properties of experiential predicates.

References:


Word-minimality vis-à-vis morphophonemics of genitives in NKB

Shyamal Das, Professor, Dept. of English, Tripura University, Suryamaninagar, Tripura 799022
lingphon16@gmail.com

The constraint of word-minimality stipulates that every grammatical word is a prosodic word: every content word must have a foot. Preferred foot type across languages is a disyllabic (mostly) trochee. Bimoraicity sneaks in in the absence of required segmental melody phonologically. The Noakhali Dialect of Bangla (now on NKB) metrics predominantly falls in line with this general preference. In it (a) scansion takes place left-to-right; (b) only one foot—a syllabic trochee is constructed aligned with the left edge; (c) all syllables are light (L) i.e. long vowels or diphthongs are not distinctive in terms of weight value (both monomoraic) and codas are not moraic. Corollary: no iterativity is practiced; quantity sensitivity/insensitivity is ruled out. Compounds are kept out of consideration here. Focusing only on genitive suffixation the paper explores the NKB mechanism/s to fulfill word minimum requirements: vowel epenthesis, phonetic lengthening of the root vowel; or converse, UR vowel deletion of the suffix.

- Monosyllabic CV and CVC words acquire prosodic word status through formation of moraic trochees via phonetic lengthening since coda is non-moraic and indifferent to metrification processes. Under suffixation of genitives, sequences with one or more syllables with final CVC take the suffix –ɛr/-ɛr: /han-ɛr/ ‘of beetle leaf’, /beʃ-ɛr/ ‘of cane’ etc.
- Thus a priori there exists an ‘allomorphic’ relation between the two SR forms of the genitive –ɛr/-ɛr and –r.
- Superficially, this general equation is driven by the cross-linguistic phonotactics of syllabification.
- However, things get problematized when one notices exceptions galore to these phonotactically determined rules.
- Why do CV bases take –ɛr/-ɛr (VC) as the suffix at the apparent cost of bimoraicity acquired through phonetic lengthening of the base vowel, and Maximal Onset Principle? How does the theory account for the presence of the immigrant mora? Is no lengthening deployed? Is the lengthened vowel curtailed to one-mora size? Or is it treated as extrasyllabic?
- Answer to these queries is found out to be as the following:
a) No phonetic lengthening is deployed once the suffixal vowel enters the scene with its own mora, and consequently, ensures bimoracity at the syllabic or/and moraic level.
b) In between a high and a low vowel (at the morpheme boundary) a glide [j] emerges fulfilling the phonetic requirement for an onset in the second (VC) syllable.

- The contention for –r as the UR suffix fails on the principal count that often native speakers do not prefer it in case of monosyllables with CV despite the process and result thereof being phonotactically amiable: bimarcity is ensured on moraic/syllabic (LL) count with –er/-er.

- If not totally discounted, -r is treated as optional coexisting with –er/-er. Further, in words like /gía-t/ ‘of cream color’, /sia-r/ ‘of musk rat’ the UR form of the genitive is –er/-er. It is clandestinely present with the non-high vowel –e/e- assimilated to second part of the diphthong in the base i.e. -a. A clinching and clear-cut piece of evidence is derivable in this respect from the optional co-existence of –r and –er/-er in /kia-r/ ~ /ki-er/ ‘of what’: -e supplants base a and asserts its moraic value.

So, the paper arrives at the following findings:

- NKB has an undominated constraint for syllabic trochee of (LL) type. If the base is monosyllabic phonetic lengthening is resorted to ensure moraic binarity for a trochee.
- Under affixation of the genitive –er/-er the incoming mora of the suffixal vowel is put into service as mora-provider, and simultaneously moraic lengthening of the base vowel is discarded/not required.
- –er/-er is the UR form of the genitive suffix, not –r. The latter is at the most an allomorph of the former.
- Word minimality condition is satisfied at every level through a disyllabic/bimoraic foot.
- As for morpho-phonemics, vowel lengthening at phonetic level is resorted to as and when needed. Neither vowel epenthesis nor deletion is attested. Vowel assimilation creeps in and ensures the clandestine presence of the suffixal vowel –e/e.

References

Translation and bilingual lexicography: the case of French to Urdu and Urdu to French, analysing some specialised bilingual glossaries

Alain Desoulieres
INALCO/CERLOM (HR)
Université Sorbonne Paris Cité
Visiting Fellow JNU, JMI New Delhi

After Dictionaries from pen to keyboard changing and unchanging patterns: the case of Urdu a paper on history of Urdu lexicography and its defining (changing) categories (lexical and syntactic analysis) presented at WSSANLP workshop within COLING Conference 2016, that took place at Osaka (11th December 2016) I wish to reflect on my present long task work, not yet published of a French Urdu Dictionary, with Linguistic and Literary Introduction, and bilingual specialised glossaries an essential dictionary, from six thousand French main entries translated into Urdu lexical equivalents and phrases (following my Petit Lexique Français-Ourdou an Elementary French Urdu Dictionary Désoulières Inalco 2008, 177 p. 21x29cm)
So I think it is time to tackle with a critical approach of the main use of such documents for French speaking students and researchers of Urdu, that is the need of building a working knowledge of Urdu essential and modern vocabulary for Urdu writing and expression purposes, but also to provide a quick guide to essential French for Urdu speaking students and researchers who have to translate from French. The use of reversing bilingual glossaries (lexical as well as semantic based) is also to be studied.
So my theme of reflexion is Translation and bilingual lexicography, the case of French to Urdu and Urdu to French, analysing some specialised bilingual glossaries. What strategies to use a limited corpus? The transition from transliteration to nastaliq Urdu script. The essential notions of Urdu lexicology and neology should be examined as well as North Indian lexical derivation in Urdu orthography and Urdu received pronunciation (also the case of Sanskritic lexemes and proper names). How to make it an efficient and predictive tool for translation? E.g. how to guess a neologism? Categories for Indo Persian lexicology in Urdu (including Urdu orthography and pronunciation of lexemes from Arabic through Persian and even from Modern Arabic) are also to be examined. Not excluding the role of English neology and Indo Persian calques of English technical terms and the question of English loan words. An essential monolingual bibliography and other sources should be suggested to the reader.

References

1. Main sources (lexicographic corpora):
Firoz Ud Din, Al Haj Maulavi, Firoz ul Lughat, new edition, Lahore s.d; c. 1995 and further edition 2005 (monolingual and etymological modern and classical Urdu dictionary) (over 150,000 words), and Educational Publication House, New Delhi
Khaweshgi, Muhammad Abdullah Khan (1937) Farhang e Amra (dictionary of Arabic and Persian loanwords in Urdu) 2nd ed. Islamabad 1989
Qureshi, Bashir Ahmad (n.d. c.1980) Kitabistan’s 20th-Century Standard Dictionary Urdu into English, second ed. s.d., Lahore and Delhi (50,000 words)

2. Related publications:
Dissociation of Case and Agreement in Santali

Mayuri Dilip and Rajesh Kumar
Indian Institute of Technology Madras

Agreement is a morpho-syntactic complexity in natural language, discussed in studies such as Saksena (1981), Comrie (1984), Chomsky (2000), Butt (2001), Bhatt (2005), Rau (2007), and Baker (2008), among others. The analysis of Agreement, has developed over the past few decades, and discovered certain significant aspects such as the Agreement information flows from noun to verb, specific loci in a tree structure that trigger Agreement, feature interaction between the noun and the verb, and spec-head configuration of the noun and the verb. Taking into account the previous analyses, we investigate Agreement in Santali, with the application of Baker’s parameters, and Bhatt’s operation AGREE. The investigation reveals a unique feature of Agreement in Santali, that is, Case, and Agreement are dissociated, when the arguments trigger object agreement. The evidence for such dissociation is the fact that the arguments triggering object agreement, show variation in Case. This can be contrasted with the arguments triggering subject Agreement, where, no such dissociation is found, and as a result, they show similarity in Case.

We investigate the Agreement structure of subject, direct object, indirect object as in (1); and Non-Nominative Subject (NNS) in constructions manifesting psychological states as in (2), by applying Baker’s parameters (Baker 2008) such as i. The Direction of Agreement: F agrees with DP/NP only if DP/NP asymmetrically c-commands F; ii. The Case-Dependency of Agreement: F agrees with DP/NP only if F values the case feature of DP/NP or vice versa.

1. baha$_i$ gidra-ko$_j$ arel-em lagid ema-d-ko(w)$_j$-a-ci
   Baha.SUBJ child-pl.DO Arel-to.IO give-[+tr]-OAM-[+fin]-SAM
   ‘Baha gave the children to Arel’.

2. baha$_i$ raba$h$ nam-aka-d-e(y)$_h$-a
   Baha.NNS cold.NOM.OBJ have-pst-[+fin]-OAM-[+fin]
   ‘Baha caught cold.’

Adopting Kidwai (2000), we postulate that the locus of the subject agreement is [spec, AgrsP] occurring above TP and the locus of the object agreement is [spec, AgroP] which occurs below vP and above VP. However, the functions of [spec, AgrsP] and [spec, AgroP] are further modified based on the evidence obtained by Santali. We state that [spec, AgrsP] is a locus of both subject agreement and nominative case, and the evidence comes from the one to one mapping of Agreement and Case, where the argument that depicts subject agreement, also manifests nominative case. [Spec, AgroP] is the locus of object agreement and not Case, and the evidence comes from the one to many mapping of Agreement and Case, where the arguments that trigger object agreement, do not manifest one particular type of Case. That is, NNS, DO and IO show differences in Case, however, they show similarity in object agreement. In such case, the internal arguments achieve Case at base-generated position. The head V assigns Case to one of the two specifier positions of vP and VP, based on the thematic hierarchy (T-hierarchy) such as experiencer > theme/patient > recipient > agent, where the accusative case is assigned to the argument based on its higher thematic ranking in the hierarchy. Since, the verb can assign case only once to one of the two specifier positions, the argument that remains Caseless moves to [spec, AgrsP] for Case and Agreement. The selection of one of the internal arguments to trigger object agreement involves the application of T-hierarchy, for example, experiencer outranks theme in psych-predicates, theme outranks
recipient in ditransitive sentences, theme outranks agent in transitive sentences, and agent is not considered at all for Case and object agreement in intransitive sentences. Since, the verb can assign Case only once, the internal argument that remains Caseless moves to [spec, AgrsP], which is a locus for default Case.

References

Baṅgaṇī ergativity: attempt of an in depth study

Andrea Drocco – University of Turin
(andrea.drocco@yahoo.it)

Baṅgaṇī, an Indo-Aryan language spoken in the so-called Baṅgaṇ area, a land enclosed by the Pabar and the Tons rivers (Uttarkasi district, Uttarakhaṇḍ state), shows in its grammar and lexicon some peculiar features (Zoller 1989) still rather controversial (cf. also Van Driem & Sharma 1996). The debate is still in course, due to the lack of enough documentation available, as the majority of scholars complains about (Abbi 1997; Cardona & Jain 2003; Hock (ed.) 2016). Moreover Baṅgaṇī, among the Western Pahāṛī languages of New Indo-Aryan, is now esteemed as a critically endangered language by the UNESCO Atlas of the World’s Languages in Danger. After a paper in print (Drocco 2016), the aim of my talk is to present the additional results of the analysis of i) the material collected through a fieldwork research with Baṅgaṇī mother-tongue speakers and ii) the unique elicited text available. In particular after having shown the description of Baṅgaṇī pronominal and nominal declension, focusing on the case marking and agreement system of the Subject-like and Object-like arguments of intransitive and transitive clauses (in perfective and non-perfective tenses), I will try to understand if in Baṅgaṇī the ergative case-marking is widespread also in some non-ergative clauses. Moreover the comparison between the good amount of data collected enables me to advance some hypothesis about the origin of the different Baṅgaṇī ergative case-markers.

References

A Linguistic Analyzer for Building a Database of Old Marathi

Oliver Hellwig, SFB 991, Düsseldorf University, Germany
Ashwini Deo, Department of Linguistics, Ohio State University, U.S.

Background and Motivation

The general scope of research to be addressed by this project is detecting changes in tense and aspect marking, and in case marking patterns of Old Marathi. Such changes cannot be studied simply by looking at printed books, because it is essential to see how different forms map onto similar semantic functions over time. Plain text corpora do not provide an authoritative understanding of the grammatical structure of forms. Therefore, we are working on a morphologically analyzed corpus — specifically to be able to search for instances of abstract grammatical and semantic categories rather than phonological strings alone. The presentation will introduce the corpus, its linguistic motivation, and the underlying technology from the perspective of Natural Language Processing.

Old Marathi is one of the few Indo-Aryan languages that offers a relatively continuous record of the evolution of New Indo-Aryan since 1280 CE, while, for instance, Hindi or Bengali do not provide such a comparatively homogeneous material. Old Marathi literature is special also in terms of having a substantial prose component due to the Mahānubhav literary tradition. Finally, the language exhibits several syntactic and semantic diachronic changes that can be studied with finer resolution through a database.

Specifically, patterns of variability in which multiple forms are used in a layered way to express similar meanings can be best studied if there is a diachronic structured corpus. The first step for this is getting a handle on the earliest available Marathi grammar which is the main scientific aim in this research project.

Building Resources

As is the case for other early NIA languages, there exist virtually no digital resources for the study of Old Marathi. The first step of our project thus consisted in adapting an existing OCR system for digitizing printed Old Marathi texts. Using this OCR, we have digitized the complete Līlācaritra and major parts of the Pañcopakhyān, manually corrected the OCRed texts, and transformed them into a structured XML representation. In addition, we were able to obtain a digital version of an Old Marathi dictionary (Tulpule et al., 2000), which forms the basis of our lexical database.

Technical Background

We have created a web-based application for annotating texts, and searching in annotated texts. The application combines a PHP interface with an analyzer written in Python, a MySQL database, which stores language specific data, and a corpus of source texts encoded in XML. We tried to keep the linguistic analyzer as flexible as possible by not integrating any linguistic rules that apply only to Old Marathi. In this way, it can easily be reused for the analysis of other MIA or NIA languages, given that a digital dictionary and reliable descriptions of phonetic and morphological processes are available.

The analyzer consists of three main components:
1. The **lexicon** stores lemmata and their orthographical variants, irregular nouns and verbal forms, and meanings for each lemma together with their sources in the secondary literature.

2. The **rule base** collects the inflectional classes, and the inflectional affixes that modify lemmata.

3. The **language model** stores smoothed probabilities of lexical and morphological n-grams extracted from the annotated part of the corpus.

Linguistic analysis of a sentence is performed in two steps. First, the analyzer generates morpho-lexical proposals for each string in the sentence by applying the default morphological rules stored in the rule base, and by searching for irregular forms. Compounds are also split in this step. In the second step, the analyzer searches for the optimal morpho-lexical path through the resulting trellis using the language model and a trigram-based dynamic programming approach (Brants, 2000).

**Next Steps**

Apart from increasing the size of the annotation database, we will focus on transferring linguistic knowledge from the huge (unstructured) digital corpora of modern Marathi to Old Marathi. We expect that the controlled transfer of linguistic structures will decrease the error rate in the analysis of Old Marathi. In addition, this step is also highly relevant for research in Natural Language Processing, where “non-standard” textual representations – including spoken and historical forms of language – currently receive increasing attention.

**References**


Multi-layer annotation of the Ṛgveda – Methods and Scholarly Applications

Oliver Hellwig, SFB 991, Düsseldorf University, Germany
Sven Sellmer, Institute of Oriental Studies, South Asia Unit, Adam Mickiewicz University, Poznań, Poland

The Ṛgveda is one of the oldest pieces of Indo-European poetry. The text has, most probably, been composed over several centuries by groups of authors, and reflects in parts an oral mode of literary production. Due to its high age, it is of central importance for Comparative Linguistics and Indological Studies. Therefore, a digital annotated version of the Ṛgveda is a strong scientific desideratum.

Our presentation reports about an ongoing project in which the Ṛgveda is annotated with stacked layers of linguistic information, including sentence boundaries, phonetic processes (Sandhi), morphology, lexicographic information, word semantics, and verb-argument structures. As our data come from heterogeneous sources – a scenario quite typical for Digital Humanities and Philology -, the presentation will put a strong focus on the question of how such data can be merged and represented.

The second part of the presentation will address some areas in which we are planning to use the annotated data, including statistical approaches to author identification of Vedic hymns and detection of formulaic language as an indicator of orality. We will sketch how such questions can be addressed by using existing Open Source software.
This paper describes a case of variation in Meeteilon. The data of variation in use is taken from Meeteilon(s) spoken in two neighbouring regions of Manipur i.e., Thoubal (T) and Imphal (L). The variation in question is not related to prestige, (non)-standard, age, sex or any other social markers because T and L are considered to the same standard variety of Meeteilon. The data below may be considered to be a case of Microparameter (Kayne, 2005) which looks for variations within the closely related languages.

1. (T & L)
madu tou-hái-de
That do-sanction-NEG
That is not sanctioned.

2. (T & L)
Tomba madu tou-khi-de
Tomba that do-DISTAL-NEG
Tomba didn’t do that.

1’. (T & *L)
naŋ madu tou-hái-nu
You that do-Yet-IMP+NEG
Don’t do that now/ immediately.

2’. (*T & L)
naŋ madu tou-khi-nu
You that do-Yet-IMP+NEG
Don’t do that now/ immediately.

1”. (T & *L)
ây madu tou-hái-roi
I that do-Yet-FUT+NEG
I won’t do that now/ immediately.

2”. (*T & L)
ây madu tou-khi-roi
I that do-Yet-FUT+NEG
I won’t do that now/ immediately.

3. (*T & *L)
tou-hái-nu
do-sanction-IMP+NEG

4. (T & L)
cat-Khi-roi
go-DISTAL-FUT+NEG
Won’t go away

(1) and (2) show us that the knowledge of the two morphemes (-hái- & -khi-) is shared between the speakers of T and L as permitted and past/distal respectively without any variation when they are followed by the morpheme –de (Negation). The variation arises between the morphemes in (1’ & 1”) and (2’ & 2”) when the morphemes are followed by –nu (Imperative Negation) and –roi (Volitional + Negation/ Future + Negation). L chooses –khi- and T chooses –hái- to derive the same expression. (the tonal differences in -hái & -hái- (permitted and immediately) and khi- & -khi- (past/distal and immediately) to be noted.) However, there are no constructions such as *tou-hái-hei- OR *tou-khi-khi- while speakers in T (not L) can use tou-khi-hái-nu and tou-khi-hái-roi. It indicates that the position of -hái- or -hái- will be higher than -khi- in the cartographic hierarchy of the numerous verbal morphemes of Meeteilon. (3) and (4) also indicate that -hái- cannot occur under the
IMP\RATIVE+NEG marker –nu while it can occur with FUT+NEG (i.e. tou-\hei-roi). Because –nu being an imperative morpheme, the speaker shows an authority towards the addressee while –\bai- marks a god/ society/ convention to be the inherent authority by default. So there is a conflict between the two available authorities (god/ convention vs. the speaker). When it comes to the question of T choosing –\bai- and L chooses –\kh\i- as the candidate for expressing immediately in the environment of negation discussed above, it is a common phenomenon among the other varieties of Meeteilon to have phonologically different forms of morphemes for a single grammatical function but the case (\bai & \kh\i) in T and L is challenging because of the closeness of the two comparing varieties, similarity of the tone (’), the usage of both (\bai & \kh\i) by both the varieties and lack of understanding about the features associated with different verbal morphemes of the language in general.

References


Out of India? A new look at the linguistic arguments and evidence

Hans Henrich Hock
University of Illinois

In a 1999 paper, titled “Out of India?”, I addressed some of the arguments presented by Indian nationalist authors for the assumption that the Indo-European languages (IE) originated in India and that, therefore, the so-called Aryan Invasion Theory (AIT) must be wrong and the relation between Indo-Aryan/Sanskrit and the other IE languages must be explained as resulting from movement Out of India (OIT). This paper takes a more detailed look at linguistic arguments presented by these authors, adds discussion of more recent publications, and examines how well the OIT proposals hold up.

Arguments against the AIT range from denying the validity of the Comparative Method (e.g. Jha & Rajaram 2000, Kalyanaraman 2009; see also Marcantonio 2009ab), to claims that arguments against an Indian origin of IE can be overcome by pointing to parallels with Romani (Misra 1992), arguments that Sanskrit features attributed to Indian substratum influence can be considered to be inherited (Talageri 1993), or attempts to refute Hock’s argument against the OIT (based on IE dialectology) by proposing an alternative scenario that roots the IE languages in India (Talageri 2008). Major focus will be on the claims and arguments of Marcantonio, Misra, and Talageri.

Marcantonio’s attempt to question Indo-European comparative-historical linguistics in its entirety rests on two major claims: The methodology of comparative-historical linguistics is circular, with sound changes invoked merely to rescue the method; and Verner’s Law, supposedly the foundation of comparative-historical linguistics, fails to predict correct outcomes. Both claims are dubious. Not invoking sound changes would fail to explain linguistic relationships and language history and moreover would violate the Uniformitarian Principle—all languages undergo sound changes in their observable history. Further, Marcantonio’s claim regarding Verner’s Law is shown to be just plain wrong. (It confuses absence of evidence, attributable to other changes, with counterevidence.)

As regards Misra’s claim, I add to Hock’s linguistic arguments that Romani does not provide a parallel with the supposed out-migration of IE languages. Arguments include a more detailed examination of Romani vowel changes as well as evidence that Romani preserves distinct reflexes of retroflexion — something not found in early IE outside the greater South Asian area.

This issue links with Talageri’s (1993) claim that the non-Indian IE languages simply lost the retroflex : dental distinction. The fact that words like mīḍha ‘recompense’ correspond to western IE forms with sibilant (e.g. Gk. misthos), and words like mīḍha ‘urinated’ to forms with velar (e.g. Lat. mictus) shows that this claim cannot be correct. Rather, the relationships between the words involve irreversible phonological mergers in the history Sanskrit, which fed into the development of retroflexion.

Talageri’s (2008) attempt to refute Hock’s (1999) questioning of the OIT in terms of PIE dialectology and migration, and to propose instead a set of migrations out of India, runs afoul of
Occam’s Razor: Hock operates with a simple, centrifugal movement of IE languages out of the Eurasian Steppes to their later geographical locations, whereas Talageri has to invoke a complex scenario of six different stages, with five migrations and reaffiliations, to account for (more or less) the same geographical distribution. Moreover, Talageri’s attempt is parasitic on Gamkrelidze and Ivanov’s (1995) proposal to locate the PIE homeland near the Caucasus, without justifying why an Indian root should be preferable to a Caucasus one.

I conclude with the observation that an “Into India” hypothesis better accords with the evidence of retroflexion and with IE dialectology than an Out-of-India one. This evidence must be added to the positive evidence of retroflexion: This feature is absent from the PIE ancestor of Indo-Aryan (and of neighboring languages); it is widespread in South Asia, from the Andaman Islands to Afghanistan; it is absent farther north and west; and therefore it cannot be explained other than through convergent developments within South Asia after migration of Indo-Aryan speakers to the area.

References


1. Kannada, is one of the Dravidian Languages, which is spoken in South India. There are number of dialects in Kannada Language. Many Proto-Dravidian features are retained in these dialects.

2. In Dravidian languages, by touching the tip of the tongue to the roof of the mouth retroflex sound is produced. The feature of retroflex according to Peter Ladefoged: “the tip of the tongue is curled up and back to touch the roof of the mouth behind the alveolar ridge. The symbols are a long $t$ /ʈ/. These sounds, called retroflex sounds, also occur in Hindi and Sindhi.”

retroflex approximant is produced when the tongue curls up, and not touching the roof of the mouth, but stays near the roof of the mouth or in the liquid state. Because of this liquid nature, different sounds are produced with little difference. Example: They are retroflex lateral approximant, retroflex voiced stop, retroflex voiceless stop.

3. Discussions are done regarding the properties of retroflex approximant, whether it is approximant or fricative.

Bh. Krishnamurti: The phoneme * $z$ (IPA= $\zeta$) which was apparently pronounced as a retroflex approximant ($|$ frictionless continuant).

Zvelebil: the retroflex * $r$ ($ɻ$) in the older stages of the literary languages was a kind of retroflex fricative.

I have considered this as retroflex approximant ($ɻ$) in Kannada.

4. The reflexes of retroflex approximant in Kannada, which are pointed out by scholars is as follows:


b. P.S.Subrahmanyam (1989): $ɻ$ (IPA: $ɭ$) was preserved until about the middle of the 10th century A.D. in that period, it changed to $|$ (IPA: $|$) in the intervocalic and final positions and to $r$ (IPA: $ɽ$) before later assimilated to the following consonant.

c. Zvelebil (1970): * $r$ is preserved as a phoneme till about the middle of the 10th cent. Later, it has been replaced in the prevocalic and word final position by $|$ or $n$. Since these developments are complementary in distribution, they may be regarded as a single development in the system.

The above Scholars have identified reflexes as: retroflex approximant ($ɻ$) > $ɭ$ and $r$, $n$. According to my observation, apart from these reflexes the retroflex approximant ($ɻ$) shows other changes also.

Retroflex Approximant ($ɻ$) > Retroflex Voiceless Stop ($ʈ$) > Retroflex Voiced stop ($ᶑ$)

E.g., $piṛ(ɻ)ukkai$ = Dung of sheep. Ma.$piṛukka$, Ka.$piṛke$, pike $Havyaka$ Ka. $piʈʈ$ e. Tu. $piʈʈe$

Te.$piḍuka.(DED)$

$nōɻ$ = Ka. to precede, take the lead. Kui. $nōka$ Kuwi. $nōkali$

The archaic form for $nōl$ is $nōɻ$ in Kannada.

Ta. $nōʈtam$, $nōkku$ = to see, look at, view

Ma. $nōʈtam$ = view

Ka. $nōdu$, $nōta$ = to look at

So $nōɻ > nōl > nōta > nōdu$

5. The retroflex approximant is the oldest form. And retroflex voiceless stop ($ʈ$) is the intermediatory stage, which is native in its origin. Changed to retroflex voiced stop ($ᶑ$). Examples found in inscriptions.

E.g., $beɖdale$ = rain fed agricultural land

$beʈʈina gaɖle$ = wet land
In the Sandhi, root form shows the original phoneme. Eg., kāɻ(forest) + kicchu(fire) = kāɻkicchu/ kāḍkicchu = Forestfire.

Because of its liquid state, retroflex approximant changes its form in gemination, while past tense marker added, Ꞵ changes to its suffix form. Eg., eɻ(raise) + tu(past marker) = eɻtu > eddu

But Grammarians of Kannada, Kēʃirāja (13th Cent), Dr. Shankarana Bhatt (2007) have the following opinion: retroflex voiced stop(ᶑ) changes to retroflex approximant(ɻ). But examples that I have quoted above show a different rule.

Kavirājamārga, the Oldest Kannada work on Poetics, while dictating the language of the poetry, insists that poetry should use the words which are in current usage. According to him Northern variety is suitable for this. Southern variety retains most archaic form. So can’t be used.

Northern Southern
Eg., kōḍuven (to join) kōɻpen
kōɻpen is the archaic form, it has changed to kōḍuven. This change is due to the influence of the Prakrit, where voiced consonants were introduced to Kannada from that language.

Eg., Ka. kumbaɻa (pumkin) > kūʂmaɳḍa
Ka. dāɻimbe (Pomagranate) > dāḍima

Here, retroflex Approximate Ꞵ (l) changed in to retroflex voiced stop (ɻ) because of the language contact.

The same phenomena happen in the Dravidian languages.

Ka. kudike (pot) - Ta. kuṭṭikkai
Ka. adike (arcanut) - Ta. Aṭṭikkai

The above research Ꞵ (native) > Ꞵ (native) > Ꞵ (borrowed) leads to many research possibilities. It leads to the probe as to why in Dravidian languages and in particular in the Kannada language, retroflex voiceless stop is less in the initial position of the word. This will be probed in detail in my paper.

References

Emeneau M.B. 1970. Dravidian Comparative Phonology, Annamalai University, Annamalainagar.
P.S. Subrahmanyam 1983. Dravidian Comparative Phonology, Annamalai University, Annamalainagar.
Language plays complex roles in society. The previous studies (Fisher 1958, Labov 1963, Labov 1965, Labov 2001, Trudgill 2002, Chambers 1993 among others) on language and society explain their interconnected, complementary interactions and, those studies were primarily focused on the variations in the language. Variation being the fundamental nature of languages, the question of personal and social identity navigated through language variation and, established that there is an interconnection between language variation and identity. In the South Asian context, Indian linguistic complexity presents a composite cluster of variation as described in Emeneau (1956).

This paper analyses the variation in language at the lexical level and how the lexical choice of the speaker(s) affects in shaping their identity. It obtains primary data from the lexicon of the Mappila dialect of Malayalam spoken by the members of Mappila (Muslim) community of Kerala. The variation in the lexical choice is analysed by collecting data from the speech samples of 15 minutes from four different age groups of Mappila dialect speakers. Various contexts were analysed and the frequency of borrowed words in each instance is calculated to reach a conclusion on how the variation is happening in the speech community. The paper shows how the lexical choice of the speakers could be socially motivated and involve in shaping and changing identities. Lexical items or vocabulary clearly signal the group identity and personal identity.

Mappila dialect of Malayalam was rich in their frequent use of borrowed words from Arabic, Persian and Urdu. Even though there was a trade link with the Arabs and it influenced the language, there was a deliberate attempt to show their identity as a Mappila community member, which was derived from the socio-political situation during those days. This made a clear variation between the Mappila dialect and other dialects of Malayalam at the surface level, which was motivated to create and establish the identity of a person as the member of Mappila community. For example, the following words were borrowed from Arabic and used extensively by the speakers of Mappila dialect even though there are Malayalam expressions for the same meaning.

i) /khalba/ ‘mind’ or ‘heart’
ii) /hima:r/ ‘donkey (used as a term to tease or abuse people)’
iii) /subahə/ ‘early morning’

Historically, these kinds of linguistic variation were highly motivated because of the socio-political factors and, intertwined with the historical facts about the origin and spread of Islamism in the region; people from the Mappila community highly motivated to project their identity as a Mappila because of the social insecurities they had to face before accepting that religion. Thus the deliberate inclusion of Arabic, Persian and Urdu words in their speech
helped in showing their identity. However, the socio-political situations and factors at the origin of Mappila community have been changed over a period of time. The social motivation for indicating their identity as a Mappila no longer exist and thus the frequency of borrowed words from Arabic, Persian and Urdu have been reduced from their speech. Apart from the religious terms, the borrowed words from these languages are very few at present. The words given in examples (i), (ii) and (iii) are almost replaced by words like /manasə/ ‘mind’, /kaɻət̪a/ ‘donkey’ and /pularçça/ ‘early morning’, which are used by the speakers of other dialects of Malayalam. The analysis is carried out by the changes in the language of the people according to their age and found to have significant variations between generations.

Similarly, in the Namboothiri dialect (Brahmin Dialect) of Malayalam is another instance in which the borrowings from other language creating the community identity associated with the speaker. Namboothiri dialect has borrowings from Sanskrit and the phonological system of the dialect is also highly Sanskritised by the use of aspirated plosives, more clusters and so on, leading to identify the speaker of this variety as a Namboothiri community member. The Hindi-Urdu controversy existing in India is another example of language variation associated with the identity of the speaker(s) and communal consciousness. Even though the structure of both the languages found to be similar in the linguistic analysis, they are considered to be different because Hindi is mostly claimed as the language of Hindus whereas Urdu is claimed as the language of Muslims.

This paper shows that the motivation for language variation highly depended on the socio-political factors, especially among the speakers of Mappila dialect and, manifested in their language at the lexical level and thereby helped in moulding their identity. Borrowings from Arabic, Persian and Urdu helped in shaping their identity as Mappilas because of the correlation between these languages and the Muslim religion. But the speakers of the community show a deliberate withdrawal of these borrowed words over a period of time according to the new socio-political scenario and thereby, result in an alteration in expressing their identity. The need of projecting a specific identity of an individual would vary according to the change in the socio-political scenario and a variation in language can shape the identity in order to go with the varying socio-political situation.

**References**


Emeneau, Murray B. "India as a Linguistic area." Language 32.1 (1956): 3-16.


Labov, W. 1965. Social Stratification of (r) in New York City Department Stores.


Different types of dictionaries present everyday tools for people who work with languages. Bilingual, monolingual, advanced learner’s, concise, with or without pictures, dictionaries have their own place in the life of a linguist as well as of a language learner. With the advent of technology, electronic dictionaries have also become important tool. However, their appearance raises the question of dictionary design (how many elements should be included in its design), authorship (one or several, known or anonymous, etc.) and accessibility (free or paid, for example, downloadable or not, etc.).

The research team wanted to learn more about dictionary usage habits of Croatian students who specialize in one of following Asian languages: Hindi, Korean, Japanese and Sanskrit. The ultimate goal was to study how many students would wish to be active in e-dictionary design as participators and also to study how well the vocabulary that is included in learning material corresponds with their interests and needs. The research team has started out its study of students enrolled in at least one of aforementioned Asian languages in Croatia with following assumptions:

1. students prefer usage of on-line dictionaries due to accessibility and speed
2. students prefer usage of on-line dictionaries which offer translation and search options in their mother-tongue
3. students participate actively in the creation of on-line dictionary once they are given a chance
4. students’ participation enhances quality of on-line dictionaries aimed at learners’ as it helps create a base of words relevant for students’ need.

To test our thesis the research team has created a survey for students before they were given access to specially designed on-line dictionary that includes languages they study with Croatian translation as well as the option to search for a particular lemma through Croatian language or one of mentioned Asian languages. To make dictionary entries relevant to student population, team has analyzed learning material used in classrooms and created entries that represent core of the on-line dictionary for each language pair. That ensures students’ readiness to try the new dictionary as well as detection of lemmas that admin could add to the on-line base once the students suggest them with or without description. In that sense, the team predicts that after longer usage, e-dictionary will reflect vocabulary needs of students in a particular program. The paper will describe the process of lemmas’ inclusion, description and results of students’ activity in several months from dictionary’s activation in March 2017.

References:
Recognition and multi-layered analysis of converbs in early NIA
Rafał Jaworski and Krzysztof Stroński
(Adam Mickiewicz University)

The aim of the present paper is twofold: firstly we are going to demonstrate how computational tools such as Ordinary Least Squares (OLS) Regression and Latent Dirichlet Allocation facilitate analysis of untagged, digitalized early NIA corpora and secondly we are going to present results of preliminary multi-layered analysis of converbs in one early NIA variety, namely early Awadhi.

Converbs are non-finite verb forms marking adverbial subordination (Haspelmath 1995). In the context of Indo-Aryan languages they occupy a special place since they constitute a pan-Indian feature defining so called South Asian linguistic area (cf. Masica, 1976). Therefore converbs have been investigated from various perspectives and with regard to various languages, (cf. Dwarkesh 1971; Tikkanen 1987 and Subbarao 2012 among others). Even though the scope of research on converbs has been quite large we still do not have corpus-oriented studies dealing with early NIA dialects which show transitional stage in the development of morphosyntactic alignment. Since there is a scarcity of tagged early NIA texts, and digitalized NIA texts are becoming more accessible it is a necessary step to facilitate the analysis of select dialectal corpora by specialized computational tools. This analysis is possible thanks to recent advances in the process of text digitisation (Seljan et al. 2013) and the application of computational, mathematical, machine learning methods. The methods described in the present paper have been employed on the basis of a prior linguistic analysis of one of early Awadhi but further on the research will be carried out on other dialects such as Braj and Pahari. Firstly, 5000 word early Awadhi corpus comprising Jayasi’s ‘Padmāvat’ verse text has been tagged by means of IATagger (Jaworski 2015) with regard to morphological, syntactic, semantic and pragmatic levels. The tagged text has become the basis for the automatic recognition of converbal forms in the untagged portion of the text.

The proposed mechanism capable of fully automatic converb recognition relies on the concept of supervised machine learning. In this scenario, the input data consists of the aforementioned 5000 words labelled as converbs or non-converbs. In the first step of processing, the so called statistical features are extracted for each word. These features include: suffix (last 2 letters of the word), prefix, information if the word is in the first or last position in the sentence, previous word in the sentence, next word in the sentence and others. Then, by the means of OLS Regression, Latent Dirichlet Allocation and other statistical methods, a mathematical model is constructed. This construction is performed with the help of a robust, state-of-art Vowpal Wabbit tool (Agarwal et al. 2014). The resulting model is able to calculate the probability of a word being a converb on condition that, e.g. its suffix is “āi”, it is located in the middle of the sentence and the previous word is “lobha”. This model can be applied to an unannotated portion of the text in order to automatically extract words for which a high probability of being a converb is calculated. Before running this tool on the unannotated text, however, we evaluated it on the annotated data by the means of cross-validation. The tests revealed that the mechanism is able to identify 66% of all converbs in a text.

The application of the supervised machine learning tools on the unannotated data resulted in a corpus of untagged converbal chains which have then been tagged semi-automatically. Multi-level tagging enabled analysis of main argument marking in converbal chains with respect to animacy hierarchy, definiteness, control properties of converbs and pragmatics of converbal chains, (for similar analysis of early Rajasthani cf. Stroński and Tokaj 2015).

In the present paper we are going to focus exclusively on morpho-syntactic and semantic properties of converbs, i.e. differential subject marking (DSM) in converbal chain contructions and syntactic lability of converbs. Preliminary research shows that:
DSM has primarily a grammatical motivation, i.e. it is dependent on the transitivity of the main verb, but secondary factors such as animacy and definiteness also contribute to the variation in A marking (cf. Khokhlova 2006 for Rajasthani; cf. also Verbeke 2013 for more general discussion);

there are attestations of the violation of the "same subject constraint" which are predominantly semantically motivated (similarly to earlier stages of IA cf. Tikkanen 1987 or to other languages of the region cf. Tikkanen 1995); converbs can then form a kind of an absolute construction which is formally very similar to the one based on the participle;

We do think that the proposed multi-layered analysis, even though based on a quite homogenous corpus could give a broader picture of non-finites and also give some implications of a more general character. Then, extrapolating results of such analysis to untagged digitalized corpora is our desideratum and tools presented in the paper have been designed to facilitate this process.

References


Polysemy and cognitive linguistics. A case of vána

Joanna Jurewicz

University of Warsaw

The aim of this paper is to address the problem of polysemy of the Sanskrit words on the example of the meanings of the word vána used in the Rgveda („a tree, wood, forest, fire drill, vessel for Soma, water and material of the world”). I will show that methodology of cognitive linguistics is very useful to analyse the rational background of polysemy and its conceptual consistency. The basis for my analysis are three assumptions accepted by cognitive linguistics: 1. the meaning of words reflects thinking about the designate, 2. thinking is motivated by experience and cultural beliefs, 3. the associations between semantic aspects of the word can be modeled as conceptual metonymy, conceptual metaphor and conceptual blending. On the basis of these assumptions, I will reconstruct the semantic structure of the word vána. It is a radial category the centre of which is constituted by its most literal meaning which are tree and its metonymic extensions, i.e. wood and forest. The meanings of things made of wood (i.e. fire drill and vessel) are also close to the central meaning and are metonymic extensions. The meanings of water and the material of the world are the metaphoric extensions of the central meaning and more peripheral. They are based on cultural beliefs and models shared by the Rgvedic poets. I will also argue that the Rgvedic poets consciously shaped the semantics of the word vána by using it in contexts which forced the recipient to activate its less literal meanings. Thus they could create a general concept of the hiding place of desirable goods, such as fire, Soma, the sun, and the world.

References
Vedic texts didn't exist in a vacuum, but enjoyed specific uses in society; the texts were often reworked and repurposed to reflect the different needs of reciters across time and space, and scarcely anywhere is this more visible than in the lexicon. A major source of variation between Vedic recensions is the substitution of certain words and phrases to adapt an old hymn to new uses; in my talk, I will look at an especially conspicuous example of this, namely Atharvaveda Paippalāda 18.1–14 (i.e. the first two *anuvākas* of book 18), of which I've prepared a new critical edition. This hymn recycles the text of R. V 10.85—Sūryā's bridal hymn—into a new composition optimized for use by the Paippalādins, in line with their priestly functions and aspirations. Their textual substitutions not only leave clues to their period and location, but also reveal their political ambitions and the role they wished to assume in Vedic society. The mechanisms by which they adapt this lofty narrative of celestial myth for ritual praxis play a pivotal role in Vedic textual history and transmission.
Is ‘English supply’ the only route to livelihoods? Language and work in a multilingual economy
National Institute of Oriental Languages and Civilisations, Paris, France

Language is a key factor in questions of educational access and socio-economic mobility in multilingual societies. However, studies on these issues have largely focused on linguistic minorities, and little on the developmental stakes of a more equitable terrain for widely-spoken but economically marginalised languages - a typical feature of post-colonial societies.

The striking divide in India between English and the major regional languages (covering 95% of the Indian population) in the quality and scope of education and training has a well-recognized impact both on individuals’ economic opportunities and outcomes. Should this language divide be palliated through equalizing policies such as ‘English for All’ (EFA) or by promoting linguistic rights for marginalized ‘vernacular’ educated groups?

In a context of rising expectations, acquisition of English at any cost is perceived as a sine qua non for participation in economic growth, cutting across all social classes. This has led to widespread moves towards EFA on the part of local governments all over India, responding to public demand, taking the form of the abandonment of instruction in local languages in favor of ‘English medium’. A recent instance was the electoral promise of ‘KG to PG’ English medium (PG meaning Post-graduation) by the Telangana government in the 2014 parliamentary elections - a policy currently being applied, presaging the disappearance of Telugu from schooling. On the other hand, a typical response on the part of ‘vernacular’ educated groups is to demand special quotas in government employment and other rights to palliate their perceived disadvantage.

Surprisingly, the alternative angle of countering the imperatives of Englishization with policies for enhancing the economic salience of the ‘vernaculars’ to narrow the opportunity divide with English, has hitherto received little explicit attention from policy makers or researchers. More importantly, the widespread perception of the ‘English advantage’ by both individuals and policy makers simply ignores the fact of the economic weight of Indian languages and the actual linguistic requirements of work in different sectors and levels of the economy, between agriculture, industry and services.

In view of the Englishization pressure, we propose that language policies could be developed in relation to human capital formation to foster the creation of livelihood paths that enable social mobility for vernacular-educated populations. A well developed political economy of language could well serve to put language forward as a primary variable in economic planning, rather than as an after-thought.

In view of this, we propose some tools and concepts to approach linkages between language, economic activity and livelihoods that we may qualify as the ‘linguistic structure of the economy’. In particular, we envision the description of the multilingual economy as a composition of ‘vernacular economies’, each of whose size may be approached with a modified version of the concept of ‘Gross Language Product’, broached by D.Graddol (1995). Further, we consider the notion of ‘linguistic mapping’ of a local economy, that is a representation of the distribution of activities in an economy by language.

We examine these concepts in the context of the two Telugu-speaking states of south India (which include also Urdu speakers and a smaller presence of other Indian languages). Within
a project of description of the linkages between language, education, livelihoods and economic sectors we present the results of an exploratory survey of language use in the pharmaceutical sector in the city of Hyderabad, which raise questions on the impact of the linguistic divide with English evoked above on communicational, learning and performance inefficiencies in production. Measuring the impact of language-linked inefficiencies on productivity could have a bearing on highlighting the importance of the language factor in economic growth.

References
Syntactic Variation: A comparison between Meeteilon and Non-Meeteilon speakers
Alfina Khaidem
CASL, Delhi University

The paper discusses how variation can be analysed using Exoskeletal framework where the exoskeletal theory states that the syntactic structure with abstract syntactico-semantic properties containing feature matrices are generated independently of the concrete lexical and functional items, which are inserted into those structures as exponents at the later stage in the derivation (late insertion). The main focus of this paper is on the variation of question particle, -ra used as wh-particle and nominaliser -Nai in Non-Meeteilon variety in comparison with the standard Dialect, Meeteilon, a Tibeto-Burman language, the official language spoken in a state called Manipur, in North eastern part of India. The following are some of the examples of variation in wh-question and nominaliser:

<table>
<thead>
<tr>
<th></th>
<th>Meeteilon (Standard)</th>
<th>Non-Meeteilon</th>
</tr>
</thead>
<tbody>
<tr>
<td>YES/NO question</td>
<td>cak ca-ra-bə-ra</td>
<td>cak ca-ra-bə-ra</td>
</tr>
<tr>
<td></td>
<td>rice eat-PERF-NZR-YES/NO</td>
<td>rice eat-PERF-NZR-YES/NO</td>
</tr>
<tr>
<td></td>
<td>Have you eaten rice</td>
<td>Have you eaten rice</td>
</tr>
<tr>
<td>Wh-question formation</td>
<td>kəri ca-ri-no/-ge</td>
<td>kəri ca-ri-(bə)-ra</td>
</tr>
<tr>
<td></td>
<td>what eat-ASP-WH</td>
<td>what eat-ASP-WH</td>
</tr>
<tr>
<td>Nominaliser</td>
<td>su-na-bə</td>
<td>su-niN-Nai</td>
</tr>
<tr>
<td></td>
<td>wash-PUR-NZR</td>
<td>wash-wish-NZR</td>
</tr>
<tr>
<td></td>
<td>in order to wash</td>
<td>in order to wash</td>
</tr>
</tbody>
</table>

Meeteilon is a Pro-drop language. The language has two wh-question particles –ge–no. The variation can be seen between Meeteilon and Non-Meeteilon speaker variety where yes/no question particle -ra is used throughout in question formation. And, the nominaliser -Nai with -niN- is used in adjunct clause but ungrammatical for standard Meeteilon. Some examples of wh-question in standard Meeteilon are given below:

1. naN kəri tau-ri-no/-ge/*ra
   you what do-PROG-Q
   What are you doing?

2. naN kəri-gi lak-i-no/-ge/*ra
   you what-GEN come-IND-Q
   Why are you coming?

The –ra particle is ungrammatical with the wh-question formation in standard Meeteilon. The morphemes -niN-Nai cannot be used in the eventive reading in Standard Meeteilon but is possible in Non-meeteilon variety.

3. caobi [phi su-na-bə ] cət-khre (Meeteilon)
   Chaobi cloth wash-PUR-NZR go-PERF
   Chaobi left inorder to / to wash clothes.
4. caobi [ phi su-niN-Nai ] cət-khre (Non-Meeteilon)

Chaobi cloth wash-wish-NZR go-PERF

Chaobi left inorder to / to wash clothes.

In addition to this, the reading with -niN-Nai can be used as infinitival complement in non-Meeteilon speakers as shown in the examples below.

5. ai [kaithel cət-niN -Nai ] hot na-i/ tou-ri (infinitival)

I market go-wish- NZR try-NON-FUT/do-PROG

I tried to go to the market/ I am getting ready to go to the market.

Non-meeteilon speakers use -niN-Nai in all the adjuncts as well as in the infinitival complement clause. In non-Meeteilon speaker, syntactically, -Nai- is a nominalizer because it can occur in adjunct context as well as infinitival complement unlike standard Meeteilon.

The hypothesis is that the structure contains feature matrices only for the functional projection. Here, the question particle –ra is a functional Morpheme which shows syncretism between yes/no and wh- question structure in Non-Meeteilon speaker, which is also explained in terms of Vocabulary Insertion of Distributed Morphology where an operation takes place in the PF component of the grammar where one feature is deleted in a given context, supporting the theory that variation occurs at the PF component.

The next section is to consider that adjunct is a variable structure and not part of the functional structure which is constant in Exoskeletal Model and so variation can happen at adjunct frame which is an open slot. The discussion is focused on the adjunct structure discussed above in (3) and (4). The parametric variation is studied using the Exoskeletal Model from Borer (2005a) and Áfarli (2015a). Since, inflection decides the structure in this model, the paper attempts to establish the structures of the YES/NO and WH-questions between the two dialects as well as the structure of adjunct based on the features of the derivational morpheme, -Nai and -bə. Apart from this, the feature of -niN- is kept into consideration on how it combines with -Nai.

Since, the discussion is all about the inflectional morpheme, the question particle and derivational morpheme i.e., nominaliser, the paper analyses how variation takes place at the functional morphemes typically linked with the fixed extended projection corresponding to the abstract features bundles where one of the variants, the exponent matches with the superset features is inserted and realised. Secondly derivational morphemes like –ation, able etc. are considered as categoriser which defines the category of the root. Roots make up the open class or “lexical” vocabulary which are categorised first and inserted into the frame. So, in Meeteilon –bə is a derivational morpheme which has –Nai as its variant and considered to be a categoriser supporting Embick (2015).

References


The concept of resultative constructions as a linguistic category has long been quite diffuse. However, mainly due to the work of Nedjalkov ([1983] 1988), Dahl (1985), and Bybee et al. (1994), it had been possible to meaningfully distinguish resultative constructions from perfects, although the two categories do overlap in many languages. Resultatives, in contrast to anteriors or ‘perfects’, are defined as states that ‘exist as a result of a past action’ (Bybee et al., p.54). Resultative constructions in New Indo-Aryan have been studied particularly by Khokhlova (Khokhlova & Singh 2003, Khokhlova 2008).

In Hindi-Urdu, both anteriors (perfect, pluperfect etc.) and resultatives (present, past) are built on the second, or anterior, participle. The Old Indo-Aryan predecessor of this form (the ta/na-participle), when used as a finite verb (both on itself or in combination with an auxiliary), typically denoted resultative states, anteriors being expressed by other tenses. During the Middle Indo-Aryan phase, this participle took over the function of both a simple past and a present perfect, but at the same time continued to express resultative states. This ambiguity is still retained in Hindi-Urdu:

Consider:

(1) (a) वह बैठी है।
She has sat down. (anterior, present perfect)
She sits. She is sitting. (resultative, present)

(b) वह बैठी हुई है।
She sits. She is sitting. (resultative, present)

The combination of the simple second participle with the auxiliary है ‘is’ in (1a) can express both a present perfect and a resultative state. On the other hand, the marked form in (1b) – the participle with the adjectival marker हुआ in combination with the auxiliary है ‘is’ – is only compatible with a resultative reading.

The function of the adjectival marker can be analysed as blocking a finite, i.e. simple anterior reading of the participle and restricting it to its participial, i.e. stative or resultative value.

Hindi-Urdu distinguishes three basic resultative constructions, all of which may take the adjectival marker हुआ with the participle:
— one with intransitive verbs, the verb aligning with the subject (2a);
— one with transitive verbs without an agent, the verb aligning with the direct object (2b);
— one with transitive verbs with explicite agent, the auxiliary aligning with the agent and the participle taking the adverbial form in -ए (2c):

(2) (a) उसकी आवाज़ रुधी (हुई) थी।
His/her voice was choked.

(b) आंगन में नीले रंग से भरा एक नांद रखा (हुआ) था।
In the courtyard, there was put a vat filled with blue colour.

(c) तिया उसके साथ-साथ चली थी, लेकिन न तो उनका हाथ पकड़े होती।
Tia would come along with him, but she would not hold his hand.
The construction in (2c) is particularly remarkable for its non-ergative structure, as against the regular anterior tenses that are ergative with transitive verbs. On the other hand, other than Standard Hindi-Urdu, modern Delhi Hindi has developed an ergative resultative construction, obviously under Punjabi influence:

(3) कामिनी ने लोग तो देखे हुए थे।
Kamini had seen = knows people. (Jagannathan 1984: 425)

The paper will discuss this issue in detail and with constant reference to the types of resultative contructions in Standard Punjabi.

Besides standard resultative constructions, Hindi-Urdu has also structures of the type labelled Continual Resultatives by Nedjalkov (1988: 57), using रहना ‘remain’ and रखना ‘keep’ as light verbs. These structures hardly ever use the adjectival marker हुआ:

(4) (a) वह बैठा रहता है।
He remains seated.
(b) ज़फ़र झामोश से फ़र्श पर नज़र जमाए रहा।
Zafar kept fixing his glance on the floor.
(c) ड्राइवर ने बस को रोके (हुए) रखा।
The driver held the bus = The driver waited with the bus.

The adjectival marker हुआ in standard resultative constructions in Hindi-Urdu has an interesting parallel in Nepali. Here, we find two sets of anterior tenses, one built on the adverbial form of the short participle (गरे gare), the other one built on the extended form of the participle (गरेको gare-ko), which, apparently, under Tibeto-Burman influence contains the genitive marker को ko as adjectival marker.

If time permits, I shall also address combinations of the ta/na-participle with the adjectival marker sat in ‘late’ Sanskrit (Vetālapañcaviṃśatikā of Śivadāsa) reflecting clearly the New Indo-Aryan structures.

REFERENCES

The quotative marker in Telugu, ani, identical in form to the absolutive/past part. of the verb anu ‘to say,’ is claimed to introduce not only direct/quoted speech but also indirect speech translated variously as ‘that’ clauses or ECM clauses (Krishnamurti & Gwynn 1985; Bossé & Bossé 1990.). Crucially, ani clauses are the sole (apparent) exception to the constraint on multiple finite clauses within a sentence in Telugu, normally limited to one – the matrix clause. In this paper, I argue that ani is syntactically a true quotative marker and that the surface CP-type structure, while seemingly parallel to other embedded participial clauses, masks a unique underlying structure which gives rise to its different behavior. I first provide evidence that ani is distinct syntactically from the homophonous absolutive with examples containing matrix clause Force operators (interrogative, imperative) as well as the presence of p/n/g-marked verb forms which are, in all other contexts, ungrammatical outside of a matrix clause.

Evidence for ani as a lexically-distinct functional (quotative) marker comes from the following: 1) it can be omitted, whereas the absolutive cannot (1-2); 2) it can be immediately followed by any other verb of speaking’s absolutive form, whereas other verbs of speaking cannot be so combined (3); and 3) it allows within its domain a p/n/g-marked verb, unlike absolutive markers or any other embedded clause heads (contrast (4) with (7)).

(1) nuvvu na:to: ippudːe: vellːu (ani) ceppaːvu
   2sg-nom 1sg-inst now-emph go-imp (quot) tell-past-2sg
   ‘You said “Go now!”’

(2) pro reɳɖu maːtalu *(ceppi) Sridhar iŋtːiki vellːaːdu
   pro two word-pl tell-abs Sridhar house-dat go-past-3sg
   ‘After saying 2 words, Sridhar went home./Sridhar said two words and went home.’

(3) Sridhar iːni miːni maini moː ani ceppi (*ceppi ani) pro
    sridhar-nom eney meeny miney mo quot. tell-abs (*tell-abs say-abs) pro
    iŋtːiki vellːaːdu
    hous-dat go-past-3sg
    ‘After saying eney...mo, Sridhar went home./Sridhar said eney...mo and went home.’

(4) Sridhar iŋtːiki vellːi/*vellːaːdu pro bhoːjanam tinaːdu
    sridhar-nom house-dat go-abs/*go-past-3sg pro dinner eat-past-3sg
    ‘Sridar went home and ate dinner.’

To account for these phenomena, I propose that ani clauses are merged only as the DP/NP argument of the matrix verb with a phonologically null (D/N) head. As such, they are not clausal arguments per se and therefore do not pattern with all other non-matrix, non-finite clauses.

In addition, material contained within the ani clause is not subject to grammaticality constraints, unlike material in all other absolutive (and embedded) clauses. Direct/quoted speech may range from nonsense words, to onomatopoetic sounds, to word salad – it is entirely outside any domain of
grammaticality. (See Aikenwald & Dixon 2011 and citations therein for comparable cross-linguistic data and the many similarities of Vedic Sanksrit *iti* clauses.). Conversely, indirect speech, introduced by ‘that’ in English inter alia, must conform to the syntax or else produce an ungrammatical sentence. (5) illustrates these contrasts.

(5) Nenu i:ni mi:ni maini mo:/telugu ist franse ani ñeppa:nu 1sg-nom eeny meeny miney mo/Telugu ist francais quot. say-past-1sg ‘I said (*that) “eeny meeny miney mo/Telugu ist francais”’

In addition, ‘that’ clauses and indirect question (‘whether/if/etc’) clauses are ungrammatical with matrix-domain Force elements such as interrogative and imperative markers (see Rizzi (1997), for example, for matrix CP elements). In contrast, *ani* clauses are grammatical with either of these. [(6) is Krishnamurti and Gwynn’s (1985:364)]


(7) nuvvu na:to: re:pu ramm ani ñeppa:vu 2sg-nom 1sg-instr tomorrow come-imp quot. tell-past-2sg ‘You said to me “Come tomorrow!”’

Finally, several referential anomalies seem to offer further support e.g., an unexpected mismatch between subject and verb agreement in the *ani* domain where a 3sg pronoun (unmarked for deixis but not reflexive, see Kissock 1995), *tanu*, is matched with 1sg. p/n/g on the verb (though the expected 1sg subject pronoun is also grammatical under the same reading).

Taken together, the above data suggest that *ani* clauses are syntactically distinct from both matrix clauses and from other, morphologically similar participial clauses – the former because *ani* introduces a domain opaque to the syntax and the latter because the presence of p/n/g agreement does not result in ungrammaticality. I argue that these two distinctive syntactic behaviors are due, respectively, to the quotative semantic features of *ani* and to the fact that the *ani* clause, while itself a CP, is complement to an NP that satisfies the thematic argument requirement of the matrix verb.

**References**


In this paper we propose a ‘Feature interactive model’ (FIM) based on a set of interactive semantic features for translating the cuisine verbs of Telugu and Bangla, the two widely spoken Indian languages belonging to Dravidian and Indo-Aryan language families respectively. The need to develop large scale ‘feature data bases’ across various semantic domains is motivated by the fact that they play a crucial role in various natural language processing tasks like word sense disambiguation, information extraction and retrieval, machine translation etc. The ever expanding computational activities, the opening up of global and local communities facilitated by internet, reduced cost, new markets and due to several other factors research in machine translation is growing at a rapid rate. A lot of research in Indian languages is aimed at developing grammatical resources like morphological analyzers, parts of speech taggers, parsers and lexical resources like Wordnet etc.

The present work proposes that the existing dictionaries used in NLP will have to be annotated for semantic features and each lexeme be linked with the appropriate semantic label. E.g., boil will have to be tagged for the features such as { [+Heat], [Medium=±Liquid]…} etc. This approach yields more accurate word level translations.

A semantic domain approach would also give less scope for ambiguity due to the lexical coherence that is visible among the words belonging to a particular semantic domain.

The feature-based approach involves the identification and listing of features, and is by itself challenging as both deterministic and delimiting criteria need to be evolved. Feature matrices (see Table 1) are of immense help in tasks like translation, where cross-linguistic mapping of the maximum number of features will assist in selecting an appropriate equivalent. The main objective of the paper is to unify the established semantic theories like componential analysis, semantic domains and the implementation of lambda calculus.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Telugu</th>
<th>Bangla</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Intense [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Medium [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Low [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flame</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. High [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Sim [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Low [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Soft [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. crisp [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. dry [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. Evaporated [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. Melted [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Null [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Liquid [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Steam/vapour [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liquid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Water [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Oil/Fat [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Viscosity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Flat [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. deep [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Cook [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Fry [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Boil [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. simmer [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cooking Time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Long [1/0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Short [1/0]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1
It is also a fact that the terms in a semantic domain enter into relations with each other both on syntagmatic and paradigmatic axis and can be defined as synonymy, antonymy, hyponymy and so on. The most general term is ranked highest in the hierarchy and as the related words become more and more specific, they are placed at lower levels. However, certain related words are also capable of being substituted across levels. In this regard it may also be observed that the number of levels crossed to replace a word, have to be as minimum as possible to achieve a plausible replacement in the languages. The following hierarchical tree in Figure 1 (for Telugu) illustrates the arrangement specified above.

![Figure 1](image_url)

Figure 1

Telugu: *kaacu* ‘to boil (liquids)’ cannot be replaced by *vaMDu/vaMTa ceyyi* ‘to cook’ as *kaacu* at it is a specific cooking process associated with liquids which includes a general process of cooking, *uDuku* ‘boiling (both solids and liquids)’ which is a general process which is situated in between *kaacu* and *vaMDu/vaMTa ceyyi* in the hierarchy.

The present research (work in progress) also attempts to implement this approach in a Natural Language User Interface (NLUI) being developed and tested. The currently utility of the model can be attributed to its scalability to other semantic domains as well. This would also account for new terms entering a semantic domain and thereby enrich the existing databases.

A semantic domain approach combined with feature databases and lambda calculus would be an ideal approach for lexical transfer in NLP.

**References:**


This paper proposes a Dynamic Syntax (Cann et al. 2005, Kempson et al. 2001, 2011) account of Hindi clause structure and seeks to explain the interplay between word order, agreement and case within the dynamics of left-to-right parsing in context. Under the dynamic approach, syntax is seen as the particular steps taken in the incremental process of accumulation of semantic information. This is formally represented with gradually unfolding semantically transparent tree structures as information is compiled in a strictly word by word fashion.

The dynamic approach to Hindi clause structure differs significantly from more static syntactic approaches in several respects. First, the verb is not taken to be the sole projector of predicate-argument structure and thus case is not seen as ‘selected’ by the verb but as providing information in narrowing down the relation the marked NP holds with respect to other expressions in the clause ahead of parsing the verb (Kiaer 2007). Second, no correlation between case and agreement is assumed, i.e. information from case markers and information from verbal agreement morphology contributes independently to utterance interpretation. Third, contextual information (which includes the immediate linguistic, as well as extra-linguistic context) can inform processing choices.

What emerges as key for the analysis of head-final structures such as Hindi is the role of case as resolving structural underspecification in the strictly incremental process of establishing an interpretation for a string of words. Moreover, parsing a case marker indicates completion and results in the impossibility for further development of a node decorated with information projected by a case-marked expression. This is reflected in the clitic-like behaviour of Hindi case markers (Sharma 1999), as well as the impossibility for case-marked discontinuous NPs where the head precedes the modifier (Fanselow and Féry 2006).

References

Kempson, R., E. Gregoromichelaki and C. Howes (eds.) 2011. The Dynamics of Lexical Interfaces. CSLI Publications/Center for the Study of Language and Information.
Labile verbs in Old Indo-Aryan: their origin, decline and fall
(A typological perspective)

Leonid Kulikov

Adam Mickiewicz University in Poznań / University of Pavia – Marie Curie / Ghent University

The topic of the present paper is the origin and history of certain aspects of transitivity oppositions in Indo-Aryan languages. It focuses on the evolution of the system of labile verbs (= verbs that can show changes in syntactic pattern with no formal change in the verb, such as English *The door opened* ~ *John opened the door*). According to the *communis opinio*, several ancient Indo-European languages, including Old Indo-Aryan (Vedic Sanskrit) had a considerable number of labile verbs and verbal forms, cf. *rudrāḥ ptāsyā sādanaśu vāyrydhūḥ* (RV 2.34.13) ‘Rudras have grown [intransitive] in the residences of the truth’ ~ *índram ukthāni vāyrydhūḥ* (RV 8.6.35) ‘The hymns have increased [transitive] Indra’. This paper offers a general survey and analysis of the Vedic verbal forms for which labile patterning is attested. I will subject to critical analysis the hypothesis on the high degree of lability of the ancient Indo-European syntax (e.g. Hirt 1937), elucidating the position of the Indo-Aryan branch from the point of view of the history of transitivity and lability.

Combining typological analysis with text-critical and philological analysis of the relevant texts and integrating data from Indo-Aryan with evidence from other Indo-European languages that attest drastic changes in the system of labile verbs (such as Germanic or Greek), I concentrate on mechanisms responsible for the emergence and decline of lability, uncovering the main scenarios of the rise and disappearance of lability attested in the history of Indo-Aryan, focusing mainly on the Old Indo-Aryan period.

On the basis of the methods described above, I offer an adequate description and historical analysis of drastic changes in the history of transitivity (in particular, lability) in Indo-Aryan languages. In particular, I argue that, for most of labile forms, the secondary character of lability can be demonstrated. Thus, for many labile forms with middle inflection (especially, forms belonging to the present system, such as class I presents *svādate* ‘makes sweet (for oneself) / is sweet’ (cf. RV 9.74.9 svádasva ‘be sweet [for Indra]’ ~ RV 3.54.22 svádasva); námate ‘bends [transitive] (e.g. ‘he bends arrows for himself) / bends [intr.]’; see, for instance, Gotô 1987: 193–196), labile patterning results from the polyfunctionality of the middle diathesis (self-beneficent / anticausative).

I argue that basic scenarios of the rise of lability include, in particular, transitive-causative reanalysis of ‘content accusative’ constructions with some fundamentally intransitive verbs, such as *pusyati* ‘prospers; makes prosper’, in accordance with scenarios of the type ‘X prospers (in) Y’ → ‘X makes Y prosper’. I also demonstrate that the main diachronic process responsible for decline of lability is the development of the overt causative/anticausative morphology and the expansion of productive morphological (anti)causatives.

I will further demonstrate that, within the Old Indo-Aryan period, we observe the decline of the labile type. Already in the second most ancient Vedic text, the Atharvaveda, the number of labile forms considerably decreases. Thus, most of the active perfects which show labile syntax in the Rgveda are either attested in intransitive usages only, or in transitive usages only, or do not occur at all.

At the end of the Old Indo-Aryan period we observe the emergence of some new labile verbs, such as *śṛjyate* ‘is made’ / ‘makes’ or *sūyate* ‘is produced’ / ‘produces’ (see Kulikov 2012:
508ff.), foremost due to the collapse of the Old Indian morphological system and the influence of Middle Indo-Aryan languages. Nevertheless, even in spite of these phenomena, the general tendency towards the decline of labile syntax continues to prevail.

I argue that the decline of labile syntax documented in the history of Indo-Aryan languages can be regarded as part of a more general tendency towards the non-syncretic type of encoding of transitivity oppositions, attested in some Eastern branches of the Indo-European language family, in particular, in Indo-Iranian. Typologically, the Eastern type shares more features with some non-Indo-European families, such as Turkic, rather than with such Western branches of Indo-European as Germanic, Romance or Greek (see e.g. Lavidas 2004).

An historical and typological analysis of the development of labile verbs in Indo-Aryan furnishes important evidence for diachronic typology of labile verbs and transitivity oppositions. Such a diachronic typology will serve as a reliable basis for an adequate description of labile verbs and transitivity both in Indo-European and in other language families.

References
The translatability of *kavíkratu* and the lexical diversity of the Rigvedic poetological lexicon
Frank Köhler, University of Tübingen

The translation of archaic texts is a challenging task for any translator, as is the search for different lexical layers within it. In this talk it will be argued that with regard to the Rgveda (RV) a certain subset of its vocabulary, namely that of ‘poetology’ (i.e. those terms which the rigvedic authors have used for depicting their own activity, namely composing poems) may be used to discover an additional parameter for organizing the rigvedic lexicon. It will be argued that the uneven distribution of terms of this subset may be seen as revealing differences not only with regard to historical, regional or sociolinguistic strata of the terms in question, but also with regard to differences of the world views which form their basis. The term *kavíkratu* will serve as an example. This *bahuvrīhi* compound is attested ten times in the RV (1.1.5; 3.2.4; 3.14.7; 3.27.12; 5.11.4; 6.16.23; 8.44.7; 9.9.1; 9.25.5; 9.62.13), seven times as an attribute for Agni, thrice for Soma. Its morphology (adjectival possessive compound) and the meaning of its two constituents *kavi* “poet” and *kratu* “resolve” are more or less undisputed, hence the translation “with a poets’ resolve”. However, in spite of the rare occurrence and the apparent lucidity of the morphology and semantics of *kavíkratu*, such a translation may not be appropriate. The uneven distribution of this term and its marked position within the stanzas where it occurs in combination with the semantic ambiguity of Sanskrit compounds may be taken as an indication, that this compound possesses a more intricate structure, and that this intricacy is the reason for its occurrence.

The former part of this compound, *kavi*, makes it likely that this term belongs to the rigvedic poetological vocabulary, and for this reason it will be subjected to an analysis in order to determine its meaning(s), its significance within the world view of the rigvedic poets, and in order to attempt to give an explanation for its uneven distribution.

Finally, the potential consequences of singling out a specific lexical subset to make it the base for a subdivision of the RV with regard to its lexicon will be outlined.
Northwestern Indo-Aryan and the Rise of Diversity in the Hindu Kush-Karakoram

Henrik Liljegren (Stockholm University)

Building on Nichols (2003), this is an attempt at characterizing the multilingual Hindu Kush-Karakoram region (northeastern Afghanistan, northern Pakistan and northernmost Kashmir) in terms of genetic stability and structural diversity, and a further development of the microtypology suggested in Liljegren (2017). It also addresses the issue of areality or sub-areality. The structural features surveyed include grammatical gender, alignment, kinship and geomorphic systems. As the dominant phylogenetic component of the Hindukush-Karakoram is Indo-Aryan, regional representatives of that particular group (of which the majority were collectively referred to as “Dardic” in the past) are the main focus of the study, but naturally references will be made throughout to languages belonging to other genera in the region (Iranian, Nuristani, Tibeto-Burman, Turkic and Burushaski) as well as to Indo-Aryan in general. The features have been selected in order to represent (relatively) independent variables, each of them illustrating a unique (areal and sub-genetic) distribution, shaped by a variety of factors and competing forces at work.

The inherited sex-based gender system largely prevails in Hindu Kush Indo-Aryan (henceforth HKIA), with most of the languages making a two-way masculine vs. feminine distinction in their noun lexicons. However, at closer inspection, these languages – in spite of their relatedness – display a few signs of significant diversification: 1. The pervasiveness of such sex-based gender is stronger (and perhaps further strengthened) in the Southeast than elsewhere, i.e. among languages spoken adjacent to the main Indo-Aryan belt. 2. It is missing altogether in two languages spoken in the opposite geographic extreme, i.e. the Northwest (as earlier pointed out by Emeneau (1965:68–71) and Bashir (2003:823)), and is on the retreat in yet another neighbouring language (all three characterized by an animacy-contrast of low complexity). 3. The Southwest stands out with a few languages that instead combine their inherited sex-based gender with animacy-related distinctions and thereby form highly complex agreement patterns.

Figure 1: Map of the Hindukush-Karakoram region with languages plotted and labelled with three-letter ISO 639-3 codes.
As for alignment patterns, the HKIA languages display a great range of variation (as laid out in further detail in (Liljegren 2014)). The diversity is primarily evidenced in the case marking of core argument noun phrases and verbal person marking properties. As many as six distinct alignment types have been identified, each reflecting contact-induced changes that can be attributed to three significant areas that conflate in the region: 1. A large Persian-dominated area overlaps with the Western part of the region, characterized by overt patient marking. 2. An area in the East, with e.g. ancient Tibetan influences, is characterized by overt agent marking. 3. An area in the South, bordering on the influential Hindi-Urdu belt, is characterized by patient agreement in the perfective.

For kinship as a feature, the present study is restricted to the lexical items used for one’s parents and their siblings. Here, too a great deal of variation is displayed, with a total of six verified configurations. In essence, however, the distribution is the result of three competing systems (the remaining three constituting hybrids), each with a clear geographical distribution: 1. A maximum differentiating terminology, with six different terms (F≠FB≠MB/M≠MZ≠FZ) dominates in a southern belt, thus aligning itself with Punjabi kinship systems. 2. A pattern F=FB≠MB/M=MZ≠FZ is an eastern or northeastern feature, possibly reflecting the ancestral terminology of Burushaski (Parkin 1987:165) and the one used in Balti, the nearest Tibeto-Burman neighbour; if looking at the distribution of F=FB only, it appears typical of the languages spoken in an uninterrupted central belt, stretching all the way from the extreme Southwest to the extreme Northeast. 3. An “aunt” and “uncle” terminology (F≠FB=MB/M≠MZ=FZ) is found in the Northwest (consistently so in a single HKIA language), with obvious reflexes in adjacent non-Indo-Aryan communities in the Pamir.

Although deserving a more careful cross-linguistic study, a preliminary survey reveals the presence of a geomorphic system of spatial reference in a few of the HKIA languages spoken in a subarea in the West, whereas it seems virtually absent in other parts of the larger region. Languages in this subarea (along with neighbouring Nuristani languages) linguistically encode the inclination of the mountain slope, the flow of the river as well as boundary-crossing. This partly confirms Palmer’s (2015) so-called Topographic Correspondence Hypothesis, predicting that a language’s system of spatial reference will reflect the topography of the surrounding landscape. However, that the emergence and pervasiveness of such a system is further conditioned by language contact, is evidenced by the subareal clustering of a few other structural features –retroflex vowels, pronominal kinship suffixes and bisyndetic contrast marking (Liljegren & Svärd Forthcoming) – coinciding with the boundaries of “Peristan”, an area that until relatively recently constituted a pre-Islamic cultural sphere with Nuristan (previously referred to as Kafiristan) as its most prominent local centre of influence (Cacopardo & Cacopardo 2001: 249–250; Klimburg 2008; Jettmar, Jones & Klimburg 1975: 394).

References


Some aspects of noun morphology in Toto
Patrycja Ava Markus, Jadavpur University

Toto is a highly endangered Sino-Tibetan language (around 1400 speakers), spoken by the tribe Toto in Totopara in the upper part of the state West Bengal, India. Nowadays, this language is strongly influenced/superseded by languages such as Bengali, Nepali, Bhutanese. The main objective of this paper is to provide new inputs regarding nouns in Toto – whether there is any distinction between them, what are the case markers, how we create plural forms, classifiers, how do nouns behave with adjectives and verbs. I have collected my data from several native speakers of Toto in Totopara: Shobha Toto, Shanti Toto, Sanchita Toto, Bhugol Toto and children from the Dhanapati Toto Memorial High School. Here are some of the results of my work:

There is no gender marking in Toto. Words like /pika/ (cow) or /dikka/ (buffalo) are nowhere close to each other. Both share the same ending, there’s no visible root which both of these words could share. That applies to pronouns as well:

<table>
<thead>
<tr>
<th></th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singular</td>
<td>ka</td>
<td>nați</td>
<td>aku</td>
</tr>
<tr>
<td>Plural</td>
<td>kabifia</td>
<td>națišia</td>
<td>abifia</td>
</tr>
</tbody>
</table>

There’s no suffix which we could add after /ka/ or /aku/ to tell exactly whether the speaker or the described person is a female or male.

Plural forms in Toto can be created by adding a prefix /uiʃe/ or suffix /bi/. There’s no clear rule as to which of them applies to animate or non-animate objects, native speakers seem to apply them quite randomly. Examples: uiʃe ku𝑖tui (eggs), uiʃe giŋŋakoeʃi (spiders), ajubi (mothers), lapabi (leaves), eʧabi (horses).

Numbers in Toto:

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 itʃo</td>
<td>11 ʧutʃi</td>
<td></td>
</tr>
<tr>
<td>2 niʃo</td>
<td>12 ʧuni</td>
<td></td>
</tr>
<tr>
<td>3 jʊnʃo</td>
<td>13 ʧuʃi</td>
<td></td>
</tr>
<tr>
<td>4 ʤiʃo</td>
<td>14 ʧuʤe</td>
<td></td>
</tr>
<tr>
<td>5 natʃo</td>
<td>15 ʧuja</td>
<td></td>
</tr>
<tr>
<td>6 ğuʃo</td>
<td>16 ʧuru</td>
<td></td>
</tr>
<tr>
<td>7 niʃo</td>
<td>17 ʧuʤu</td>
<td></td>
</tr>
<tr>
<td>8 jaʃo</td>
<td>18 ʧodʒe</td>
<td></td>
</tr>
<tr>
<td>9 kuʃo</td>
<td>19 ʧuʤeh</td>
<td></td>
</tr>
<tr>
<td>10 taiʃo / tjaʃo</td>
<td>20 niʃo</td>
<td></td>
</tr>
</tbody>
</table>

Nowadays people mostly use Bengali or Nepali numbers. Interesting fact is that numbers from 1-10 share the same ending, while numbers from 11-19 share the same initial consonant. Also number 2 and 20 are exactly the same.

Toto consists of a long list of classifiers which apply to certain objects:
<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>kuĩĩĩ/kuĩĩĩ (egg)</td>
<td>uĩũũ</td>
<td>nuĩĩĩ</td>
<td>ʃũũũ</td>
<td>ʃũũũ</td>
<td>nũũũ</td>
<td>ʃũũũ</td>
<td>nũũũ</td>
<td>jũũũ</td>
<td>kuĩĩĩ</td>
</tr>
<tr>
<td>ʰĩĩ (book)</td>
<td>ɨʃo</td>
<td>nɨʃo</td>
<td>ʃɨʃo</td>
<td>ɡɨʃo</td>
<td>nətʃo</td>
<td>ɣɨʃo</td>
<td>nɨtʃo</td>
<td>ɣɨtʃo</td>
<td>ʃɨtʃo</td>
</tr>
<tr>
<td>ɾɛʃ (rupee)</td>
<td>iɾeŋ</td>
<td>nʃe</td>
<td>ʃiʃe</td>
<td>ʃiʃe</td>
<td>nʃe</td>
<td>ʃiʃe</td>
<td>nʃe</td>
<td>jʃe</td>
<td>kuʃe</td>
</tr>
<tr>
<td>ʰɨʃ (person)</td>
<td>ɨʃo</td>
<td>nɨʃo</td>
<td>ʃɨʃo</td>
<td>ɡɨʃo</td>
<td>nətʃo</td>
<td>ɣɨʃo</td>
<td>nɨtʃo</td>
<td>ɣɨtʃo</td>
<td>ʃɨtʃo</td>
</tr>
<tr>
<td>fʃeʃ (tree)</td>
<td>ɨʃeŋ</td>
<td>nɨteŋ</td>
<td>ʃiʃeŋ</td>
<td>naten</td>
<td>ʃiteŋ</td>
<td>nuteŋ</td>
<td>ʃuteŋ</td>
<td>ʃiteŋ</td>
<td>ʃiteŋ</td>
</tr>
<tr>
<td>ʃeʃ (animal)</td>
<td>ipu</td>
<td>nipu</td>
<td>ʃipu</td>
<td>ʃipu</td>
<td>napu</td>
<td>tupu</td>
<td>nipu</td>
<td>japu</td>
<td>kupu</td>
</tr>
<tr>
<td>luĩĩũ (stone)</td>
<td>iruŋ</td>
<td>niruŋ</td>
<td>ʃiɾuŋ</td>
<td>ɡiɾuŋ</td>
<td>naruŋ</td>
<td>ɾuɾuŋ</td>
<td>niruŋ</td>
<td>jaruŋ</td>
<td>kuɾuŋ</td>
</tr>
<tr>
<td>fa (house)</td>
<td>iʃa</td>
<td>nɨʃa</td>
<td>ʃɨʃa</td>
<td>ɡiʃe</td>
<td>natʃo</td>
<td>ɣɨʃo</td>
<td>nɨtʃo</td>
<td>ɣɨtʃo</td>
<td>kuʃo</td>
</tr>
</tbody>
</table>

In Toto, when we put the noun and adjective together, there’s no alternation. Examples:

As we can see in the table below, possessive adjectives also have their own form:

<table>
<thead>
<tr>
<th></th>
<th>1&lt;sup&gt;st&lt;/sup&gt;</th>
<th>2&lt;sup&gt;nd&lt;/sup&gt;</th>
<th>3&lt;sup&gt;rd&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singular</td>
<td>kabko (exc.)</td>
<td>nəko</td>
<td>ako</td>
</tr>
<tr>
<td>Plural</td>
<td>jaŋko (inc.)</td>
<td>naʃiʃibliŋko</td>
<td>abiʃko</td>
</tr>
</tbody>
</table>

Case markers in Toto:

- Nominative – created by adding /fia/ to the noun: aku-fia (he/she), gan-fia (song)
- Accusative/Dative – created by adding /fiŋ/ to the noun:
  - aku mæmæ fəŋ-fəŋ kəŋʃiŋnə
  - 3SG girl tree-ACC look-PRS-PROG
  - “She is looking at the tree”
- Genitive – created by adding /ko/ to the noun: ʃoʃia-fia ʰĩĩ – Shobha’s book,
  - apa-ko ɡəʃkən – father’s shop
- Locative – created by adding /ʃa/ to the noun:
  - ako mæmæ-fia ʃiŋiŋ-ʃa ʃiiʃiŋnə
  - 3SG girl-NOM rope-on hang-PRS-Prog
  - “She is hanging on the rope”
- Instrumental/Ablative – created by adding /ʃo/ to the noun:
  - abifia ʃaraʃʃa ʃiʃa
  - 3PL plates-INST eat-PST
  - “They ate from plates”

Except for the Locative case marker, in Toto we can find postpositions to describe a certain location of an object:
“behind the…” - created by attaching the Genitive marker /ko/ and postposition /nuniŋa/ to the noun:
 complète-ča kom-ko-nuniŋa
buffalo-NOM house-GEN-behind
“The buffalo is behind the house”

“next to…” – created by attaching the Genitive marker and postposition /abætɑ/ to the noun:
 complète-ča pukri-ko-abætɑ
temple-NOM pond-GEN-next to
“The temple is next to the pond”

These are in short the results of 4 days of fieldwork in Totopara in September 2016. There is much more yet to document and I hope that one day I will again have the opportunity to do so.

Bibliography


Discourse particles in Hindi and the relation between grammar, enunciation and “modal” meanings

Annie Montaut
Emeritus professor, INALCO, Paris, UMR Structure et Dynamique des Langues (CNRS/IRD/INALCO)

The use of particles is particularly profuse in spoken Hindi (as in any spoken language, yet with a quite limited number of particle items in Hindi) which are obviously associated with the packaging of information structure, all the more crucial in spoken statements since they cannot be separated from the enonciative frame of their uttering. Word reordering is the most common way in Hindi for focalizing and topicalizing a constituent, without morphological restructuration. A topic is thus fronted and a focus is preverbal. However, the language also displays particles for this purpose, particularly but not only, with constituents already respectively fronted and preverbal in the unmarked order (SOV). The paper bears on the most frequent of such particles, the topic enclitic particle to and the focalizing enclitic particle hī. Yet both also display ‘modal’ (argumentative, polemical) meanings which cannot be reduced to the marking of topic and focus respectively, and both also behave as conjunction/resumptive pronoun and restrictive particle respectively. The third most frequent particle bhī, which shares with the first two a double function/nature, namely an array of rather grammaticalized functions (coordination, concession, indefiniteness) and a purely enonciative function, will help analyzing the relation between what is often treated as modal particles (Abtönungspartikel) in Germanic languages or discourse/enonciation particles elsewhere, and their non-modal counterparts in language.
Does denaa GIVE as a V2 render benefactive meaning in Hindi?:
A corpus-based comparative analysis in Hindi and Japanese

Miki Nishioka
Osaka University

In this paper, I will focus on the verb denaa GIVE as an auxiliary verb, that is, a V2, and investigate whether denaa GIVE as a V2 renders benefit or benefaction, compared to Japanese counterparts such as ‘yaru / ageru / kureru’ GIVE. Masica (1991: 326-30) points out that certain special V2s in V1 + V2 concatenations have been called intensifiers, operators, explicators, or vectors; namely GO, COME, GIVE, TAKE, etc., all of which express manner-specification. Masica counted ‘benefaction’ as the manner-specification likely expressed by GIVE. Snell (2010: 189) who also ascribes a benefit function to GIVE verbs such as देना, suggests that the benefit of an action “flows away from the doer”.

Numerous works in many languages have studied benefactive construction, both individually and typologically. Of these, Shibatani (1996) submits that benefactive construction is based on a GIVE schema: [NP1 NP2 NP3 GIVE], and also discusses the [NP1 NP2 NP3 V1 GIVE] concatenation in Japanese using the same schema. Creissels (2010: 35) states that the verb GIVE underwent grammaticalization, and treats the lexical verb GIVE as a marked-Vlex type V2, as observed in Indo-Aryan languages. As we know, Hindi, an IA language, has the same construction as Japanese.

The semantic meaning of the Japanese GIVE, as we see in Shibatani (1994) and Yamada (2004), is limited to the benefactive and malefactive. In addition, it is easily negated with the negative marker -nai. An illustrative example follows.

(1) haha-wa chichi-ni gohan wo tsukutte age-nai.
mother-NOM father-DAT food ACC make.PTCP GIVE-NEG.NON-PAST
‘My mother does not make food for my father.’

This sentence renders as: It is not a fact [that my mother makes food for my father]. The scope of negation includes the whole sentence. What about Hindi translations, then? Assuming that Hindi employs almost the same V2 functions as Japanese, Hindi sentences should be negated without any obstacles.

(2) a. "maA baap ke lie khaanaa nahil banaa detii hai.
mother father DAT food NEG make.STEM GIVE.IPFV COP.PRST
‘My mother does not make food for my father.’

b. maA baap ke lie khaanaa nahil banaatti hai.
mother father DAT food NEG make.IPFV COP.PRST
‘My mother does not make food for my father.’

[p.c. Rajesh Kumar and Gunjan Sharma]

Sentence (2) a. sounds inapplicable or strange to my informants, who are native speakers of Hindi. The applicable option is (2) b, to the affirmative: “maA baap ke lie khaanaa banaa detii hai”. This is what Jagannathan (1981: 272) and Snell (2010: 190) have pointed out. They had already mentioned that auxiliary verbs such as GO, GIVE, TAKE never or hardly ever appear negated. According to Nishioka (2013), based on a web corpus of Japanese, this is also true of Japanese V2 shimau ‘PUT AWAY’, which bears a similar ‘manner’ of action (to quote Snell) to
that is, various forms of ‘completeness’. It was also roughly true of jaanaa GO in Hindi, in a web corpus research by Nishioka and Akasegawa (2015), except in ho- and rah- as a V1.

Thus the question arises: Does the Hindi denaa as V2 bear benefaction or benefactive meaning, as in Japanese and other languages? To examine this question, I searched a web corpus using the criterion of co-occurrence of V2 GIVE with a negative marker: nahiI, na, mat, to check for the same situation with the two patterns [P1. NEG + V1 + GIVE] and [P2. V1 + NEG + GIVE]. Only unmarked declarative sentences were targeted. Exclamatory, interrogative, and imperative sentences were excluded from the investigation, as were passive sentences and noun phrases.

The main findings are:

1. In P1, the number of indicative sentences is almost one-eighth of subjunctive or subjunctive-related sentences. In P2, it is roughly one third. This means that the frequency of V2 denaa ‘GIVE’ with negative indicative sentences is very low.

P1. [NEG + V1 + GIVE]

<table>
<thead>
<tr>
<th>(1) Sentence Type</th>
<th>(2) Mood</th>
<th>(3) Declarative Sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exclamatory</td>
<td>Indicative 130</td>
<td>Indicative 42</td>
</tr>
<tr>
<td>Interrogative</td>
<td>Subjunctive 348</td>
<td>Subjunctive 326</td>
</tr>
<tr>
<td>Declarative</td>
<td>Imperative (infinitive usage included) 22</td>
<td>Total 368</td>
</tr>
<tr>
<td>Imperative</td>
<td>Total 500</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

P2. [V1 + NEG + GIVE]

<table>
<thead>
<tr>
<th>(1) Sentence Type</th>
<th>(2) Mood</th>
<th>(3) Declarative Sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exclamatory</td>
<td>Indicative 62</td>
<td>Indicative 47</td>
</tr>
<tr>
<td>Interrogative</td>
<td>Subjunctive 181</td>
<td>Subjunctive 172</td>
</tr>
<tr>
<td>Declarative</td>
<td>Imperative (infinitive usage included) 33</td>
<td>Total 219</td>
</tr>
<tr>
<td>Imperative</td>
<td>Total 276</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Data compiled from Corpus Of Spoken Hindi

2. In most of the indicative sentences, a negative marker does not negate an affirmative proposition in itself, but negates a part of the proposition (a partial negation) as in ex. (3). Here, adding the particle (PTCL) to to mijaaaj topicalizes ‘nature’, although old age can change other things:

(3) buRhaapaa     insaan    kaa   mijaaaj  to   nahiI  badal
detaa.  GIVE-IPFV
‘Old age does not change human nature.’

http://www.hindisamay.com/contentDetail.aspx?id=3177&pageno=1
3. The tendency towards the co-occurrence of *denaa* with a negative marker is very similar to that of *jaanaa*. This allows us to assume that the use of *denaa* as a V2 indicates a manner of ‘completeness’ rather than a ‘benefit’ or ‘benefaction’ in Hindi, unlike in Japanese (and maybe other languages).

References


Exploring the language layer of the Rigvedic dānastuti genre
Georges-Jean Pinault
(Paris, EPHE/Sorbonne, PSL)

The linguistic variety of the ancient Vedic texts is a well-known fact. This can be observed within the Rigveda itself, the most ancient collection of hymns, and if one compares the language of the Rigveda with the one of the Atharvaveda. Glimpses of Vedic dialects can be detected in several passages and words, although the poetic language shows a high degree of convention and normalizing. Among the hymns of the Rigveda few specific features can be attributed to the different families of bards, even though one can surmise that they belonged to different regions of the Vedic world. It is also likely that some families or so-called “branches” were linguistically mixed. The hymns resort to different genres of discourse. The dānastuti, lit. ‘praise of the gift’, makes a distinct part of the poetic competence. The passages in question, which are often limited to a single stanza, while others are more developed, making a substantial part of the poem, are devoted to the praise of the generosity of the patron, who is expected to reward appropriately the poet for his work. There exists a comprehensive survey of these parts of the hymns of the Rigveda in the dissertation of Patel (1929), who has been the student of Karl F. Geldner. This meritorious book describes mostly cultural, historical and ritual features: the references to the poets, to their social entourage, the types of gifts which constitute the reward of the poets, the gods mentioned in these stanzas. On the other hand, the familiar, and in cases crude or mischievous, tone of these pieces has been noted by several commentators of the Rigveda. It would be however too simple to consider that these parts reflect faithfully everyday speech. The paper will explore if the dānastutis have linguistic traits which contrast with the standard layer of the Rigvedic poetry at all levels: phonology, grammar (use of tenses and moods), syntax, vocabulary. On the level of stylistics and poetics, one will also investigate if they show a specific phraseology, or some kind or rearranging of the standard phraseology of the praise of the gods and goddesses.

References

The Early Vedic pīpáya-type

Roland Pooth
Ghent University

The Early Vedic language of the Rgveda exhibits a reduplicated and thematic TAM stem with reduplication, guṇa (full a-) grade of the root (e.g. pīpáya-) and thematic vowel which has been rendered both “subjunctive” (pīpáya-tas) and “pluperfect” (á-pīpaya-). There are several linguistic problems with the philological separation of two stems. The function of this stem will be the topic of my talk.

References

The Role of Pitch in Gurung Tone
Danielle Ronkos, The Graduate Center, CUNY

Gurung is a threatened Sino-Tibetan language spoken in central Nepal. Like other languages in the Tamang-Gurung-Thakali-Manange (TGTM) subfamily, it has lexical tone. Although linguists working on the TGTM languages agree that the tone-bearing unit is the phonological word (Glover 1969; Glover & Glover 1972; Hildebrandt 2003, 2007a & 2007b; Mazaudon 1988 & 2014; Sprigg 1997), the exact acoustic properties of these tones have yet to be determined. Speakers of these languages have intuitions about tonal category membership and can reliably produce and distinguish minimal tone pairs; however pitch alone does not seem to predict a given word’s tone. Recent work by Mazaudon (2014) on Tamang suggests that these tones are best treated as bundles of acoustic features which speakers use in tandem to perceive and indicate tone category membership. She identifies F₀, breathy voice, and onset consonant voicing as the relevant features for Tamang tone. This supports earlier work such as Hildebrandt’s (2003) dissertation on Manange tone, which concluded that F₀ was an inconsistent predictor of a given word’s tone across Manange speakers.

Although both Hildebrandt (2003) and Mazaudon (2014) include acoustic analyses of words collected from speakers of Manange and Tamang, respectively, there has not yet been any extensive acoustic analyses of words collected from speakers of Gurung. This paper aims to fill this gap by taking the methodology outlined in Hildebrandt’s (2003) dissertation and applying it to recordings collected by Swenson (2015) during a recent sociolinguistic survey of Gurung in Nepal. Analyses of variance of the F₀ measurements for the vowels in 137 tokens produced by each of the six Gurung speakers surveyed by Swenson indicate that, like other TGTM languages, F₀ is not a consistent significant predictor of tone category membership for Gurung. This suggests that Gurung tone may also be best treated as a bundle of acoustic features, and paves the way for further research to identify the relevant acoustic features for Gurung tone.

References
SIL International.
áhir budhnyaḥ and dundubhi: the Serpent of the Deep and the earth-drum: hypothesis of etymological and/or cultural connections

Paola Maria Rossi
University of Milan

The present paper aims at examining whether the well-known Greek doublets Pythôn ~ Typhôn/Typhôeus, (Ivanov-Toporov 1974; Watkins 1992; more cautiously West 2007), respectively from *phuth- and *thuph-, considered as the equivalents to the I.E. doublets *bhudh- ~ *dhubh-, referred to the I.E mythologem of the Dragon/Serpent or primordial Monster of the Deep, could be compared to an OIA doublets. Actually, the relationship between the OIA budh-n-, from I.E. *bhudh-, as budh-nā “bottom, ground, foot, root”, and the Greek pythmên is ascertained (Beekes 2010), as well as the related adjective budh-n-yā “belonging to the bottom”, especially occurring in the Vedic syntagma áhir budhnyaḥ “Serpent of the bottom/deep”, has already been compared to the syntagma Pythô ophis by Watkins (1992; 1995: 460-463). On the contrary, the case of the Greek Typhôn/Typhôeus and the related etymological interpretation seem to be more uncertain, since the linguistic and cultural interferences are manifold: it seems to be mainly associated with the Mediterranean and Anatolian context, with no evident connection to the OIA references. However, in the Vedic culture, the mythologem of the Serpent of the Deep and the related Dragon Slayer is connected to the ritual of the royal consecration (Heesterman 1959), and to the mythical motif of the rising sun (aja ekapād: Horsch 1965) with the release of waters/cows (Witzel 2005): the Water Dragon is conceived of as an obstructing being in the depth of the earth/caves which can be defeated only by breaking and splitting rocks and obstacles, making them resound and thunder. The motif of the sound as warrior weapon to defeat the enemies is well attested in the Vedic repertoire, but it assumes a specific value in connection to a particular rite, the Mahāvrata or festival of the Winner Solstice: on this occasion musical instruments are used, especially the famous “earth-drum” or bhūmi-dundubhi: it seems to evoke iconically the mythologem of the Serpent of the Deep, especially as a resounding and thunderous being (Parpola 1999), reminding some images of the Greek Typhôeus (i.e. Hes. Theog. 823-835). Therefore, the point is whether the term dundubhi, generally analyzed as an onomatopoeic term belonging to a substrate language, probably related to the Muṇḍa languages (Kuiper 1991), could be also interpreted as an I.E. derivative, associated not only mythologically with the representation of Typhôeus, but also etymologically to that *dhubh- as allotrope of *bhudh-.
In the oldest Indo-Iranian poetry – the Veda and the Avesta – just like in the language of the Homeric epics, we can discover forms of language-internal development analysable on various levels: diatopic (with some tangible elements of dialectal differentiation), diachronic (both on the scale between real archaicty and innovative features and on the level of “invented archaicty”, viz. conscious, sometimes artificial archaization up to hypercorrect “wrong improvements” emulating ancient forms of expression), as well as diastratic (above all, in a ‘sociolinguistic’ perspective of using ‘lower’ vs. ‘higher’ stylistic features). One of the aspects partially ignored so far is the field of tension between language archaicity and language change in the textual corpora within a matrix linked by genre, i.e. by comparing representatives of hymnial, ritual, incantational etc. poetry with regard to the lexical and grammatical development within representatives of the same literary and/or pragmatic genre – e.g. lexical and grammatical elements contained in the liturgical portions of the Rigveda as compared, on the diachronic and diatopic axis, with their later versions in the ritual corpora of the Yajur-Veda of the various local schools and, on a different diastratic level, with their modified variants in the mantras of the Atharvaveda.

The aim of the present paper is to fill this research gap: In the relatively less studied, but extremely archaic and highly interesting textual corpus of the RV-Khilas, we find four different liturgical complexes that are recited by the priests ‘in important moments’ of the public ritual: the litanies called Nivid-, Puroruc-, Praiṣā- and Kuntāpa-. At least two of these rituals of the RV-Khilas correspond to relevant ritual modules from the Avestan Yasna: Above all, these are the Nivid rituals (transmitted in the RV-Khila 5,4): They consist in the invocation of a deity by name + enumeration of short characteristics of the deities invoked. Portions of these hymns are attested both in more archaic strata of the Rigveda, the Family Books with the so-called Āprī- rituals, and in more recent ritual complexes such as the Ṛtu- yāja- liturgy included also in later Vedic ritual texts. On an intralinguistical level, within Vedic, these different time sections give a very useful diachronic axis to research the development of the specific liturgical vocabulary and ritual phraseology connected e.g. to the cult Fire and Waters.

On an interlinguistic level, the rituals from the Rigveda(-Khilas) mentioned so far exhibit a series of correspondences with the oldes Iranian ritual poetry, the Avesta, especially its first main module, the Long Liturgy of the Yasna. So, Y. 1–6 contains a ritual catalogue of the Thirty-Three deities, consisting of names and short characterizations extremely similar to and partly containing literal correspondences with to the Nivid rituals mentioned above; the Ṛtu- yāja- liturgy contains surprising similarities with the Avestan Yasna and Visprad liturgies. The RV-Khila 5,7, the Praisādhyāya, is the text concerning the election of the priests for the big Soma sacrifice and the distribution of their functions. In the Haoma sacrifice, Yasna 14, its correspondence is the election of and the distribution of the functions to, the priests for the Haoma sacrifice.

The Vedic language corpus more often than not contains lexical items that are attested in a very limited quantity and semantics is frequently object of serious polemics. Therefore, in order to be able to reconstruct lexical and grammatical change in Vedic on its Indo-Iranian background, it is necessary (1.) to analyze the exact morphological correspondence of the formations and (2.) the phraseological combinatorics of their constituents. The present paper
illustrates this approach on the material of texts from the RV., RVKh., AV. and YV. Moreover, what is essential from semantic viewpoint (3.), is to investigate the pragmatical context of the usage of the words and expressions concerned, as components of ritual poetry that have to be situated in the liturgical pragmatics of the Vedic and Avestan rituals.
A comparative study of the Anaphors in Meiteilon, Nyishi and Toto

Atanu Saha

Jadavpur University

The paper surveys the anaphorical constructions in Meiteilon, Nyishi and Toto and shows that these languages use similar strategies predominantly nominal one for reflexivity. On the other hand, the languages use contrastive strategies to yield the reciprocal constructions. After discussing the objectives of the paper in section 1, section 2 of the paper discusses the reflexive strategies and constructions and the next section focuses on the reciprocal constructions. Conclusions are summarized in section 4.

Keywords: Anaphor, reflexives, reciprocals, Tibeto Burman languages

1 Objectives
In this paper, I have discussed reflexives and reciprocals of Meiteilon, Nyishi and Toto. All the three languages belong to the Tibeto Burman Language Family. The main research questions are the following:

1) How is reflexivity and reciprocity expressed in the languages? Geniušien’ė (1987) emphasizes on the following aspects of reflexivity e.g. a dedicated reflexive morphology for the language in question, number of morphemes to show reflexivization, relation of the reflexive marker with pronouns, Agreement with respect to the number and gender, Syntactic positioning with respect to the verb, transitive intransitive distinction, lexical or grammatical strategy, binding domain and the effect on valency.

2) How the forms and the functions of those two types of constructions (reflexivity and reciprocity) are related as classified in Subbarao (2011) i.e.

(i) A nominal anaphor (reflexive or reciprocal), or
(ii) A verbal anaphor (a verbal clitic for the reflexive or reciprocal), or
(iii) Both nominal and verbal anaphors.

2 Reflexives
As mentioned above, a nominal strategy for reflexivity is found in all the three languages as shown in (1-3).

1. John-ŋa masa-bu u-jø-i (Meiteilon, Own data 2011)
   John-ERG self-BNF see-REF-ASP
   ‘John saw himself’.
2. no aŋa-ŋam morum-(su)-ḏen (Nyishi, Own data 2009)
   you self-ACC love-RFL-CONT
‘You love yourself.’

3. ʃanti-ha ṭaṭi-raŋ yacpa-din-na  
    santi self –NS hit- CONT-PRS  
    ‘Santi is hitting herself.’

Meiteilon also requires a verbal suffix /ə/ additionally to mark the reflexivity. In Nyishi, the verbal reflexive suffix is optional.

3 Reciprocals

Reciprocity shows up with verbal suffixes in Meiteilon and Nyishi as shown below:

4. makʰoi kʰoi-na-re  
   they divorce-RECP-ASP  
   ‘They are divorced.’
5. bulə na-miŋ-su-pa  
   they marriage-COM-RFL-ASP  
   ‘They are married.’

The difference between (4-5), indicates that in Meiteilon reciprocity is expressed by a single suffix while in Nyishi reciprocity is a composite structure comprises of a comitative suffix -miŋ and a reflexive suffix –su.

The reciprocal construction shows up with the nominal strategy as in (6) in Toto.

6. ɛbiha ebe-eja ga-diŋ-na  
   they one one-LOC sit-CONT-PRS  
   ‘They are sitting one after another’.

Another morphological variant can also be noticed in the following example:

7. i deabi taibiləra-hiŋ tæ-ɖiŋ-na  
   these people each other-ACC touch-CONT-PRS  
   ‘They are touching each other.’

4 Conclusion

To summarize, in case of reflexives, Meiteilon, Nyishi, and Toto primarily use a nominal strategy. In case of reciprocals, Meiteilon and Nyishi use a verbal strategy in contrast to the sole nominal strategy used for reciprocals in Toto. The facts from the three languages against the issues raised in objectives section are summarized as below:

<table>
<thead>
<tr>
<th>Identifying reflexivity</th>
<th>Reflexives</th>
<th>Reciprocals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Meiteilon</td>
<td>Nyishi</td>
</tr>
<tr>
<td>dedicated reflexive morphology</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>number of morphemes to show reflexivization</td>
<td>Single</td>
<td>Single</td>
</tr>
<tr>
<td><strong>Identifying reflexivity</strong></td>
<td><strong>Reflexives</strong></td>
<td><strong>Reciprocals</strong></td>
</tr>
<tr>
<td>----------------------------</td>
<td>---------------</td>
<td>----------------</td>
</tr>
<tr>
<td><strong>reciprocalization</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>relation of the reflexive marker with pronouns</strong></td>
<td>It can be used a pronoun</td>
<td>It can be used a pronoun</td>
</tr>
<tr>
<td><strong>Agreement with respect to the number and gender</strong></td>
<td>Number</td>
<td>Number</td>
</tr>
<tr>
<td><strong>Syntactic positioning with respect to the verb</strong></td>
<td>Preverbal ; SOV and a suffix</td>
<td>Preverbal ; SOV and an optional suffix</td>
</tr>
<tr>
<td><strong>transitive intransitive distinction</strong></td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td><strong>lexical or grammatical strategy</strong></td>
<td>Grammatical</td>
<td>Grammatical</td>
</tr>
<tr>
<td><strong>the effect on valency</strong></td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Form function</strong></th>
<th><strong>Meiteilon</strong></th>
<th><strong>Nyishi</strong></th>
<th><strong>Toto</strong></th>
<th><strong>Meiteilon</strong></th>
<th><strong>Nyishi</strong></th>
<th><strong>Toto</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>i) A nominal anaphor (reflexive or reciprocal), or ii) A verbal anaphor (a verbal clitic for the reflexive or reciprocal), or (iii) Both nominal and verbal anaphors</td>
<td>Both nominal and verbal anaphors</td>
<td>Both nominal and verbal anaphors</td>
<td>Only Nominal</td>
<td>Verbal</td>
<td>Verbal</td>
<td>Only nominal</td>
</tr>
</tbody>
</table>

What is also noted is that Nyishi does not have a dedicated reciprocal morphology; hence it uses a compositional strategy. Reflexives and reciprocals agree in numbers and no gender agreement is found in general in these languages.

5 References


In a 2013 article, Acharya has shown that the original pre-Kauṇḍinya, pre-philosophical pāśupatavrata essentially consisted in imitating the behaviour of a bull, and that traces of such a vow can be found in Vedic sources.

In my research I have investigated the hypothesis that this practice may have been informed by cultural models originating in the Indo-European institution of the Männerbund and its initiatory practices involving ritual masking. My argumentation relies on textual evidence to illustrate the persistence of a consistent set of cultural traits, from the reconstructed Indo-European past all the way to early Śaivism via the mediation of Vedic Vrātya culture through a series of re-interpretations and adaptations.

After summarising my views, findings and methodology, I will focus on the sixth anuvāka of book 17 of the Paippalādasaṃhitā of the Atharvaveda (AVP 17.27–43), of which I have prepared a new critical edition and a first English translation based on Bhattacharya’s 2011 editio princeps as well as additional manuscript sources. This anuvāka is a long prose text dedicated to illustrating the mythological origin and the efficacy of a so-called anaḍudvrata (‘the vow of the draft ox’).

As first pointed out by Bisschop & Selva (2016), this text shows some striking textual parallels with the Pāśupatasūtra. In the current talk I will focus on how its content, style and formulaic language suggest that it may have been composed in a Vrātya environment, making it the missing link between Vedic Vrātya culture and Pāśupatism.

This text, especially if read together with AVP 3.25 (~ AVŚ 4.11) — of which I have also prepared a new edition — might also provide some evidence in support of Pontillo’s (in press) hypothesis of a connection between the pravargya or gharma ritual (for which an unorthodox origin has long been suspected) and the Vrātya culture.

References


A pragmatic account of definite and indefinite descriptions in Hindi –
the case of 'hai' and 'hotā hai' forms
Ghanshyam Sharma, INALCO-MII-UMR-7528-LABEX-EFL, Paris

Hindi displays two present tense singular forms of the verb honā ‘to be’, namely hai (present form of the auxiliary) and the conjugated hotā hai (i.e. imperfective participle of honā plus its present form as an auxiliary), both of which happen to have the same rendering in English and other European languages. The received wisdom is that the simple form, hai, indicates a present quality of the subject it is predicate of whereas the conjugated hotā hai (also known as the ‘long’ or ‘full’ form) indicates a general phenomenon or the universal truth or at least an indefinitely-lasting quality of the subject it refers to, as in (1) and (2) respectively.

(1) gilāś=kā pānī khārā hai
   glass=GEN water.M.S salty.M.S aux-PRES.S
   “The water of the glass is salty.”

(2) samudra=kā pānī khārā hotā hai
   sea=GEN.M water.M salty-M.S be-IMPFTV.M.S aux-PRES.S.
   “The water of the sea is salty.”

It is assumed that the hai-form in (1) indicates the present, actual or temporary quality (i.e. salinity) of the water in question whereas the hotā hai-form in (2) is employed to indicate a long-term or universal quality of sea water. However, this does not seem to be the case in the following examples.

(3) sanskrit bhāṣā mahatvapūrṇa hai/
    Sanskrit language-F.S important aux-PRES.S./
    be-IMPFTV.F.S aux-PRES.S
    “The Sanskrit language is important.”

(4) hindū sādī=mē sanskrit bhāṣā mahatvapūrṇa hai/
    Hindu marriage=in Sanskrit language important aux-PRES.S/
    hotī hai
    be-IMPFTV.F.S aux-PRES.S
    “In a Hindu marriage the Sanskrit language is important.”

It appears that even though the Sanskrit language in (3) is important at all times, the conjugated hotā hai is a bit anomalous. Whereas in (4), the conjugated form is preferred even though ‘the Hindu marriage’ is not a case of universal truth. A careful study of the phenomenon, however, reveals that Hindi – an articleless language – makes use of the distinction between the hai-form and the hotā hai-form to cope with the phenomenon of generics as well as definite and indefinite descriptions which are found in the articulated languages such as English. In fact, it would be incorrect to define the hotā hai form as ‘long’ or ‘full’ form. The hai-form and the hotā hai-form seems to be related to the question of particular and generic references, respectively. The aim of the paper is to make an attempt to analyze the two Hindi forms in the light of definite and indefinite descriptions as well as generics. Furthermore, an attempt will also be made to ascertain those pragmatic factors which are responsible for the phenomenon in question.

References
Carlson, Gregory, 1977, “A Unified Analysis of the English Bare Plural” Linguistics and Philosophy 1: 413-456
A widely observed and frequently commented on aspect of contemporary language use in South Asia is the use of more than one language in a single discourse, in fact, even within the same sentence. This phenomenon has been studies under the rubric of Code-Switching and Code-Mixing (eg. Annamalai 1973, Sridhar 1978, Sridhar and Sridhar 1980, Kachru 1977, Gumperz and Wilson 1971, among others; see Kachru, Kachru, Sridhar 2008 for a detailed bibliography) and lately, “translanguaging” (Garcia 2009).

In this paper, I will analyze mixed language use data from a wide range of discourse types to explore their implications for an integrated model of language performance in structural, sociolinguistic, psycholinguistic, and cultural terms. I will first use morphological, lexical, syntactic, and discourse properties of this language use to establish their intricate patterning, which is not easily described in a conventional monolingual model of linguistic competence. I argue that mixed language use requires a multilingual discourse grammar (Ferguson 1979) and a psycholinguistic model where both grammars are simultaneously active, interactive, and integrated (Sridhar and Sridhar 1980).

I will then go on to discuss the social and cultural dimension of the use of mixed languages. I contrast two evaluations: On the one hand is the widely prevalent attitudes of both lay persons and language professionals, who see mixing as “corruption,” a result of laziness and a sure sign of the coming extinction of their language. They even go so far as to call for protectionist measures on the part of the government. In contrast, I show that mixing is a versatile and effective communication strategy that cleverly exploits the full range of resources of the multilingual community. It extends the verbal repertoire of a language to accommodate the demands of modernization and globalization. Although it presupposes multilingual competence, used among proficient multilinguals who are at home in all the languages of the community’s repertoire, it poses no more problems of intelligibility than monolingual usage.

Contrasting the extremes of pure monolingual language use in either direction, which are considered unacceptable because of associations of puritanism on the one hand and pretentiousness on the other, mixed language users may be said to be adopting a language maintenance strategy, rather than a sign of language loss. Furthermore, the mixed language is a cultural imperative to express the essential hybridity of the modern South Asians, who see the larger world (represented through English in South Asia) as part of their sensibility while maintaining their regional identities. This point is reinforced by the embrace of mixed language by cutting edge language practitioners, such as media writers who see it as an expression of youth culture. On the other hand, the interlarding of regional South Asian elements in South Asian varieties of English (eg., Hindi, Kannada, Bengali, etc.) has the complementary effect: it localizes English, makes it less alien and more acceptable. Thus, the use of a mixed language, in either direction, is best understood in terms of a functionalist explanation, as a civilizational response to globalization and colonization.

I support the foregoing arguments with a structural and functional analysis of a wealth of data from English-Kannada and English-Hindi mixing drawn from a wide range of discourse contexts, including the following: ordinary informal conversations between friends, television interviews and FM radio talk shows, plays, short stories and novels, film dialogues, print and TV advertising, Internet chat rooms, social media postings, and SMS messages, news reporting, academic lectures, public speeches, and songs, and so forth.

I end on a historical note: Language mixing has been a strategy in India from the earliest
times. It accounts for the extensive linguistic convergence, leading to the emergence of India as a linguistic area; emergence of new contact-based languages such as Urdu, and Indian English; and the enrichment and modernization of regional languages with profound infusion from Sanskrit, Persian, and now English. Thus, language mixing has been central, not peripheral to defining the character of Indian civilization.

Select references


Language Maintenance among Thanjavur Marathi speakers

Kamal K. Sridhar
Stony Brook University

This paper examines language maintenance in a small, little-known community, known as Thanjavur Marathi speakers. They migrated from a Marathi-speaking area of Maharashtra to the Tamil-speaking area of Thanjavur (formerly Tanjore) in Tamil Nadu in the late 1600s, and have continued to maintain Marathi as home language in the overwhelmingly Tamil-speaking area, despite their small numbers, lack of political or economic power, and, also, in many cases, dispersal or re-migration to distant regions in India and even abroad. This project explores the role of a range of sociolinguistic and socio-cultural variables in explaining language maintenance, a major dimension of Indian multilingualism and should be of interest to sociolinguistics and the sociology of language, from both a theoretical and descriptive point of view.

India has been regarded as a unique site for the study of language maintenance (eg. Fishman, Pandit, Srivastava, Kachru, Pandharipande, Sridhar) for several reasons, including, especially:

a. The normative nature of maintenance of minority languages across centuries and even millennia;
b. The extra-ordinary complexity of social structure involving hundreds of castes and sub-castes;
c. The complex interplay of linguistic factors with and cultural factors, such as, purity and pollution, the ethnic separateness of home life, and pragmatic hierarchy of language functions;
d. The role of long-standing intensive and extensive multilingualism, resulting in multi-level borrowing and convergence with or without, pervasive code-switching and code-mixing.

The case of Thanjavur Marathi is of additional interest because of

e. The historical persistence of the language contact situation, documented across at least 350 years; and

g. The existence of potentially contradictory attitudinal factors, such as pan-Indian tolerance of linguistic and cultural diversity, on the one hand, and the area-specific, century old movement of Tamil nationalism and anti-Brahmin and anti-Aryan sentiment on the other; and

h. the dispersal of the community in various regions of India and abroad (a case of triple diaspora).

The maintenance of Marathi in this small community in these diverse circumstances raises many interesting questions. I will examine the usual sociolinguistic variables, such as, the existence of a critical mass and the language loyalty of the minority community; the attitude of the majority community and power relations between the two; interactional issues, such as the ethnic separateness of home life and opportunities for interaction in domains such as
neighborhood, education, religion, and marketplace, I will also examine questions such as how much displacement a language community can take and still survive, and the minimum amount of environmental support with which a language continues to be maintained.

These and related questions will be analyzed with reference to a variety of data. These concern specific questions of linguistic proficiency, language use, and sociolinguistic attitudes toward language switching and mixing, language maintenance policies in the family domain, socialization patterns, and institutional resources. The data were gathered from children and adults from three generations. The methods include a variety of empirical modalities, including elicitations, grammaticality judgements, translation exercises, ethnographic participant observations, and questionnaires. These were supplemented with oral history interviews with members of the community in various sites across India and the U.S., and archival data from Sarasvati Mahal Library.

The results reveal a consistent pattern of maintenance across diverse contexts and support a socio-cultural model of maintenance that goes beyond the traditional models.

References


Sridhar, Kamal K. 1989. English in Indian Bilingualism, New Delhi: Manohar,


Labelling Finiteness as a Semantic Feature  
Shubham Kumar Srivastava* and Sarada Bandana Biswas  
University of Delhi  
*Corresponding author e-mail: shubham.glasses@gmail.com

Finiteness, studied in terms of tense and agreement, has been thought to be part of the Inflection node (INFL) of a basic minimalist tree. A certain set of languages show evidence that this is in fact not true universally. European Portuguese shows instances of non-finites that show agreement patterns (Raposo 1987 & Ledgeway 1998). Stowell (1982) proposed that English infinitival subordinate clauses have an internally determined tense. These pieces of evidence indicate that defining finiteness on the parameters of tense and agreement cannot be construed as a universal truth.

Finiteness is a feature that is more than often embedded in the speaker’s attitude towards an utterance. This attitude is reflected in the grammatical mood and the illocutionary act behind the construction.

The primary hypothesis of this paper is that finiteness is a semantic/pragmatic feature that gets the Complementizer node (C) in a basic minimalist tree. Mood is a much more salient feature in human language than tense. I would eat bananas when I was a baby vs I would eat bananas if I lived in Poland. The tense form of the two modals are identical. The difference is in the mood that these modals employ: The former has a realis indicative mood in the embedded construction, whereas the latter uses an irrealis subjunctive mood in that place. (Barron 2000)

This paper draws a non-discrete scale of finiteness in terms of moods. It places various constructions on this scale depending on their mood and judge the relation of these moods with the degree of finiteness.

Indicative > Interrogative > Imperative > Subjunctive (Optative > Conditional > Counterfactual) > Participial > Infinitive

Revisiting Stowell’s (1982) examples:

(i) Jenny; remembered [PROi bringing the wine.]
(ii) Jenny; remembered [PROi to bring the wine.]

The embedded clause in (i) has an indicative mood, while that in (ii) is unmarked for any mood which makes (i) more finite than (ii). This, in fact, reinforces Stowell’s (1982) judgement that embedded clause in (ii) is unrealized for any tense, while that in (i) has a tense pliable to the matrix verb.

The above scale also justifies why certain infinitival constructions in Hindi can exist as independent construction.

(iii) čае lаnа  
Tea bring.inf-IMP  
‘Bring tea!’

In this sentence, finiteness based on the mood of the construction supersedes it being an infinitival construction. Imperatives stand as complete sentences, and (iii) is an instance of that in Hindi. Similar instances can be seen in French, and Italian.
At the illocutionary layer, the subjunctive verb, and the full infinitive indicate the irrealis mood. Different languages employ different strategies to indicate this phenomenon.

\textit{Jill} wants [\textit{PRO}, to leave/ \textit{Jim} to leave/*that \textit{she}, leaves/ *that \textit{Jim} leaves].

As one can see, English uses the to-infinitive in the subordinate clause irrespective of whether or not the subject of the main clause and the subordinate clause refer to the same entity. However, the subjunctive is not normally permitted in these situations. (Barron 2000) One of the few exceptions being a certain kind factive predicates like \textit{to regret}, and \textit{to be aware of} which have an \textit{indicative} mood and an \textit{assertive} force at the illocutionary layer.

Looking at similar examples in Hindi, one notices a different pattern emerging.

(vi) \[jɪl \ jə-ŋə \ cɑ̂̃tɪ \ hɛ\]
\[
\quad \text{Jill.3FS-Nom} \quad \text{go-INF} \quad \text{want.FS} \quad \text{be.Prf}
\]
\[ ‘\text{Jill wants to go.’} \]

(v) \[jɪl \ cɑ̂̃tɪ \ hɛ \ ki \ jɪm \ ɛlə jaye\]
\[
\quad \text{Jill.3FS-Nom} \quad \text{want.FS} \quad \text{be.Prf} \quad \text{that Jim.3MS-Nom} \quad \text{walk go.SUBJN}
\]
\[ ‘\text{Jill wants that Jim leave.’} \]

Here, \textit{jim-ka jana} is the direct object of the main clause rather than being the subordinate clause. (Also, one can notice that infinitive form of verb has been nominalized to be part of a genitive.)

(vii) \[jɪl \ jɪm-ka \ jə-ŋə \ cɑ̂̃tɪ \ hɛ\]
\[
\quad \text{Jill.3FS-Nom} \quad \text{Jim.3MS-Gen} \quad \text{go-Inf-Acc} \quad \text{want.FS} \quad \text{be.Prf}
\]
\[ ‘\text{Jill wants Jim’s leaving.’} \]

Hindi uses full (-\textit{na}-) infinitive when the subject of the main clause and subordinate clause are same (iv), and subjunctive otherwise. Even if one co-indexes the pronominal in the subordinate clause with the subject of the main clause (vii), it will still be assumed to have its referent in a different entity. This is found to corroborate the divide shown in the non-discrete scale of finiteness in terms of moods. The paper concludes that finiteness should be treated as a semantic/pragmatic feature that manifests in the grammatical mood.

\textbf{Keywords:} tense, finiteness, modals, mood, subjunctive, infinitives

\textbf{References}

Light verb constructions consisting of a converb (the main verb) and another verb (light verb) are an areal feature of modern Indo-Aryan (IA) languages. The present paper is an attempt to collect morphosyntactic data on LV in early stages of New IA languages.

There has been a vast literature on the category of ‘light verb’ in IA; however, only a few studies display a diachronic typological bias. They usually explore possible scenarios of the transition from a converb + V complex to a light verb, accompanied by various implications pertaining to morphosyntactic and semantic changes (cf. Hook 1993; Butt 2003, Butt and Lahiri 2013; Slade 2013).

In the present paper I investigate light-verb constructions in early stages of New Indo-Aryan (Rajasthani, Awadhi, Braj). Data comes from the IA corpus consisting of prose and poetry documents from XIV c. till XVIII c. I focus on morphosyntactic differences. I go through the set of properties proposed by Slade (2013) to place light verbs in a bigger typological frames. Among the properties are: interruptibility, recursion, construction-specific restrictions, dominance.

Preliminary research has shown that light-verb constructions in investigated languages seem to be quite stable when all four properties are concerned.

Compound verbs in most cases are not interrupted by any word. However in cases of compounds expressing ability the negative particle occurs between main verb and light verb (1). There is also one instance in Braj of a noun interrupting the compound (2).

1) pachai mātaṁg|i māthā tima karī
   after outcast.INS.M.SG head.M.PL this way do.CVB
rājā nai dekhāryā jima rājā|i
king.OBL.M.SG DAT show.3PL.PST.PTCP so that king.INS.M.SG

---

1 Following the VanValin’s terminology, in case of light verb constructions we deal with nuclear juncture where distinct nuclei are involved in a single core creating a complex predicate, see VanValin&LaPoll (1997:441-516)
2 Or an absolutive as in Slade (2013) where the absolutive is a fixed, indeclinable verbal form. Although Haspelmath (1995) defines the converb as nonfinite verb form whose main function is to mark adverbial subordination, he points out constructions where a superordinate verb combined with a converbal form functions as an applicative auxiliary (Haspelmath 1995:44-45). I believe that in the early stages of investigated languages we deal with converses not with stems: “The form in -i(-i) as a part of a light verb started being replaced by the root form around the 18th century” concerning Rajasthani in the upcoming Stroński, Tokaj, Verbeke.
3 Rajasthani texts: XIV-XVIII c.; Awadhi: XVI c., Braj: XVI/XVII
4 ‘interruptibility’, the ability of elements to intervene between the main verb and the light verb of the CV construction; ‘recursion’, whether a main verb may be modified by more than one light verb; ‘construction-specific restrictions’, i.e. which syntactic constructions light (vector-type) verbs can occur in; and ‘dominance’, whether agentive marking of subject of transitive verbs in perfective tenses is controlled by the transitivity of the main verb or the transitivity of the light verb.” Slade (2013:13)
"Then outcast showed heads to the king in the way that [they] couldn't be recognized by the king."

2) ghaṭatahi ghaṭata amāvasa
decrease.PRS.PTCP decrease.PRS.PTCP the night of the new moon.F.NOM.SG

bhaī dui dina lāja gāri bhuiṁ
be.F.PST.PTCP two day shame hide.CVB earth.LOC.F.SG
gāī
go.F.PST.PTCP

“Gradually [moon] started fading away that finally new moon came. [Moon] was so ashamed that [he] hid [himself] beneath the earth.”

Furthermore, when the recursion is concerned all of the investigated languages do not allow more than one light verb per main verb. However, there are instances of reduplicated main verbs (Awadhi) modified by one light verb (3) and two main verbs (where one is a causative verb) modified by one light verb (Awadhi) (4).

3) raci raci rākhe caḥdana
make.CVB make.CVB keep.M.PL.PST.PTCP sandalwood.NOM.M.SG
caurā pote agara meda
platform.NOM.M.SG smear.M.PL.PST.PTCP aloe.NOM.M.SG musk.NOM.M.SG

au kevarā

and screwpine.NOM.M.SG

“They made sandal wood platforms (cabutara) and smeared them with aloe, musk and screwpine.”

4) hāru gaṁvāi so aisehin rovā heri
Necklace.NOM.M.SG lose.CVB that like that cry.3.SBJ search.CVB

herāi lehu jaum khovā
make to search.CVB take.2PL.IMP if lose.M.SG.PST.PTCP

“Having lost the necklace why do you cry? If (it is) lost search and make (others) search for it.”
Moreover, research shows that in those languages converbs can only be formed from simplex verbs.\(^5\) What is more, marking of the main arguments depends on the light verb not on the converb used in compound.\(^6\)

Preliminary results show quite stable features of light verbs in early NIA. Additionally, LVs in investigated languages show the same features as LVs in Modern Hindi which would suggest a balanced development in LV constructions through the centuries.

References:


---

\(^5\) Nepali allows compound verbs to be in CVB form:

\[
\text{Bhāt} \quad \text{khā-i} \quad \text{sak-era} \quad \text{u} \quad \text{sut-na} \quad \text{gay-o}
\]

Food eat-ABS finish-CONV he/she sleep-INF.OBL go.PST-3SG

"Having finished dinner, he went to sleep." (Peterson (2002: 108))

\(^6\) However, in the early stages of languages the marking of agent in perfective domain is not stable. (See Stroński 2011) At least up to the 16th century – variation between marked and unmarked forms of A (S stands for an intransitive subject, A for a transitive subject and O for transitive object, following Dixon 1994: 6.)
The phenomenon of co-argument sensitivity in Indo-Aryan. A case study of the role of referential features in the argument structure of Kashmiri
Saartje Verbeke (FWO-Vlaanderen/ Ghent University)

Since Silverstein (1976) and Comrie (1978, 1989), referential hierarchies have been related to case marking issues. In Kashmiri, referential hierarchies seem to play a role in both the case marking as well as in the agreement pattern. First, the pattern of case marking in Kashmiri is dependent on the relation between the two main arguments of a transitive clause.

Table 1 Marking of animate O’s in the imperfective construction

<table>
<thead>
<tr>
<th>A 1st person</th>
<th>O 2nd person</th>
<th>O 3rd person</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOM</td>
<td>NOM</td>
<td>NOM</td>
</tr>
<tr>
<td>OBJ</td>
<td>OBJ</td>
<td>OBJ</td>
</tr>
</tbody>
</table>

This table makes it clear that the marking of O differs according to its relation with the A argument: a higher ranked A argument leads to nominative marking, a lower ranked A argument leads to objective O marking. This pattern fits in with the observation of Witzlack-Makarevich et al. (2016) and Gildea & Zúñiga (2016), who argue that the underlying assumption to define an alignment is the comparison of the three core arguments S, A and O in different clauses of one language, not the comparison of two arguments in the same clause. In this paper, I approach the case marking in Kashmiri from this angle, and argue that “co-argument sensitivity” is the driving factor behind the Kashmiri case marking, not the universal idea of a referential hierarchy.

The second way in which referential properties seem to influence syntactic features of Kashmiri is found in its pattern of verbal indexing. Kashmiri has pronominal suffixes added to the verb which are cross-references to pronominal arguments. In other words, the verb indexation is dependent on the referential category of person. Furthermore, there is the more complicated case of the ergative pronominal suffixes. Ergative suffixes are added to the verb when the ergative subject is a pronoun. However, they can also refer to an object argument. In an imperfective construction, O takes the nominative case when it is lower in person than the subject, or inanimate (see Table above). Under these circumstances, the verb is followed by a suffix cross-referencing the pronominal nominative O, and this suffix shares its form with the ergative suffixes.

(1) \( bi\ səzə-th\ tsi\ tôr\)
    I.NOM.SG send.FUT.1SG-2SG you.NOM.SG there
    ‘I will send you there’

(2) \( bi\ mar-an\ su\)
    I.NOM.SG kill.FUT.1SG-3SG he.NOM.SG
    ‘I will kill him.’

In these examples, the suffixes \(-th\) and \(-n\) respectively refer to a second person and third person O. Formally, they belong to the paradigm of ergative suffixes. They are obligatory.

On the surface it seems that the verbal indexing is determined by the hierarchy between the arguments, and the question arises of how to mark the different arguments when both are third person. In such constructions, one makes a typological distinction between the proximate and obviative argument. The two terms refer to the two core arguments of a construction as well as to the marking associated with the features they bear. The argument marked as proximate is the
argument highest on the prominence/salience hierarchy, the argument marked as obviative is the lower one (Henze and Zimmerman 2010: 26).
In other words, obviative markers indicate the not-salient nature of an argument (Aissen 1997). They show the opposition of one argument with the salient and topical one, the proximate argument. Nichols (2011: 661), describing the language Ingush, defines obviation as following: “Obviation is the obligatory ranking of third person nominal (nouns, pronouns) based on discourse function, syntactic relations, and semantic properties such as animacy.” This broad definition of obviation can be applied to the pattern found in Kashmiri. Then, the ergative suffix –(a)n in the following example is not only obligatory to show its difference from the nominative subject argument, as Hook and Koul (1984) argue (a definition in the negative), it is also obligatory because it indicates the obviative argument, the argument that is the least salient in the construction of all the arguments. Indeed, the suffixes only indicate animate arguments lower in person, or less prominent, than the A argument. In that sense, these suffixes fit the definition of obviative markers.

\[\text{(3) } na \text{ chu-s-an bi sui zanān ti} \]
\[\begin{array}{llllll}
\text{not} & \text{AUX.PRS-1SG-3SG} & \text{I.NOM.SG} & \text{he,NOM.SG} & \text{know.PTCP.PRS} & \text{and} \\
\text{na} & \text{sui me} & \text{not he,NOM.SG} & \text{I.OBJ.SG} \\
\end{array}\]

‘I don’t know him, and he doesn’t know me.’

The examples from Kashmiri seems to indicate that the concept of “obviative” is not unique to the Amerindian language family, it occurs in various forms in different language of the world, functioning as a signal of low salience.

References

Short synopsis:
In this study residents of Pithoragarh, a remote Himalayan town in North India, were recorded responding to a series of interview prompts which were designed to illuminate their attitudes towards English. The concept of bahar ‘outside, away’ emerges as one of the central themes from qualitative analysis which reveals how Pithoragarh inhabitants position themselves as backward vis-à-vis more proficient English speakers from outside their mountain region. Speaker narratives also present English as essential for getting a job and for affording both self-confidence and self-expression.

Detailed synopsis:
This study focuses on the residents of Pithoragarh, a remote Himalayan town in North India, and seeks to identify the dominant themes in speakers’ discourse about English. With a population of about 60,000, semi-urban Pithoragarh is a buzzing educational hub in a predominantly rural district. Most residents of Pithoragarh are native to the region, the Kumaon Hills, and speak either Hindi or the regional language, Kumaoni, as a first language. In addition, the town is home to tribal peoples native to Kumaon, long-term immigrant communities from Punjab and Nepal, and a recent influx of newer migrant labourers from Bihar. Hindi functions as a lingua franca for interactions between these different language groups as well as between speakers of different Kumaoni dialects. As one of the two languages of schooling, Hindi occupies a place of higher prestige than Kumaoni, with English at the top of the language hierarchy.

There has been a long history of language contact between English and both Kumaoni and Hindi in Pithoragarh, since the establishment of the Kumaon Regiment of the Indian Army in 1813. More recently, the preference for English-medium (EM) rather than Hindi-medium schools across North India (Chaise LaDousa, 2014) has resulted in the foundation and expansion of large numbers of EM schools in Pithoragarh. Hindi speakers in Pithoragarh show an increasing tendency to mix English and Hindi in colloquial speech, as has been attested in other parts of North India (Borowiak, 2012; Si, 2010; Parshad et al, 2016), although the style of mixing in Pithoragarh differs from that spoken in metropolitan cities.

It is against this backdrop of linguistic pluralism and increasing societal focus on EM education that this study was conducted.

A sociolinguistic questionnaire was used as a prompt during interviews in which participants selected using the ‘snowball’ method commented on the role of the English language in their lives and its relative importance to Hindi. Alongside questions on language use and educational background, participants were asked to state to what extent they agreed with twelve sentences designed to ascertain the value placed on English both personally and as a prestige language within Indian society. The interviews were conducted (by the author) in Hindi, in a location familiar to the speakers, and were digitally recorded. The recorded interviews were transcribed in ELAN and the transcribed texts subsequently analysed using a qualitative data analysis based on grounded theory (Glaser and Strauss, 1967/2009) in which texts were scoured for dominant themes and coded accordingly.
Thematic analysis of the interview transcripts reveals a universally positive view towards English, which is projected as essential for one’s future, regardless of the age of the respondent. One key theme that emerges is the construction of local identity by reference to the conceptual Other who live bahar ‘outside, away (from one’s home district)’, i.e., outside their mountain region. Speakers present themselves and their society as comparatively backward, and problematize the struggle to compete for jobs with people who live in other parts of India. Their discourse reveals an inflated view of the English spoken by a supposedly more developed population living bahar. Perhaps as the anecdote to this sense of inadequacy, English is strongly positioned as the means to self-expression and growth in confidence. These latter themes mirror attitudes towards English of final year students at a girls’ school in an underprivileged part of Delhi (Vaish, 2008).

One contrasting theme highlights the negative attitude sometimes shown towards people using English (or even Hindi) in the villages outside Pithoragarh, where speakers are accused of ‘showing off’. This corroborates findings by Gumperz on negative response to the use of a more prestige language in the village setting (1958:677).

Pithoragarh is at once unique in its location and linguistic constituency, and at the same time has enough comparable qualities to be considered representative of other remote semi-urban communities in India. As the only detailed analysis of interview data with participants from such an environment, this study contributes to our understanding of how people from remote less-developed regions of North India view English and their relationship to it.

This study is part of a broader PhD project investigating semantic differences in the use of English-origin words in the colloquial Hindi of Pithoragarh residents.

References


Echo-formation in Katias a neighbouring language of Kalasha and Khowar

Yoshioka Noboru (National Museum of Ethnology, Japan)

Kati is a Nuristani language mainly spoken in Afghanistan, but also in a few parts in Khyber-Pakhtunkwa province of Pakistan.

Echo-formation is a kind of partial reduplication, and is often thought as a characteristic of languages in the Indian Sprachbund. According to traditional definitions, echo-formation should make reduplicants, that is “echo words”, with some sound change at the word initial syllables of base words. In Yoshioka (2014, forthcoming), however, I presented that Khowar and Kalasha are languages which do not avoid identical forms between a base word with the /m/ initial consonant and an “echoword” of echo-formation, while they functionally do not differ from other echo-formation. And I advocated that we may have to revise the definition of echo-formation.

In this presentation I show you another example of echo-formation with identity allowance from the Kati language. This is a neighbouring language of Kalasha and Khowar, but a little differ from them in geneology, Kalash and Khowar are north western Indic, so-called “Dardic” languages. This is just an early report of my pilot survey on the language in 2016 autumn.

As a conclusion, identity allowance within echo-formation can be considered as an areal feature, not just as a geneological feature such as one for Chitrali subgroup of “Dardic” languages, because Kati is not an Indic language but adjacent to Kalasha at least in the Rumbur valley, where I have had the survey. This feature seems to be neither main stream in Indic nor found in Persian. It will be an interesting issue to know how widely this phenomenon stretches around and where is the origin.

References


Role clusters in Indo-Aryan
(diachronic and synchronic aspects)

Zakharyin, Boris (Moscow State University)
Khokhlova, Liudmila (Moscow State University)

The paper deals with history of semantic role clusters in the evolution of Indo-Aryan. Vedic and Sanskrit as variants of Old Indo-Aryan (OIA) were not role-dominated languages. There existed no straight-forward correspondences between semantic roles and case markers: one role might be marked by a number of cases and one and the same case might be used for different roles. The latter phenomenon will be the focus of our analysis.

The *nominative-accusative* type of alignment specific for OIA has determined the existence of the universal Agent/Patient cluster: any argument of semantically intransitive verb – be it Agent or Patient – is coded in the same way as the Agent of a transitive verb. According to Kibrik, this cluster may be defined as ‘hyper-role of ‘Protagonist’, the main participant of the situation.

There are non-specific clusters in OIA which can be found in many other nominative-accusative languages. E.g., OIA used the same case markers for Agent/Experiencer and Patient/Stimulus clusters. This can be explained as the implementation of the economy principle: Agent and Experiencer, on one hand, Patient and Stimulus, on the other, always belong to different syntactic structures and thus do not need marking by different cases. For Passive Agent/Instrument cluster the semantically based explanation may be suggested as both roles imply the sense of the cause of action. Also trivial is the Active Agent/Passive Patient cluster typical for nominative-accusative alignment.

The other OIA clusters, being less predictable, deserve special attention. Among them are: Agent/Patient cluster (only in feminine plural nouns), Stimulus/Source, Stimulus/Locus, Stimulus/Possessor, Patient/Goal, Recipient/Theme, Theme/Locus, Possessor/Recipient, Passive Agent/Source. Special attention will also be paid to clusters characterizing OIA pronominal enclitics (which function differently from their full counterparts): the Possessor/Recipient cluster characterized the pronominal enclitics of singular in OIA and in Middle Indo-Aryan (MIA) Pāli. For plural enclitics in Pāli the Patient/Instrument cluster is also significant, and in Ardha-Māgadhi (belonging to the later stage of MIA) the consolidated Patient/Instrument/Possessor/Goal cluster became typical for all pronominal enclitics.

The destruction of the old verbal and nominal systems of inflections in IA resulted in appearance of the two main syntactic patterns: the ergative one and the nominative-accusative one, both characterized by partially overlapping role clusters. Thus, in Apabhramsha, representing Late MIA, the Instrument/Locus and Source/Possessor/Recipient clusters were present in ergative and nominative-accusative domains. Some clusters were distributed across different domains: the hyper-role of Protagonist remained in nominative-accusative domain, while a new cluster was formed in ergative domain. It implied the combination of Patient of semantically transitive verbs with any argument (Agent or Patient) of semantically intransitive verbs. This cluster is usually defined as ‘hyper-role of Absolutive’, and, according to Kibrik, it implies ‘the immediate, nearest, most involved or most effected participant of the situation’ (Kibrik 1997: 292).

The hyper-role of Absolutive is wholly predictable by ergative alignment itself, while the Agent/Instrument cluster was specific for ergative domains of Apabhramsha and Early NIA, that is for languages that signified the climactic point of Indo-Aryan ergativity in general. Later on this cluster disappeared in most of New Indo-Aryan languages, but in some of them it got preserved, e.g., in Gujarati which had retained maximal amount of ergativity features.
The ergative domains of Ardhamagadhi, Apabhramsha, Old Punjabi had the unusual Agent/Patient cluster in paradigms of 2-nd person pronouns [Bubenic 1998: 89]. The same cluster in ergative domain also occurred in noun paradigms of some languages of Early NIA. Later on it has been eliminated from ergative domains of the majority of Western NIA having remained only in Rajasthani. The starting point of this elimination process were personal pronouns. The decrease of the ergative alignment in Western NIA resulted also in appearance of the Patient/Recipient cluster first in non-ergative and later in ergative domains. The development of non-demotional Passive Voice has resulted in forming the Passive Agent/Source cluster in Modern Punjabi and Gujarati.

For future research it seems useful to analyze possible historical links between role clusters existing at different stages of Indo-Aryan evolution – interesting in this respect is the phenomenon of Agent/Source cluster which had occurred in OIA (though rare there) and was retained in Old Punjabi.

References:


Kibrik, A.E. 1992. Ocherki po obschim i prikladnym voprosam yazykoznaniya (= Essays in problems of general and applied linguistics), Moscow: Moscow University, 1992.


A preliminary study of the Tamil comparative postpositions

Wojciech Zeyland, Jagiellonian University

An English sentence 'She is taller than him' is an example of comparison of inequality (Heine 1997, Stassen 1985, Ultan 1972). Such constructions may be expressed in Tamil language in a number of ways. A few of them employ postpositions, namely some specific forms of the three following verbs: pār 'see', kāṭṭu 'show' and viṭu 'leave'. In order to express a comparison they are placed immediately after the standard of comparison marked for accusative case (see Andronov 1969, Arden 1910, Rajendran 1976). Since Tamil postpositions in general have not been well studied in a historical perspective, the question of the history of these particular postpositions as comparative postpositions remains unanswered. In my presentation I shall discuss the following problems with regard to aforementioned forms and their origin: when do they appear in texts for the first time? What were the conditions that favoured the choice of these three words in their respective forms as comparative postpositions? What can be said about the distribution of these postpositions? Are they characteristic of only some registers of the language? Are they freely interchangeable? Does any of the forms visibly predominate in Tamil texts? The results of my work with the corpus of Tamil language as embodied on the Project Madurai website will be presented.

References