

Some areal features of the 110-item wordlist for the Indo-Aryan languages of South Odisha¹

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

Keywords: Indo-Aryan languages; Dravidian languages; Munda languages; basic lexicon; lexicostatistics; areal linguistics.

The aim of this study is to describe a number of specific features of the *basic vocabulary* contained in the 110-item Swadesh/Starostin/Yakhontov wordlists in the context of an entire language area (*Sprachbund*), primarily focusing on the data of the *micro-language-area* of South Odisha (India) and its position within a larger area — East India as a whole (in the text below I shall refer to East India as a *mini-area*, since the term ‘area’ would be more suitable to describe the entirety of South Asia as a whole). In these regions, three different language families are represented: Indo-Aryan, Dravidian, and Munda. Speakers of these languages are all settled in adjacent locations, where different language idioms mutually influence each other.

The importance of studying the history of basic vocabulary, such as the 110-item word list, in the context of a language area, stems from the following reasons: 1) the list contains basic vocabulary that is, as a rule, least affected by semantic change and lexical replacement; 2) its items are commonly used in glottochronology to estimate the approximate date of splitting of different language idioms.

According to one methodology of lexicostatistical evaluation of the degrees of genetic relationship between different languages, developed by Sergei Starostin and commonly employed in the Moscow School of comparative linguistics, borrowings from other languages should not be taken into account for such an evaluation, since the number of borrowed words can drastically surge during periods of intensive language contacts, whereas replacement of words from the basic wordlist with other indigenous words is assumed to proceed at a more or less regular rate, independently of language contacts. My goal here is to investigate the possibility of lexical replacement in the basic wordlist with original words from the inherited lexical stock that is nevertheless still driven by language contact.

Although the language map of South Asia (Fig. 1 in the Supplement) shows languages from a variety of genealogical groups, my research concerns only three language families: Indo-Aryan, Dravidian, and Munda. The Indo-Aryan languages are spread throughout the

¹ This research has been supported by the Russian Foundation for Basic Research, project 17-34-00018: “Sociolinguistic research of the Koraput Munda languages and development of a multimedia corpus of texts exemplifying different social conditions of language usage”.

Northern part of the subcontinent, while the southern part is largely occupied by Dravidian languages. In the north, Dravidian languages are represented by two isolated areas: Brahui in the west, Kurukh and Malto in the East. The Munda languages, belonging to the Austro-Asiatic family, are spoken in three isolated regions: border areas of the states of Odisha, Jharkhand and West Bengal in the north, the border of Odisha and Andhra Pradesh states in the south, and the border of Madhya Pradesh and Maharashtra states in the west. Archaeological and historical linguistic data allow us to assume that Dravidian and Munda languages may have occupied much larger territories in the past.

We collected data on the languages of the South Odisha micro-area during field trips to Odisha in between 2016–2018². Different datasets were collected, but one of our specific tasks was to assemble 110-item wordlists of basic lexicon. As far as possible, those lists were put together according to the specifications of semantic contexts proposed in Kassian et al. 2010, Dybo 2013.

The languages studied in this paper include³:

- 1) Munda, including two dialects of Sora, two dialects of Bonda, Didayi, Gutob, Ho and Mundari;
- 2) Dravidian, including Koya and Kui;
- 3) Indo-Aryan languages and dialects, including Desiya, Pano (two dialects), Oriya, Bengali, Assamese, Hindi, Punjabi, Lahnda, Sindhi, Rajasthani, Mewati, Gujarati, Marathi, Konkani, Kullui, Garhwali, Kumaoni, Nepali, Braj, Awadhi, Maithili.

Since my research is chiefly focused on Odisha, I need to provide some explanations regarding the idioms of this state, and specifically of South Odisha (Fig. 2 in the Supplement). Sora, whose dialectal diversity has so far not been studied in sufficient detail, is represented by Lanjiya and Sarda dialects. Of these, Lanjiya is the most popular idiom and has been extensively studied by the majority of researchers; we, however, take into account the importance of materials from other dialects as well, primarily Sarda. Lanjiya is spread in Gajapati district and Sarda belongs to Rayagada district.

Bonda is spread in the Malkangiri district. We have collected data on the Lower and Hill dialects, of which the Lower or Plain Bonda is more influenced by widespread languages than Hill Bonda. Didayi is also located in the Malkangiri district; Gutob is native to the Koraput district. All of these belong to the South Munda group. Ho and Mundari languages belong to the Kherwarian subgroup of the Northern Munda group and are spread in several northern districts of Odisha and some districts of Jharkhand.

Of the Dravidian languages, Kui is situated in the Kandamahal district and Koya is a dialect of Gondi spoken in the Malkangiri district.

Finally, of the Indo-Aryan languages, Oriya is the state language of Odisha, belonging to the Eastern group. Desiya and Pano are colloquial idioms, close to Oriya. Existing literature on Desiya is extremely scarce. It is a lingua franca for communication between caste groups and tribes in the districts of Koraput and Malkangiri. It is widely believed to be a pidgin rather than a native language for any of the caste groups and tribes living there. However, according

² I have to acknowledge here Evgeniya Renkovskaya, my constant companion and colleague, without whom this work would not be possible, for participating in the collection of data and for her sharp insights on the subject.

³ The wordlists for Sora, Bonda, Didayi, Gutob, Ho, Mundari, Kui, Koya, Kullui, Oriya and Bengali languages were collected during our field trips in 2014–2018. For other Indo-Aryan languages (Hindi, Punjabi, Lahnda, Sindhi, Rajasthani, Mewati, Gujarati, Marathi, Konkani, Garhwali, Kumaoni, Nepali, Braj, Awadhi, Maithili) we have used the lists from Kogan 2016, with corrections made in Krylova 2017.

to our field data, this is not the case, since there are influential castes, such as the Rona, for which Desiya is native at the moment. Nowadays, the Gutob tribe is switching to Desiya, and rapidly losing its native language. As for the Pano idiom, no one has ever studied it properly. It is spoken by a large caste group in the districts of Gajapati and Rayagada, adjacent to the Sora language (they are spoken next to each other in many villages). It is closer to Oriya than Desiya, but its wordlist still has some significant differences.

According to historical evidence by Piers Vitebsky (2017), during the period when most of the Sora people were monolinguals the Pano caste group occupied the niche of interpreters in the communication between Sora and Oriya-speaking government officials. In communications between the officials of Koraput and local tribes, the main language in use is Desiya, and Desiya-speaking interpreters play a key role, since most of the tribal population speaks Desiya, but not Oriya. From this point of view, Pano and Desiya are in similar positions. We have assembled two 110-item wordlists in the Pano language from the Gajapati and Rayagada districts, because it was not previously clear how the two Pano dialects differ from each other, especially since in two different districts, speakers of Pano language are neighbours to speakers of two different Sora dialects. As for the other Indo-Aryan languages taken into account, I have included these to provide a broader base of comparison.

Variations in basic vocabulary under the influence of language contacts (as in the case of the specified areas) may be traced through borrowings, partial borrowings and replacements of some original words by others (including replacements from inherited vocabulary) under the influence of language contacts. The most obvious example of mutual influence between languages in the vocabulary is, of course, direct borrowing. In the area of intensive language contact, the number of borrowings may be very significant even in the basic vocabulary. For example, I found 18 clear loanwords from the Indo-Aryan and 7 from the Dravidian family in the Munda languages of South Odisha; two Dravidian languages include 20 Indo-Aryan and 1 Munda loanwords; and in the Indo-Aryan idioms of the mini-area, 10 Dravidian and 2 Munda words were discovered. Borrowings from Indo-Aryan languages in Dravidian and Munda are more numerous than borrowings from Dravidian to Munda. Overall, the main source of borrowing is predictably identified as the Indo-Aryan languages; next in frequency are the Dravidian languages, while the contribution of Munda is insignificant.

Indo-Aryan loanwords in Kui: *kali* ‘black’, *raka* ‘blood’, *puri* ‘full’, *loku* ‘person’, *coka* ‘round’, *balu* ‘sand’, *bela* ‘sun’, *dhuali* ‘smoke’, *poheri* ‘swim’, *jorek* ‘two’, *duru* ‘far’, *basari* ‘year’.

Indo-Aryan loanwords in Koya: Koya *soben* ‘all’, *padaj* ‘foot’, *manej* ‘man’, *mansun* ‘man’, *guṇraj* ‘round’, *panqe* ‘white’, *duram* ‘far’, *worsa* ‘year’.

Indo-Aryan loanwords in Munda languages of the South Odisha: Sora *sana* ‘small’, Bonda *kaliya* ‘black’, Gutob *kakor* ‘cold’, Gutob *bosa* ‘fat’, Gutob *urei* ‘fly’, Gutob *jibon* ‘heart’, Gutob *manus* ‘person’, Gutob *besi* ‘many’, Gutob *ron* ‘red’, Gutob *jorek* ‘two’, Gutob *tati* ‘warm’, Didayi *tara* ‘star’, Bonda, Didayi *borti* ‘full’, Bonda, Didayi *bol* ‘good’, Gutob, Didayi *buk* ‘breast’, Bonda, Didayi *patla*, Gutob *potla* ‘thin’, Bonda *borso*, Didayi, Gutob *boros* ‘year’.

Dravidian loanwords in Munda languages of the South Odisha: Sora *kob/kumab* ‘ashes’, Gutob *ponrei* ‘swim’, Didayi *dula* ‘lie, sleep’, Bonda *piri*, Gutob *piri* ‘bird’, Didayi, Gutob *chendi* ‘hair’, Sora, Bonda, Didayi, Gutob *maṇḍi* ‘knee’, Bonda, Didayi *monji* ‘seed’.

Munda loanwords in Kui: *gule* ‘all’.

Dravidian loanwords in Eastern Indo-Aryan languages: Desiya *gar* ‘egg’, *chendi* ‘hair’, *maṇḍi* ‘knee’, *dul-* ‘lie’, Desiya, Pano *cher* ‘root’, Desiya, Pano (Gajapati) *āṇdra* ‘man’, Pano (Gajapati) *sok* ‘neck’, *guṇḍiya* ‘heart’, Oriya, Pano, Desiya *monji* ‘seed’, Oriya, Assamese, Bengali *kāmūṛ-* ‘bite’.

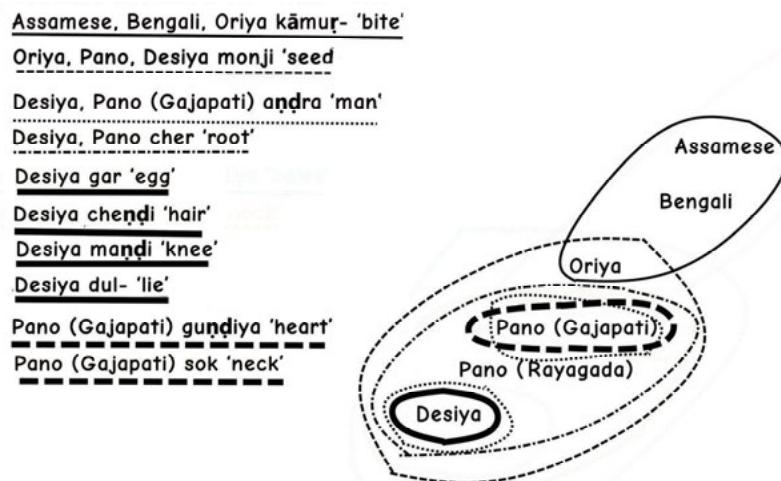


Figure 3. Dravidian loanwords in East Indo-Aryan languages

Figure 3 illustrates the spread of Dravidian borrowings across the 110-item wordlists of Eastern Indo-Aryan languages. It clearly shows that their number reaches its maximum in South Odisha.

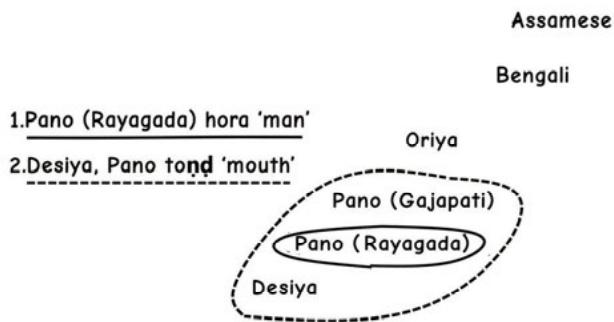


Figure 4. Munda loanwords in East Indo-Aryan languages

Borrowings from Munda to Indo-Aryan are much fewer in number: Pano (Rayagada) *hora* 'man' and Desiya, Pano *tond* 'mouth'. Here it is quite instructive to pay special attention to the latter. This root is widely represented in IA languages, and, according to Turner 1966, it had already been borrowed into Proto-Indo-Aryan:

5853 *tundā* n. 'beak, trunk, snout' TĀr. MBh., 'contemptuous term for mouth' Bādar, *tūṇḍika* 'having a trunk' AV. [Cf. *trōṭi* f. 'beak, fish's mouth' lex., and further *śuṇḍā*, *cañcu*, **cōṇṭa*. Cert. non Aryan and prob. ← Mu. EWA i 510. Poss. same as *tundi* 'navel' and *tundā* 1 'belly' with which it shares several variant forms. Some of these may be due to differences of dialect in the source language, some to phonetic development within IA. (e.g. *t̪* > *t̪̪*), some to slang deformation]

1. *tuṇḍa*: Pa. *tuṇḍa*, **aka* n. 'beak, mouth, snout'; Pk. *tūṇḍa* n., **đī* f. 'mouth, front part'; S. *tūṇī* f. 'spout'; Or. *tuṇḍa* 'lips, mouth, beak', *tuṇḍi* 'muzzle', (Sambhalpur) *tūḍ* 'mouth', M. *tūḍ* n.; Si. *tuḍa* 'beak, snout, mouth'.
2. **tuḍa*: Si. *tola* (st. *tol*) 'lip', *tolu gānavā* 'to graze (lit. rub the lips on)'. 3. **tōṇḍa*: Pk. *toṇḍa* n. 'mouth, front part'; M. *tōḍ* m. 'mouth', Ko. *toṇḍa*. 4. **tōṭṭa* 1: Or. *toṭi* 'throat'; M. *toṭi* f. 'spout'. 5. **tōṇṭa*: K. *tōth* (dat. **ti*) f. 'beak' (← Ind.?). 6. **tutta*: K. *tutu* 'long faced'; WPah. bhal. *tutt* n. 'face', bhad. *tuttar* n. 7. **tuṇḍa* 1: N. *tūro* 'mouth of animal, beak, point of ploughshare'; Or. (Jānpur) *tuṇḍa* 'lips, mouth, beak'; Bi. *tūṛ*, **rā* 'beard of wheat' (semant. cf. *sūṛ* s.v. *śuṇḍā* 1). 8. **tuṇṭa* 1: A. *tūṭli* 'fleshy protuberance under chin'; B. *tūṭi* 'throat'. 9. **tōṇṭa* 3: K. *tūṭi* f. 'spout' (← Ind.?). 10. **tūṭṭa*:

P. *ṭūṭī* f. 'spout', N. *ṭuṭo*, **ti*. 11. **ṭōṭṭa* 3: S. *ṭoṭī* f. 'mouthpiece of hookah'; H. *ṭoṭī* f. 'spout'; G. *ṭoṭī* m. 'throat'. 12. **thuṇḍa*: L. *thunn* m. 'thick lips (opprobrious term)', awāṇ. *thun* 'lips'. 13 **thuntha*: A. *thūtari* 'chin', B. *thūti*, *thūtani*; Or. *thūti* 'chin, snout'. 14. **thuttha*: P. *thūthṇā* m., *thūthṇī* f. 'mouth of horse or camel'; N. *thutunu* 'nose, snout'; A. *thutari* 'chin'; B. *thuti*, *thutani* 'chin, mouth of animal'; Or. *thuthi*, **ti* 'snout, chin, beard'; Bhoj. *thuthun* 'pig's snout'; Mth. *thūthun* 'mouth of animal'; Aw.lakh. *thūthun* 'horse's nostrils'; H. *thūthṇā*, *thūthān*, *thūthṇā* m. 'snout, mouth'. 15. **thōḍa* 1: Or. *thōṛī* 'lower lip, chin, beard'. 16. **thōḍḍa*: L.mult. *thōḍ* m. 'lip', (Ju.) *thōḍī* m., *thōḍī* f. 17. **thōṭṭha* 1: Mth. *thothī* 'mouth'; H. *thoth* m. 'snout', *thōṭī* f. 'forepart of animal's face'; M. *thotṛī*, *thodṛī* f. 'side of cheek'. 18. **thōṇṭha* 1: B. *thōṭā* 'chin'. 19. **thēṇṭha* 1: Or. *thēṇṭa*, *thāṇṭa* 'beak'. 20. **thuṇḍa*: Ku. *thūn* m. 'beak, elephant's trunk'; N. *thīṛō* 'beak'; Or. *thuṇḍī* 'lip, chin'; G. *thūḍī* f. 'chin'. 21. **ṭhūḍḍa*: P. *ṭhūḍḍā* m. 'beak of a paper kite'; H. *ṭhūḍḍī*, *ḍḍhī* f. 'chin'. 22. **ṭhōṇḍa*: H. *ṭhōṛī* f. 'chin'. 23. **ṭhōḍḍa*: S. *ṭhōḍī* f. 'chin', P. *ṭhōḍī* f., H. *ṭhōṛī*, *ṛhī* f. 24. **ṭhōṇṭha* 1: A. B. *ṭhōṭī* 'beak'; Or. *thuṇṭhi* 'lip, chin, beak'; Mth. H. *ṭhōṭī* f. 'beak'. 25. **dūṇḍa*: L.mult. *dūnn*, (Ju.) *dūn* m. 'wild pig's snout'. 26. **dūṇḍa* 1: K. *ḍōṇḍu* m. 'front of face'. 27. **dutta*: Dm. *dut* 'lip'; Woṭ. *dut* f. 'mouth'; Bshk. *dut* 'lip' (Leech "dūḍh" < **dūḍḍha* ?), Tor. *dūt*; Phal. *dut* 'mouth' (*dhut* < **dūḍha* ?); Sh.pales. *dūt* 'lip'. Forms with *r* (cf. *trōṭī* f. 'beak' lex.) in Sh.koh. *turūṭī* 'lip', jij. *thurūṭī*, gil. *thūṛū*, *ṭū* m. 'beak' (Turner 1966: 334).

In Munda languages, the root is well represented and can be traced back to the Proto-Austro-Asiatic level:

393. Kh *to'mo'd* 'Mund (mouth)', Ju *tamar*, *tamon(i)*, *tomor*, Gu *tummo*: PaV *tovo*: Re *to:m-* (KF), Pa *to:t*, So **t?o:d-ən*, *t?u:d-ən*, KF *'tam-ən* id. Pal *mu:r*, Kyusao *mu:-ε*, Omachawn *mu-ጀt*, Rum *mo:i*, Hupawng *mwō-i*, Pangnim *mot* 'mouth', Khm *măt* (*meăt*) 'bouche, gueule, bec', Sem. Jur. New. *temut*, U. Cher, *to-nut*, Sem. Beg. *lănuđ*, Sem. Martin *tenót*. U. Tem. *kə-n-ut*, Or. Hu. Joh. II. *snut*, U. Ind. *nut* 'mouth' usw. (Pinnow 1959: 178).

1001 **t¹ət* 'mouth'. A: (Khmer, Palaungic, Aslian) ~ Jehai *tənəd*, Semnam *təni:d*, Semaq Beri *kənət*; ~ (**t¹mat* >) Khmer *mət* (contaminated by 1045 **mat* eye?), Palaung dialects *muጀt*, *mot* <...> **t¹l* in view of Munda, below. <...> Sora **t?o:d-ən*, **t?u:d-ən*, &c., ~ Kharia *t¹mo'd*, &c. (Shorto 2006: 287).

Evidently, this word has to be considered as primordial for the East Indo-Aryan level. However, as a rule, in modern Aryan languages it is represented in less basic meanings, such as, 'bird's beak'. In the basic meaning 'mouth', it replaces the descendants of Old Indo-Aryan *mukha* in southernmost Odisha and, apparently not coincidentally, also in Marathi (*tōḍ* 'mouth'), which is in contact with the area where Korku is spoken. Two explanations are plausible: 1) a secondary loanword in New Indo-Aryan languages; 2) as a result of semantic shift, Common Indo-Aryan *mukha* 'mouth' was replaced by descendants of Proto-Indo-Aryan *tuṇḍa* 'beak, trunk, snout' under the influence of neighbouring Munda languages.

Partial borrowing

Partial borrowings are rare in basic vocabulary, and are more likely to represent isolated cases, though they do occur. See, i.e. Gutob *pad* *susuj* 'foot', formed from Desiya *pad* 'foot' and Gutob *susuj* 'leg'. The Gutob word *susuj* means both 'foot' and 'leg', without any lexical differentiation of the two meanings. In Desiya (and Oriya) they are expressed by different words: *pad* (*pədɔ*) 'foot', *god* 'leg'. Thus, areal influence can lead to splitting of previously inseparable meanings and to partial borrowing.

Original words that replace basic lexicon under the influence of language contact

As stated in Normanskaya 2005: 148–149, the colour designation system in the Vedic language did not transparently distinguish between 'green' and 'yellow'. The words *pīta-*, *pītaka* were used for 'yellow', whereas *hári-*, *hariṇá*, *harit-*, *harita-*, *haritála* could be used for both 'yellow' and 'green'.

For the Proto-Indo-Aryan language this is apparently not the case. In Sanskrit, the adjective *pītala* in the meaning 'yellow' is not fixed in texts, but is indicated only in lexicographical works. However, it is from this word that items with the basic meaning 'yellow' (included in the 110-item wordlists of the majority of Indo-Aryan languages) have developed:

8233 *pītala* 1 'yellow' lex. 2. **pīthala* - . 3. **pīvala* 4. **pīvada* - . [*pītala(ka)* n. 'brass' lex. ~ *pittala* 2: see *pittá*, *pīta* 2]

1. Pk. *pītala* 'yellow', Ap. *pīra*, Paš.ar. *piā lo*, dar. *pēla*, Woł. *pyēl* (m. and f.), Gau. *pīlū*, Sh.pales. *pī lo*, S. *piaro*, L. P. *pīlā* (P. also *pillā* 'pale, yellow'), B. *pīlā*, Or. *pīlā*; Bi. *pīar*, *piarā* 'yellow (of cattle)', *piyar*, *piyarā* 'half fired brick', (Shahabad) *piyar* 'gram with yellow pointed grain'; Mth. *pīar* 'yellow', *pīrī* 'yellowness', Bhoj. *pīlā* adj., OAw. *piyari* sb., lakh. *piar* adj., H. *pīlā*, OMARW. *pīla*, G. *pīlū* adj., *pīyal*, *pīli* f. 'painting women's foreheads esp. with red' (whence *pīlvū* 'to colour yellow, to dirty'). Bshk. *pāla* 'blue, green' (AO viii 307) rather < *palitā* or **palva*.
2. Gaw. *phialá* 'yellow', Sv. *phyaló*, Ku. *pīhalo*, *pīhalo*, N. *pa(h)ēlo*; Tor. *pihil* 'green' (but AO viii 307 doubts).
3. Pk. (Deśin.) *pīvala* 'yellow'; Ku. *pyūlī* 'a small wild flower', *pyuraṭī* 'a small yellow bird'; B. *piuli* 'yellow colour, the flower *Hypericum cernuum*'; Or. *piuli* 'yellow', M. *pīvlā*.
4. B. *piuṛi* 'yellow tawny colour'; Or. *piuṛi* 'yellow'.

**pītalatvanā* ; **pītalavarṇa*.

pītala 2 'brass' see *pittala* 2.

Addenda: *pītala* 1. 1. S.kch. *pīro* 'yellow'.

3. **pīvala*: WPah.ktg. (kc.) *piulɔ* 'yellow', J. *pyūlā*. (Turner 1966: 466)

While in Sanskrit *hari* means both 'yellow' and 'green', the basic adjective 'green' in most Indo-Aryan languages originates from its derivative *harita*. The meaning is more or less stable (if we exclude Dardic languages and Rigveda):

13985 *hārīta* 'yellow' RV., 'green' Śiś., *haritaka* 'greenish' MW., n. 'vegetable matter' Car. 2. *hārīta* 'green' W., *taka* n. 'vegetable matter' lex. [hāri]

1. Pa. *harita* 'green, fresh'; Pk. *hariya* 'green'; Kal. (Leitner) "hārīlek" 'green', rumb. *hā'rīla* 'brass', urt. *hari'ra* (or < *hārīta* ?); S. *haryo* 'fresh and green, luxuriant'; P. *hariā*, *harā* 'green'; WPah.jaun. *harō* 'verdant'; Ku. *hariyo*, *haro* 'green', N. *hariyo*; B. *hariyā* 'meat curry'; Or. *harā* 'green', OAw. *harī* (pl. m. *hare*), lakh. *harā* (*hariab* 'to become green'); H. *hariyā*, *harā* 'green', OMARW. *hariyaü*, G. *haryū*, *harū*. (Turner 1966: 809)

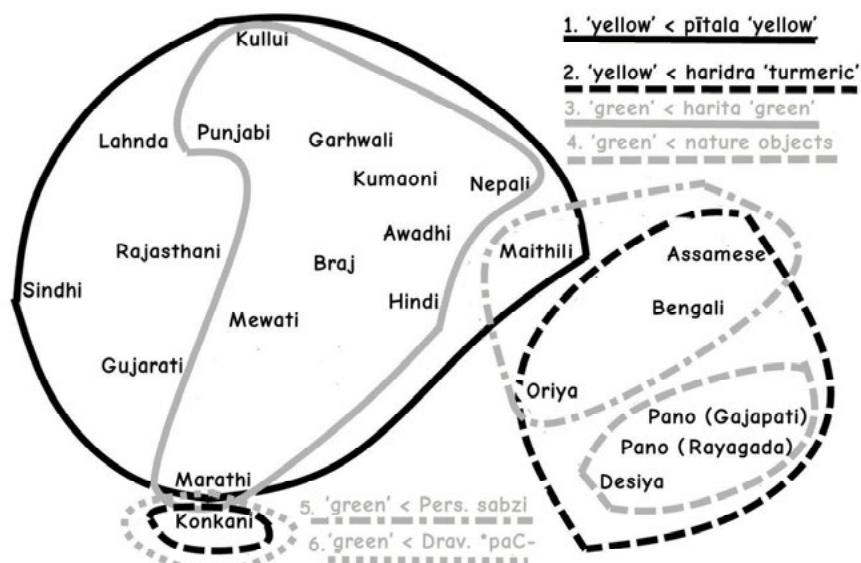


Figure 5. Origin of the basic words for 'yellow' and 'green' from Indo-Aryan languages

Fig. 5 illustrates the origin of basic words for ‘yellow’ and ‘green’ from the 110-item wordlists in Indo-Aryan. We can see that isoglosses for the descendants of *pītala* and *harita* for the most part embrace the Indo-Aryan area, but do not reach over to Eastern India. In Eastern Indo-Aryan languages, *yellow* is derived from the words for ‘turmeric’ (Proto-Indo-Aryan *haridra*): Bengali *hōlud*, Assamese *halōdhiya*, Oriya *haḷadiā*, Pano *hōldiya* and Desiya *olqi*. The same situation is also present in all Munda and Dravidian languages taken into account: Sora *sāṣṣāj*, Bonda *sāṣṣāj*, Didayi *sia*, Gutob *sāṣṣāj*, Mundari, Ho *sāṣṣāj*, Kui *sīja*, Koya *kamka rāj*, *pasp rāj* (with the loanword *rāj* ‘colour’ as the second element).

Outside Eastern India, among Indo-Aryan languages the form *halduve* ‘yellow’ is related to ‘turmeric’ in Konkani, which is in close contact with the Dravidian languages.

As for the colour ‘green’, it is borrowed from Persian *sabz* ‘green’ in the north of Eastern India: Bengali *śobuj*, Assamese *xeuj*, Oriya *sabujā*, Maithili *sabuj*. In the south of Odisha, this colour designation is primarily associated with various natural objects. The colour prototype is taken from: (a) ‘leaf’: Pano *potro kolor*, Gutob, Didei *ola kolor*, Koya *aak rāj*; (b) ‘a kind of weed or moss’: Desiya *jogliya*, Sora *kului*, Mundari, Ho *gaded*, Kui *jiven* (meanings given by Sidwell 2009–2011). As for Konkani, *pācve* ‘green’ here is a loan from Dravidian.

A strong connection of ‘yellow’ with ‘turmeric’ (also called curcuma and Indian saffron) is observed in Dravidian languages, cf.: Proto-South Dravidian **arisana-* ‘saffron; yellow’, Proto-South Dravidian **manžal* ‘yellow, turmeric’, Proto-Gondi-Kui **sirij-* ‘yellow, saffron’ (Starostin 2006).

The same connection is also present in the Austroasiatic language family, cf.:

1052 **rmit*; **rmiit*; **rmiat* Curcuma species; yellow. A: (South Bahnaric, Mon, Khasi) Mon *mīt* turmeric, Proto-Nyah Kur **muṭ* <...>, Sre *rəmīt* greenish- yellow, Chrau *rəmūt* yellowish, Biat *rmut* saffron (→ Stieng *rəmət*), (or B) Khasi *lmit-lmit* rather [yellow]; ~ Khasi dialect *shyrmit* turmeric. B: (Palaungic, ?Central Aslian) Riang-Lang *_rəmīt* turmeric, yellow, (or A) Central Sakai *rəmīt* saffron. C: (Khmer, Katuic, Khasi, Central Aslian) Old Khmer *rmyat*, Modern Khmer *rəmīt*, *lmīt* saffron, Kuy *lmī:t* yellow, Khasi *lmet-lmet* very [yellow], Proto-Semai **rmeet* yellow <...> (Shorto 2006: 299).

While Proto-Indo-Aryan shows some stability in differentiating between ‘green’ and ‘yellow’, this is not the case in Proto-Dravidian, cf.:

(a) Proto-Dravidian: **paC* ‘green, yellow; fresh, raw’ > Proto-South Dravidian **pac-*/**pas-*, Proto-Telugu **pacc-*/**pas*, Proto-Kolami-Gadba **pac-*/**pas*, Proto-Gondi-Kui **pas*;

(b) Proto-Dravidian: **pūc* 1 ‘green’ 2 ‘yellow’ > Proto-South Dravidian **pūc*, Brahui *pūsh-kun* (Starostin 2006).

In some languages, the two meanings became differentiated later through additional lexical means, without the differentiation becoming obligatory, unlike in Indo-Aryan languages, cf., for instance in Telugu:

pacca I. n. 1 greenish yellow colour; *aaku* ~ green (like a leaf); *pasupu* ~ yellow (like turmeric) (Gwinn 1990: 312).

Austro-Asiatic terms for ‘green’ are not well differentiated from ‘blue’ and sometimes ‘yellow’, cf.:

1559b **t?iir*; **t?or* ‘blue, green’. A: (South Bahnaric, Khasi) Biat *e:r* ‘yellow’; ~ (**t[l]?**iir* >) Sre *talir* ‘blue, green’ (→ Röglai), Khasi *lir* ‘blue, indigo’ (compounding form?). B: (South Bahnaric, Khmuic) Biat *vr* ‘green’, Thin *?uṛ* ‘blue’ (Shorto 2006: 416);

1585 **jñaar* (> Pre-Palaung &c. **jjiər*, **jjaṛ*); **jŋ[ɔɔ]r* ‘green’. A: (Palaungic) Kammu-Yuan *cəŋá:r*, Palaung *ŋjər*, *ŋor*, Rieng-Lang *_ŋjer*, Lawa *saja* ‘green’, Khasi *[im]* *ngarngar* ‘very [green, unripe]’. B: (Khmuiic, Ni-

cobaric) Central Nicobarese *cuŋxa* 'blue, green', Nancowry *cuŋuá*; ~ Kammu-Yuan [cəŋá:r] *cəŋɔ:r* 'green'. Vocalism obscure. Cf. Central Nicobarese *ŋə:-, -ŋə:* 'yellow', Nancowry *ŋá* (Shorto 2006: 422).

My interpretation of Eastern Indo-Aryan innovations in colour terminology is as follows: when native Munda and Dravidian speakers of East India began switching to Indo-Aryan dialects *en masse*, the unusual colour designation for 'yellow' was replaced by a calque, based on the meaning 'turmeric'. At the same time, a separate colour designation for 'green' was lost until the era of mass borrowings from Persian in the Middle Ages, when a Persian name was introduced for this colour. In the preserved Dravidian and Munda languages, new descriptive colour designations for 'green' were occasionally developed. These designations were derivatives from or words coinciding with the names of natural objects, such as the leaf or a weed (exactly, *Hydrilla verticillata*). Since then, small Indo-Aryan idioms of South Odisha, under the influence of Munda and Dravidian languages of the substrate, have also developed new calque-based colour designations for 'green', based on natural objects. Thus, this is an example of mutual influence between language families, which occurs in several stages. For Konkani, I can suggest that the word for 'yellow' was replaced by *halduve* under South Dravidian semantic influence, and that later the descendant of *harita* 'green' was replaced by a Dravidian loanword because of phonetic closeness of the original word for 'green' to the new word for 'yellow'.

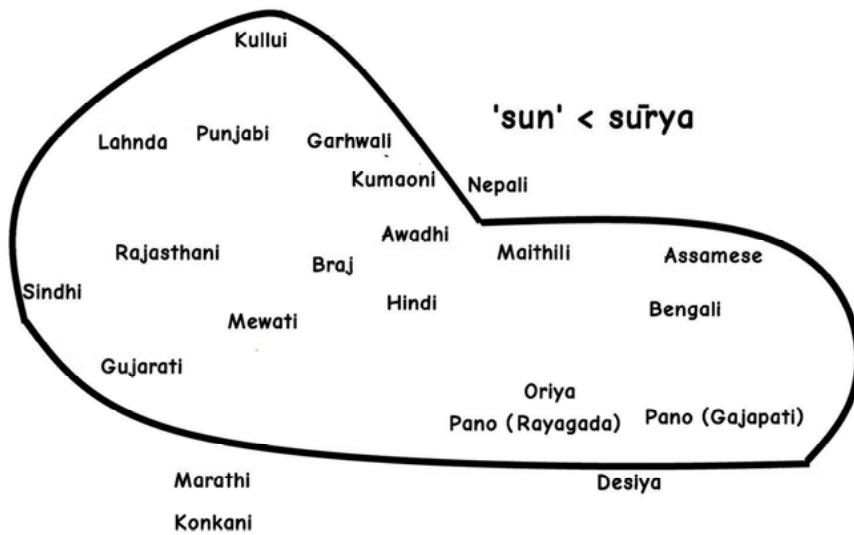


Figure 6. Descendants of *sūrya* 'sun' with the same meaning

The next example to be examined refers to the meaning 'sun'. In the vast majority of Modern Indo-Aryan languages, the main term for 'sun' goes back to Proto-Indo-Aryan *sūrya/sūriya* (Fig. 6), although it has many synonyms, primarily Sanskritisms.

However, we can see that in Desiya the lexeme *bel* is used with the same meaning. It is obviously related to Kui *bela* with the same meaning. However, the root is widely represented in various Indo-Aryan languages, as reported in Turner's dictionary:

12115 *vēlā* f. 'limit, boundary, time' ŠBr., 'flood tide' MaitrUp., 'shore, opportunity' MBh. 2. **vēḍā*. 3. **vīlī* 'tide, time'. [d appears in K. and Or. and poss. in cmpd. **ardhavēḍa*. In N. Mth. Bhoj. Aw. H. forms with r could derive from d, but more prob. represent the EH., Bi. or Mth. change of l > r. But the forms of *vēlā* and **savēlā* with -r-<-> occurring in P. Ku. G. M. (as well as those already mentioned) may be from an old **vērā* ~ **vēlā* rather than all be loans from the East Hindi Bihari area.<...> 1. Pa. *vēlā* f. 'shore, time'; NiDoc. *velanīmi*

loc. sg. m. ‘time’ (*yāñvēla* ‘when’, *tāñvēla* ‘then’); Pk. *vēlā* f. ‘shore, tide, time, opportunity’ (whence *vēlāvāi* ‘is late’); Ash. Wg. Kt. *wēl* ‘time, season, year’; Sh. *bēl*, *bil* (following pres. part.) ‘at the time of’; K. *vil* f. ‘short space of time’, *vēla* m. ‘time, hour’; S. *vera* f. ‘time, delay’, *veri* f. ‘time, turn’; L. *vel* f. ‘leisure’, *velā* m. ‘time’; P. *vel* f. ‘length of an animal’, *velā*, *be* m. ‘time’ (*ber* f. ‘time’ ← H.), *bhaṭ*. *us* *wēlē* ‘then’, *ik bērī* ‘once’; WPah.bhal. *āñ belū* ‘untimely’, cam. *bēr* ‘delay’; Ku. *bel* f. ‘meal fixed for a time, food taken at one time, meal’, *ber* ‘time’; N. *ber* ‘time’; A. *beli* ‘time, turn, year’; Or. *bēla* ‘daytime’, *lē* ‘once during the day’, *lū*, *lū* ‘from the beginning’; Mth. *beri* ‘time’, *beriā* ‘afternoon’; Bhoj. *bēr* ‘time’; OAw. *bera* f. ‘time, turn’, lakh. *bēr*; H. *ber* f. ‘limit, season, time’, *biriyā* f. ‘space of time’; G. *vel* f. ‘time’; M. *vel*, *el* f. ‘seashore’, m.f. ‘time, leisure, half day’; Ko. *velu* ‘time’; Si. *vel* a. ‘seashore, sandbank’ <...> (Turner 1966: 702).

However, the meaning of this word, as a rule, is not ‘sun’, but rather ‘time span’, ‘day-time’. Exceptions are found in the Eastern Indo-Aryan languages Assamese and Oriya, in which one of the meanings is precisely ‘sun’: Oriya *bel* ‘time, time or half part of day, the sun’ (Praharaj 1931–1940: 5902), Assamese *beli* ‘time, the sun, turn, year, delay’ (Candrakānta 1962: 700).

Of course, this term is not the most widely used, thus it is not found in the basic wordlist. In general, the polysemy ‘time/sun’ is not widespread among Indo-Aryan languages, and in Sanskrit *vela* does not mean ‘sun’, but ‘limit’, ‘boundary’, ‘period of time’ etc. On the contrary, in Dravidian languages such a polysemy is very common, cf.:

3774 Ta. *nēram* time, season, opportunity. Ma. *nēram* sun, day, light, time, season, hour, turn; *nērattu* early, seasonably. Kod. *ne·ra* sun, time; *ne·rate* early in the morning. Tu. *nērē* id. Br. *dē* sun, sunshine, day, time (Burrow & Emeneau 1984: 337).

However, the word *vela* ‘sun, time’ in Dravidian languages is represented only at the Gondi-Kui level, cf.: ‘sun’ — Konda *podu* 1 <...> Cf. also *ve:la*, with polysemy: ‘sun/time’ <...> (Sova dialect variant: *ve:ra*); Pengo *ve:ra* 2 <...>. Polysemy: ‘sun/time’; Manda *ve:ra* 2 <...> Polysemy: ‘sun/time/period/daytime/season’; Kui *ve:la* 2 <...> Polysemy: ‘sun/time/hour/ occasion/season’. Balliguda dialect: *be:la* ‘moment, time’ <...>, but the meaning ‘sun’ comes out in compounds, e.g. *be:la pe:nu* ‘sun deity’, etc.; Kuwi *ve:ra* 2 <...> Polysemy: ‘sun/age’. Sunkarametta, Parja: *ve:ra* ‘sun’ <...> Additionally, cf. also *ga:ma* <...>, with polysemy: ‘sun/perspiration’ (Starostin 2001–2019).

Since it is not represented in other Dravidian groups, we can assume the word to be a local borrowing from Indo-Aryan languages. At the same time, the meaning ‘sun’, which became its main meaning in Desiya, and its secondary meaning in Oriya and Assamese, probably developed in Eastern Indo-Aryan languages under the influence of Dravidian.

Another possible substrate phenomenon that we may have been able to come across is the word for ‘woman’, derived from the word ‘mother’ in Desiya and Pano: *mai*, *mai*. A similar development in the meaning ‘woman’ may be seen in Gutob (*yan̥kwi*, a combination of *yan̥* ‘mother’ with the feminizing suffix *kwi*).

Cf. similar semantic shifts in Mon-Khmer languages:

113 *[?]bo? ‘mother’. A: (Mon, Viet-Muong, North & Central Aslian) Old Mon ‘bo’ /mbu?/ ‘mother’, Modern Mon *b3?* ‘female of animals’ <...>, Jehai *b?*, Temiar *bo?* ‘mother’; probably → Malay *ibu*, → Javanese *ibu*; ~ Muong *bq* <...> Vietnamese *vq* ‘wife’, Semnam *babo?* ‘wife, woman’, Jehai *babo?*, Temiar *babo?* ‘woman’. Or *mb- (Shorto 2006: 96).

However, in this case, I do not have sufficient data to make any strong conclusions.

In addition to the abovementioned cases, there are cases of less easily explicable areal substitutions in East Indo-Aryan languages, some of which need to be examined in more detail.

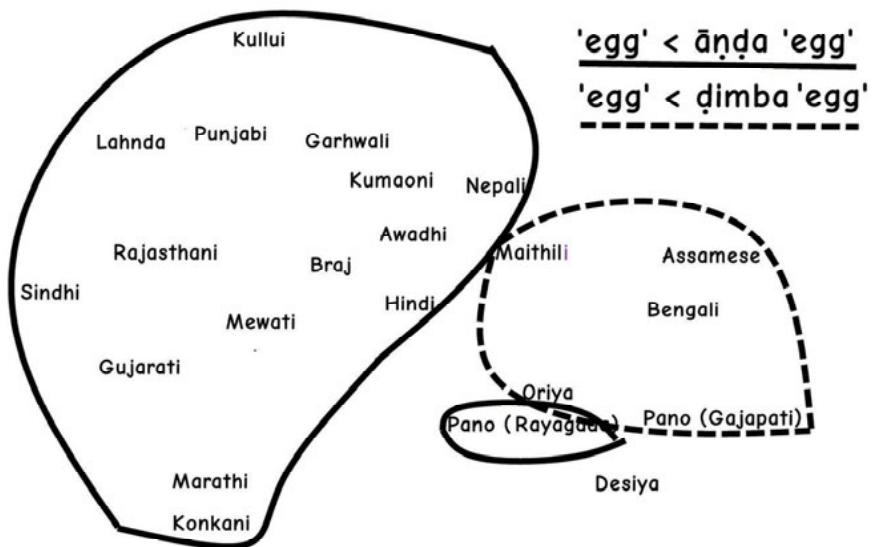


Figure 7. The origin of the words for 'egg' in 110-item wordlists of Indo-Aryan languages.

For example, in most Indo-Aryan languages the meaning 'egg' is expressed by descendants of Old Indian *āñ̥da*, while in Bengali, Assamese, Oriya and Maithili it is developed from *ḍimba*: Assamese *dima*, Bengali, Oriya *dim*, Maithili *ḍim* (Fig. 7). Curiously, Proto-Indo-Aryan *āñ̥da* is associated with the polysemy 'egg/testicle', while *ḍimba* is associated with the polysemy 'egg/foetus' polysemy, which brings it closer to the Proto-Dravidian term **karb-* 'egg, foetus':

Proto-Dravidian : **karb-* 'egg, foetus' > Proto-South Dravidian : **karu-* 'foetus' (Tamil *karu* 'foetus, embryo, egg, germ, young of animal', *karuppai* 'womb', *karuvam* 'foetus, embryo', Malayalam *karu* 'embryo, yolk', *karuntala* 'generation', Kannada *kandu* 'foetus of beasts'), Proto-Telugu **karv-* 'foetus' (Telugu *karuvu*, also *kari* 'uterus of animals', *karugu* 'an unopened ear of corn'), Proto-Kolami-Gadba **karb-* (Parji, *kerba*, pl. *kerbel* 'egg', Ollari Gadba *karbe* 'egg'), Proto-Gondi-Kui **garb-* 'egg' (Proto-Gondi **garb-* 'egg' > Koya Gondi *garba*, Su. *garbūm*) (Starostin 2006).

Although some of the listed Dravidian forms, contrary to the etymology, seem to have been borrowed from Old Indian *garbha* 'foetus' (Koya Gondi, etc.), others are more likely to represent an inherited item, or, at least, the results of contamination between an inherited root and the phonetically and semantically similar Indo-Aryan item. We can therefore suggest the polysemy 'egg/foetus' for the Proto-Dravidian level.

Among the 110-item wordlists that I have personally analyzed, this root is represented by Kui *garbaj* 'egg' and by the Dravidian borrowing *gar* 'egg' in Desiya. As to the descendants of Proto-Indo-Aryan *ḍimba*, it can be suggested that Dravidian influence may have been behind the choice of root in the basic lexicon.

The meaning 'knee' presents another example of this kind. Thus, the descendants of Proto-Indo-Aryan *goddā* 'ankle' with the meaning 'knee' are common in the middle part of the Indo-Aryan region, but not in the east, where the descendants of Proto-Indo-Aryan *aśthivāt* 'knee' dominate (Fig. 8). In Desiya, however, they are replaced by another Dravidian loan, *mañ̥di* 'knee'. Moreover, Dravidian *mañ̥di* is borrowed into all the Munda languages in the micro-area under study (Sora, Bonda, Dideyi and Gutob).

To conclude, it is extremely important to pay attention to the influence of substrate on basic vocabulary in the case of a language area, since the semantic shifts which occurred under this influence, may have a critical effect on evaluating the degree of genetic relationship between

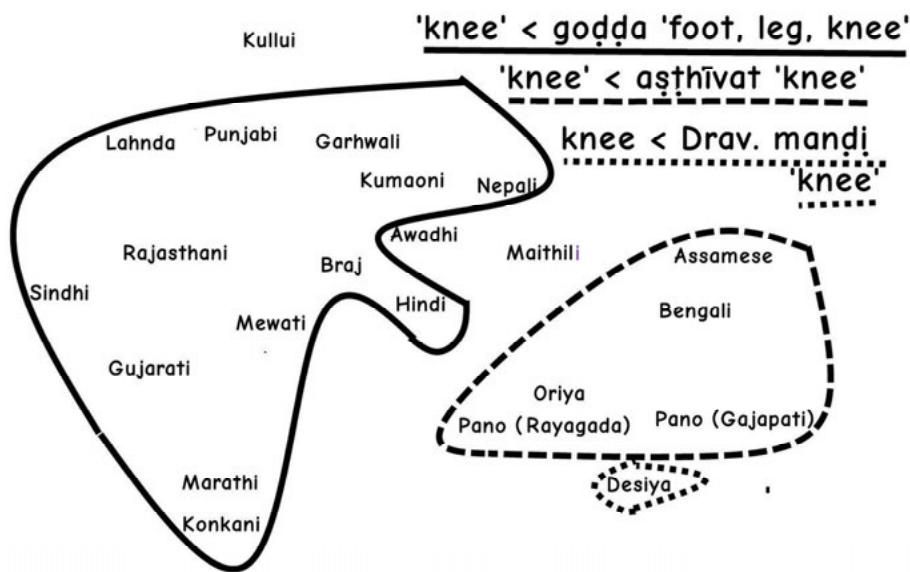


Figure 8. Origins of the words for 'knee' in 110-item wordlists of Indo-Aryan languages.

languages. Here, we can see how Eastern and Southern Indo-Aryan languages show some cases of replacement of words from basic vocabulary by another primordial words under Dravidian and Munda influence. In South Odisha, the area of very close communication between speakers of three language families, such substitutions are observed immediately in real time.

Additional materials for the paper (maps) are available in the online Supplement at: <http://jolr.ru/jlr11/krylova.zip>

References

Burrow, Thomas, M. B. Emeneau. 1984. *A Dravidian etymological dictionary*. 2nd ed. Oxford: Clarendon Press.

Dybo, Anna V. 2013. *Etimologicheskij slovar' tyurkskikh yazykov. Tom 9 (dopolnitel'nyj): Etimologicheskij slovar' bazisnoj leksiki tyurkskikh yazykov*. Astana.

Gwynn, John P. L. 1991. *A Telugu-English dictionary*. Delhi/New York: Oxford University Press.

Kassian, Aleksei, George Starostin, Anna Dybo, Vasily Chernov. 2010. The Swadesh wordlist. An attempt at semantic specification. *Journal of Language Relationship* 4: 46–89.

Kogan, Anton. 2016. Genealogical classification of New Indo-Aryan languages and lexicostatistics. *Journal of Language Relationship* 14: 227–258.

Krylova, Anastasia. 2017. Lexicostatistics and the New Indo-Aryan languages. *Journal of Language Relationship* 15: 279–298.

Normanskaya, Yuliya V. 2005. *Genezis i razvitiye sistem cvetooboznachenij v drevnih indoevropejskikh yazykakh*. Moskva: Institut jazykoznanija RAN.

Pinnow, Heinz J. 1959. *Versuch einer historischen Lautlehre der Kharia-Sprache*. Wiesbaden: Harrassowitz.

Praharaj, Gopal C. 1931–1940. *Purnacandra Odia Bhashakosha*. Cuttack: Utkal Sahitya Press.

Shorto, Harry. 2006. *A Mon-Khmer comparative dictionary*. Canberra: Pacific Linguistics.

Turner, Ralph L. 1966. *A comparative dictionary of Indo-Aryan languages*. Oxford University Press.

Starostin, George. 2006. *Dravidian Etymological database*. Available online at: <http://starling.rinet.ru> [Accessed 07.11.2018].

Starostin, George (ed.). 2011–2019. *Global Lexicostatistical database*. Available online at: <http://starling.rinet.ru/new100> [Accessed 07.11.2018].

Vitebsky, Piers. 2017. *Living Without the Dead: Loss and Redemption in a Jungle Cosmos*. Chicago/London: The University of Chicago Press.

Sidwell, Paul (ed.). 2009–2011. *Munda Languages Database*. Available online at: <http://sealang.net> [Accessed 07.11.2018].

A. С. Крылова. Ареальные особенности 110-словного списка по индоарийским языкам Южной Ориссы

Основной темой статьи является проблема замены исконных слов, относящихся к базисной лексике, под влиянием языковых контактов между индоарийскими, дравидийскими и мунда языками в Южной Ориссе и Восточной Индии в целом. Помимо заимствований, особое внимание уделяется случаям неочевидного контактного влияния субстрата в индоарийских языках, когда исконное слово заменяется на другое исконное в результате семантического сдвига, характерного для другой языковой семьи. Рассмотрение таких изменений критически важно для лексикостатистических подсчётов.

Ключевые слова: индоарийские языки; дравидийские языки; мунда языки; базисная лексика; лексикостатистика; ареальная лингвистика.