A Brief Introduction to the Turi Language of Eastern India

John Peterson
Christian-Albrechts-Universität zu Kiel, Deutschland

Abhay Sagar Minz
Dr. Shyama Prasad Mukherjee University, Ranchi, India

Prabhat Linda
Dr. Shyama Prasad Mukherjee University, Ranchi, India

Ariba Khan
Christian-Albrechts-Universität zu Kiel, Deutschland

Francis Xavier Kachhap
Dr. Shyama Prasad Mukherjee University, Ranchi, India

Manish Gari
Dr. Shyama Prasad Mukherjee University, Ranchi, India

Abstract This article presents a brief introduction to the North Munda (Austro-Asiatic) language Turi, spoken by some 1,500 speakers throughout the Indian states of Chhattisgarh, Jharkhand, Odisha, Bihar, West Bengal and Assam. After a brief introduction to the ethnic Turi group, we present a skeleton grammar of the Turi language as spoken in northwestern Odisha state, where it is still being learned by children as their home language. We then discuss the position of Turi within the Kherwarian (North Munda) group by comparing our lexical data for Turi with that for twelve other Kherwarian varieties as given in Kobayashi et al. (2003), using the software COG from the Summer Institute of Linguistics. Our results suggest that Turi is a sister language to all of the dialects of Santali and that it together with these forms the Santali-Turi branch of Kherwarian. We end with a discussion of the possible consequences of these results for the linguistic and ethnic prehistory of Eastern Central India.

Keywords Turi. Kherwarian. North Munda. Austro-Asiatic. Historical linguistics.

Summary 1 Introduction – the Turi and their Language. – 2 A Brief Overview of Turi Grammar. – 3 Turi and Its Relation to Other Kherwarian (North Munda) Languages. – 4 Discussion of the Results of the Comparison. – 5 Summary.

1   Introduction – the Turi and their Language

Turi (Glottocode: 1246; ISO code: ISO 639-3:trd), spoken in six states in eastern central and northeastern India, is a member of the North Munda branch of Austro-Asiatic. Ethnic Turi are found in larger numbers in the eastern central Indian states of Jharkhand, West Bengal, Odisha, Bihar and Chhattisgarh, as well as the northeastern state of Assam. These states and the surrounding regions are shown in Map 1. At present we are not aware of any ethnic Turi groups in Bangladesh **[map 1]**.

Map 1   The states in which larger numbers of Turi live and the larger South Asian region[1]

[1]   Many thanks to Simon Argus for producing this map for us.

Unlike most other ethnic groups which traditionally speak a Munda language, the Turi are officially a scheduled caste, not a scheduled tribe. This is likely due to the fact that the Turi are artisans who traditionally weave various products out of bamboo, such as baskets, winnowing fans, fishing equipment, umbrellas, and other items. As the caste system is closely tied to traditional occupations, and as the vast majority of Turi no longer speak the Turi language, the one obvious trait that most Turi share is their occupation.[2] Despite this official designation as a scheduled caste and not a scheduled tribe, the Turi consider themselves – and are considered by their neighbours – to be a tribal group. They have also been greatly influenced by neighbouring ethnic groups, both Hindus and practitioners of Sarna, the collective term for the religions of all tribal groups in eastern central India, and celebrate the same annual festivals as neighbouring tribal groups, such as Nava Khani, Sarhul, Karm and Sohrai.

Turi is a severely endangered language, with less than 1% of all ethnic Turis still able to speak it, and is classified as "moribund" by the *Ethnologue* (Eberhard et al. 2024), meaning that "[t]he only remaining active users of the language are members of the grandparent generation and older".[3] Fortunately, we can state that there are still young children who speak Turi on a daily basis, albeit in very few and mostly remote small villages, and most Turi nowadays have no active or passive knowledge of their traditional language.

Turi has only been rudimentarily documented until now, both with respect to its lexicon and grammar. In this study we therefore present a basic skeleton grammar of the Odisha dialect of this language as well as a list of over 200 basic vocabulary items and two short texts. Furthermore, as the position of Turi within the North Munda group of languages is still unclear, we also present the results of a comparison of native Turi lexemes with cognate forms in 12 other North Munda varieties as these are documented in Kobayashi et al. (2003), analysing these with the program COG from the Summer School of Linguistics.[4]

Our team conducted fieldwork during several trips to Turi villages in the blocks of Raidih, Bishunpar, Lohardaga and Chainpur in Southern Jharkhand, and Bundu Block in Ranchi District, as well as a few Turi families in the city of Ranchi itself. Shorter visits were also

---

2   The status of groups as "Scheduled Castes" or "Scheduled Tribes" is regulated by the Constitution of India in Articles 366 (24)-(25), and Articles 341 and 342. These are legal terms and do not result automatically from a group being socially recognized as a tribal group: Both terms can be applied to tribal groups, so that a tribal group may not necessarily be classified as a "Scheduled Tribe" but can also be a "Scheduled Caste", as is the case with the Turi.

3   https://www.ethnologue.com/methodology/#Status2.

4   https://software.sil.org/cog/.

made to Jashpur and Surguja Districts in Chhattisgarh. During this work, we also came to know of Turi-speaking groups in northwestern West Bengal and the tea garden estates of Assam, with whom we have conducted telephone interviews.

Turi from all age groups were included in our fieldwork, as were both men and women, and an effort was also made to ascertain through informal conversations the attitude of the Turi community itself towards its traditional language. Finally, a group of five adult Turi speakers from Odisha – three men and two women – were invited to a five-day workshop at the International Documentation Centre for Endangered Indigenous Languages and Cultures at the Dr. Shyama Prasad Mukherjee University, Ranchi, to document the Turi lexicon as well as possible and to record and analyse Turi songs and short stories.[5] As this work has now come to an end, at least for the foreseeable future, we have decided to publish what we have learned through our work so far in the present form, so that the data are available to scholars of Munda and all those interested in the Turi and more generally in the cultural and linguistic landscape of eastern central India.

According to the *Ethnologue* (Eberhard et al. 2024), Turi was spoken in 2007 by a total of ca. 2,000 people in all of India, out of a total ethnic population of 354,000.[6] Thus, according to these figures, Turi was spoken by about 0.56% of the total ethnic Turi population in 2007. Due to the vast area in which the Turi now live, spread throughout at least six states in northeastern and eastern central India, it is exceedingly difficult to determine the number of Turi who still have an active or passive command of their traditional language. Our own estimate is that the language is only spoken by a maximum of 1,000-1,500 people in all of India, although even this figure may be somewhat inflated, based on our own experience and on second-hand accounts, and we have direct knowledge of no more than ca. 120 speakers in Jharkhand, Chhattisgarh and Odisha. But whatever the exact number of speakers found throughout this immense area may be, Turi is clearly a severely endangered language.

A large number of languages from the Indo-Aryan (Indo-European), Munda (Austro-Asiatic) and Dravidian families are spoken in

---

[5] This workshop also resulted in the publication of Peterson, Minz 2021, a primer for Turi-speaking children to learn to write their native language.

 Due to the pandemic situation, which was still acute at that time, it was deemed better to invite a small number of speakers to Ranchi to work with our team than to travel to the communities themselves and live and work with the speakers there. Although this did impede our progress considerably, it also produced an atmosphere which was conducive to this work, as all Turi speakers and all of the authors of this study were vaccinated, which was not possible in any of the Turi-speaking villages.

[6] https://www.ethnologue.com/language/trd/.

these six states, so that the Turi have traditionally had to use a lingua franca in their daily lives, one which varies from one region to another. Thus the Turi are generally multilingual: Even those who do not speak their traditional language but a regional language such as Sadri/Nagpuri in Jharkhand and Chhattisgarh will still have some degree of fluency in Hindi, which is the official language of Bihar, Chhattisgarh and Jharkhand, or Odiya in Odisha, Bengali in West Bengal and Assamese in Assam. Those who still speak Turi in their daily lives are therefore generally at least trilingual. Finally, many Turi are also fluent in one or more neighbouring tribal languages, e.g., Kurux (Dravidian) or Mundari (Munda), depending on the region they live in and whether other ethnic groups also live in their village.

This high degree of multilingualism is certainly one of the main causes for abandoning the traditional Turi language in favour of a local regional or official language. Another possible cause is mentioned by Turi elders, who compare the state of their own traditional language with that of the Asur. The Asur, who speak a North Munda language of the same name, are traditionally iron smelters who live more isolated than the Turi in hill areas, closer to the sources of iron ore. Turi elders claim that the Turi also once lived higher up in the hills but chose to move to lowland forested areas, still near these mountains and hills, where they had better access to bamboo for their trade, but also better access to markets to sell their wares. These areas were also more conducive to agriculture, allowing more Turi to practice that as well.

According to these Turi elders, it is the continued relative isolation of the Asur that is responsible for the better rate of retention of the Asur language by the ethnic Asur than with the Turi language among the Turi: Living in the lowlands, the Turi are in constant contact with other ethnic groups, which is increasingly resulting in intermarriage between the Turi and these groups, as a result of which Turi is usually not passed on to the next generation, which then speaks a regional language at home. In contrast, the more isolated Asur have been better able to maintain their language due to less contact with other groups.[7]

The remainder of this article is structured as follows. Section 2 presents a skeleton grammar of Turi while Section 3 discusses the position of Turi within North Munda, based on our comparative study. The results of this study and their possible significance are then discussed in Section 4. Section 5 provides a summary of this study,

---

[7] The *Ethnologue* lists the status of Asur (referred to there as "Asuri") as "vigorous", meaning that it is used by all generations and that the situation is sustainable (Eberhard et al. 2024), compared with Turi's status there as moribund, mentioned in the preceding pages.

which concludes with two appendices: Appendix 1 presents a list of basic vocabulary items in Odisha Turi, ordered according to the list given in Kobayashi et al. (2003), while Appendix 2 presents two short Turi texts.

2   A Brief Overview of Turi Grammar

This section presents a skeleton grammar of Turi. Only the most basic categories can be touched upon here, such as basic forms of pronouns, number and case, the most common verb categories, lower numerals, etc. Turi, like other Munda languages, is a predominantly verb-final language. It generally differs little in its basic grammatical structures from what we find in other North Munda languages, such as Santali, Mundari or Ho, although it does show some striking differences as well. In the following, all forms are from the Odisha dialect of Turi, unless otherwise explicitly noted.

2.1   Phonology

Table 1 presents the consonant phonemes in Turi which we have been able to identify. All four affricates, all consonants with breathy voice, /ŋ/ and /ʋ/ are restricted in our data to loan words from Indo-Aryan. For ease of presentation, in all Turi example sentences aspiration and breathy voice are represented by <h>. In Table 1 and in the word list in Appendix 1, however, these are given in standard IPA, with superscripted aspiration **[tab. 1]**.

Table 1   The consonant phoneme inventory of Odiya Turi

	Bilabial	Dental	Alveolar	Retroflex	Palatal	Velar	Glottal
Stops	p pʰ b bʱ	t tʰ d dʱ		ʈ ɖtʰ	c ɟ	k kʰ g gʱ	ʔ
Affricates			ʧ ʤʧʰ ʤʱ				
Nasals	m	n		ɳ	ɲ	ŋ	
Laterals			l				
Flaps			ɾ	ɽ			
Fricatives			s				h
Approximants	ʋ				j		

Munda languages are known for their pre-glottalized, voiced stops, often with a nasal release, i.e. [ʔb] or [ʔbᵐ]. This class is also found in syllable-final position in Turi, although there is considerable variation which requires further study. Cf. e.g. from the word-list in Kobayashi et al. (2003, 353, 359) words for 'hair' (#2) from Mundari such

as [uːbˈ]/[uːb] or Santali [uːbˈ]/[uːpˈ], with Turi [uʔm], or 'eye' (#5): Ho [medˈ], Mundari [medˈ]/[med]/[meːʤ] and Santali [meːtˈ]/[metˈ]/[meˑtˈ] (Kobayashi et al. 2003, 347, 353, 359) with Turi [mɛʔn]. As can be seen, Turi tends to lose the stop entirely and retain the nasal – and often also glottalization – although alternative forms are also found in the texts, e.g., [dub], [dubʔm] and [duʔm] 'sit down'.

Table 2 presents the monophthong phoneme inventory of Turi as found in our database. The status of /ə/ as a phoneme is unclear. At present our data suggest that it is best viewed as an allophone of /ɑ/. All vowels can be nasalized; it is not yet clear if nasalization can be phonemic. Vowel length, however, is not phonemic, and all vowels other than /ə/ can be realized as long in certain environments, such as word-final positions [tab. 2].

Table 2   Monophthong vowels in Turi

	Front	**Central**	**Back**
Close	i		u
Mid		(ə)	ɔ
Mid-low Low	ɛ		ɑ

The following diphthongs are also found in our corpus in native words: ə̯ɛ, i̯a.

2.2   Nouns

There is no grammatical gender in Turi. However, object indexing on the verb is sensitive to the animate/inanimate opposition as nouns with animate reference trigger verbal agreement in the 3rd persons, while nouns with inanimate reference do not.[8] Nouns may appear in the singular, which is morphologically unmarked, in the dual, marked by =kin, or in the plural, marked by =kun.

(1)   a.   *bandra*   b.   *bandra=kin*   c.   *bandra=kun*
       'monkey'   'two monkeys'   '(three or more) monkeys'

The morphologically unmarked noun – especially if it has inanimate reference – can also have plural reference if this is clear from context or is indicated elsewhere in the text, e.g. *sɔbu ʈɔpi* 'all the hats' in (22) or *ʈɔpi* 'the hats' in (23).

---

8   One exception to this rule is the locative copula, discussed further below, which exceptionally marks subjects – including inanimate subjects – at the object position. Cf. examples (26)-(27). Otherwise, non-animate subjects are not indexed on the verb.

2.3   Pronouns

Independent pronouns are found in the singular, dual and plural. Dual forms can be reinforced by adding the form *baran=kin* 'both=DU', e.g., *aliŋ baraŋ=kin* 'we both (EXCL)', and plural pronominal forms of the 1st and 2nd persons can be pleonastically marked for *=ku* 'PL', e.g. *alɛ=ku* 'we (PL.EXCL)'. Table 3 presents the independent pronouns **[tab 3]**.

Table 3   Independent pronouns in Turi

	Singular	**Dual**	**Plural**
1ˢᵗ (EXCL)	*iŋ*	*aliŋ*	*alɛ*
1ˢᵗ (INCL)		*alaŋ*	*abu*
2ⁿᵈ person	*am*	*abin*	*apɛ*
3ʳᵈ person	*uni*	*unkin*	*unku*

2.4   Numerals

In Odisha Turi, only the numerals one-three are of Munda origin; beginning with 'four', Indo-Aryan loan words are used. Grierson (1906, 128) also cites the form *pūniā* 'four', which is not in use in Odisha Turi. These are given in Table 4 **[tab. 4]**.

Table 4   Lower numerals in Turi

miaʔn (Jharkhandi Turi: *meja*; Grierson. *mit'*, *miat'*)	'one'
barɛa (Jharkhandi Turi: *baria*)	'two'
pɛa	'three'
ʧar	'four' (IA)
panʧ	'five' (IA)

miaʔn 'one' is also commonly used as an indefinite article; cf. (2).

(2)	*miaʔn*	*duba*	*pɛndarɛ*	*duʔm-ɛn=ə=ɛ.*
	one	tree	under	sit.down-MID.PST=FIN=3SG.ANIM
	'He sat down under a tree.'			

The morph *=gɔɽ*, which is homophonous with a classifier in many neighbouring Indo-Aryan and Munda languages of the region, is also found in our texts, together with the Indo-Aryan numerals *ʧar* 'four' and *panʧ* 'five', but not with the lower numerals of Turi origin. However, it does not appear to be a classifier in Turi, as it is used in counting in our corpus (cf. (3)), but not with a following noun; cf., e.g., *miaʔn duba* 'a tree' in (2) or *miaʔn phɛrivala* 'a hawker' in example (40). We therefore tentatively gloss it as "COUNT", due to its counting function.

(3) *miaʔn barɛa pɛa   ʧar=gɔɽ   panʧ=gɔɽ   lɛkəɛ=a=bu.*
     one   two   three four=COUNT   five=COUNT   count=FIN=1PL.INCL
     'Let's count 1, 2, 3, 4, 5.'

2.5   Case

Odiya Turi has five cases:
- Unmarked – The unmarked (or zero-marked) case marks subjects (e.g., *maʤur* 'the peacock' in (9)) and inanimate objects (e.g., *sɔbu ʈɔpi* 'all the hats' in (22)).
- Objective – This case is used with primary objects with verbs of speech, cf. (4), and animate secondary objects (cf. *uni kɔɽa=kɛ* 'that boy' in example (13)).

(4) *bandra=kun=kɛ   katha-laʔ=ə=ɛ …*
     monkey=PL=OBJ   say-PST.PERF.ACT=FIN=3SG.ANIM
     'He said to the monkeys …'

=*kɛ* can also mark definite inanimate secondary objects (or P, in typological terms), as in (5), although this is not obligatory, as *sɔbu ʈɔpi* 'all the hats' in (22), mentioned above, shows.

(5) *… dan   rɛ   bandra=kun   ʈɔpi=kɛ   haʈ   sɛn=ɛŋ.*
         give.IMP VOC monkey=PL   hat=OBJ   market go=1SG
     '… Give the hats [to me], oh monkeys! I will go to market.'

- Genitive – The genitive is used to incorporate a nominal phrase into a larger nominal phrase. There are two genitive markers: =*aʔ*/=*aʔa*, used with pronouns (6), and =*rɛn*, used elsewhere; cf. (7).

(6) *iŋ=aʔ   ɲumu   prʌbhat=naŋ.*
     1SG=GEN   name   Prabhat=IDENT.COP
     'My name is Prabhat.'

(7) *phɛriʋala=rɛn   durum   bhaŋa-ɛn=a*
     hawker=GEN   sleep(n.) open(itr.)-PST.MID=FIN
     'The hawker woke up (= the hawker's sleep opened).'

With adnominal inalienable possession involving kinship terms, a different construction is used. Here, the noun is followed by the possessive marker *-ta* which is followed by an enclitic form of the independent pronoun, as with the two forms of *ba* 'father' in (8).

(8) a. *ba-t=ɛŋ*   b. *ba-ta=m*
         father-POSS=1SG   father-POSS=2SG
         'my father'   'your father'

- Locative – The locative marks the location of an entity. It has the form =*rɛ* (cf. (9)).

(9) *buru=rɛ   maʤur   susun-tan=ə=ɛ.*
     forest=LOC   peacock   dance-PROG=FIN=3SG.ANIM
     'In the forest the peacock is dancing.'

- Instrumental – The instrumental case marks the instrument with which an action is carried out or the cause of an action. It has the form =*tɛ*, as shown in (10).

(10) *uni   khis=tɛ   ɔɽaʔa   sɛn-ɔʔ-ɛn=ə=ɛ.*
      3SG.ANIM   anger=INST   house   go-MID-PST.MID=FIN=3SG.ANIM
      'He went home out of (= through) anger.'

Case markers follow number markers in Turi; cf. the form *bandra=kun=kɛ* [monkey=PL=OBL] 'to the monkeys' in (4) above or *tʃhava=kun=rɛn* [child=PL=GEN] 'of the children' below in (23).

2.6   General Introduction – Lexical Verbs

The Turi verb system shows many similarities to those of other North Munda languages, although with some differences, as we show in the following. Figure 1 gives a somewhat simplified structural overview of the affirmative finite verb in Turi and Figure 2 that of the negative verb [figs 1-2].

LEXICAL BASE-TAM.ACT/MID=OBJ=FIN=SUBJ

Figure 1 The basic schema of the affirmative finite verb in Turi

NEG=SUBJ LEXICAL BASE-TAM.ACT/MID=OBJ=FIN

Figure 2   The basic schema of the negative finite verb in Turi

The first element of the predicate is the lexical base, which usually consists of a single lexical morpheme. This is followed by portmanteau TAM/basic voice markers (= active or middle), object indexing for objects with animate reference, the finite marker /a/, and subject indexing for subjects with animate reference.

Table 5 gives an overview of the enclitic subject/object markers on the verb. With the exception of the 3rd person singular, these are all highly similar to the full forms of the pronouns given in Table 3. The two forms of the 1st person singular are speaker-specific free variants, while in the 3rd person singular in Odisha Turi =*i* indexes an object and =*ε* indexes a subject at the respective positions shown in Figures 1 and 2. In contrast, in Jharkhandi Turi =*i* indexes both subject and object. Otherwise, all indices in Table 5 can index both subjects and objects, at their respective positions within the predicate **[tab. 5]**.

Table 5 Turi enclitic argument-indices on the verb

	Singular	**Dual**	**Plural**
1st (EXCL)	=*εŋ*/=*iŋ*	=*liŋ*	=*lε*
1st (INCL)		=*laŋ*	=*bu*
2nd person	=*m*	=*bin*	=*pε*
3rd person	=*i* (object)/=*ε* (subject)	=*kin*	=*ku*

The finite marker /α/ is realized as [α] before subject markers beginning with a consonant but is elided before the 1st person singular =*εŋ*/=*iŋ*. Before the 3rd person singular marker =*ε*, the finite marker is raised and realized as /ə/. These two vowels are pronounced together as the diphthong [ə̞ɛ] Cf. the respective forms in (11).

(11) a. *gitiʔ-εn=a=lε*   b. *ɔl-ɛtan=iŋ*   c. *susun-tan=ə=ε*
      sleep-PST.MID=FIN=1PL.EXCL   write-PRES.ACT=1SG   dance-PRES.MID=FIN=3SG.ANIM
      'we slept'   'I am writing'   's/he is dancing'

The P-argument, roughly corresponding to the "direct object" with mono-transitive verbs, is indexed on the predicate before the finite marker when the reference of this argument is animate. This is shown in (12), from Jharkhandi Turi, where =*i* precedes the final finite marker =*a*.

(Jharkhandi Turi)

(12) *hɔn*   *ĩja*   *didi*   *ka=i*   *dal-ɛd=i=a.*
      child_j   REFL   elder.sister_k   NEG=3SG.ANIM.SUB_j   hit-PST.ACT=3SG.ANIM.OBJ_k=FIN
      'The child did not hit his/her own sister.'

When the finite marker is not overtly realized because the subject is the 1st person singular, the object-index directly precedes the subject index, as in (13). Note that these two indices are pronounced as two separate syllables in (13), from Odisha Turi.

(13) iŋ   uni kɔɽa=kɛ   ɲɛl-tad=i=iŋ   (< *ɲɛl-tad=i=a=iŋ)
      1SG that boy=OBJ   see-PST.ACT=3SG.   (< *see-PST.ACT=3SG.
                         OBJ=1SG.SUBJ   OBJ=FIN=1SG.SUBJ)

      'I saw that boy.'

Table 6 presents the TAM/basic voice categories that we have iden-
tified in Odisha Turi. As the Odisha Turi data presently do not con-
tain examples involving the active perfect, the active present-perfect
marker in Table 6 is from Grierson's (1906, 130) data for the Ranchi
dialect of Turi [tab. 6].

Table 6   An overview of the TAM-markers in Turi

	Middle	**Active**
Present	*-tan*	*-ɛtan*
Past	*-ɛn/-kɛn/-ɔn/-ɔʔ-ɛn*	*-ɛkɛn* – Past imperfective
		-tad – Simple past
Future	*-ɔʔ/-ɔ/-ʔ/-*	-
Present perfect	*-akan*	(*-akad*)
Past perfect	*-lɛn*	*-lɛʔ/-laʔ*

The alternate forms of the past perfect marker *-laʔ/-lɛʔ* and the mid-
dle past marker *-kɛn/-ɛn* appear to be speaker-specific (cf. (14)-(15)).

(14) a. alɛ=kun   landa-laʔ=a=lɛ   b. iŋ   ɟɔm-lɛʔ=iŋ
        1PL.EXCL=PL   laugh-PST.PERF.ACT=FIN=1PL.EXCL   1SG   eat-PST.PERF.ACT=1SG
        'we (had) laughed'   'I (had) eaten'

(15) a. gitiʔ-ɛn=a=lɛ   b. sɛn-kɛn=ə=ɛ
        sleep-PST.MID=FIN=1PL.EXCL   see-PST.MID=FIN=3PL
        'we slept'   'he went'

In a few isolated forms (cf. e.g. (16)), the middle marker *-ɔʔ,* or its
shortened form *-ɔ,* appears before the middle-voice past marker *-ɛn*[9]
with no apparent semantic distinction to the use of *-ɛn* alone. After
the short form /ɔ/, the /ɛ/ of *-ɛn* is elided and the form is realized as
[ɔn]. The decline of the use of the middle markers here is perhaps
connected to the fact that these forms are still recognizable as mid-
dle-voice past forms, distinct from the active forms, even without
the marker *-ɔ(ʔ).*

---

9 Also before *-kɛn* in Grierson's (1906, 131) data for the Jashpur dialect; cf. *jō-y-ōk'-
ken=ā* [fruit-y-MID-PST.MID=FIN] 'fruitful-was' (Authors' gloss).

(16) *sarag*   *nilija*   *ɲɛl-ɔʔ-ɛn=ə=ɛ*
      sky   blue   see-PST-MID=FIN=3SG.SUBJ
      'The sky (ANIM) looks (lit. is seen [as]) blue.'

With respect to the active past forms: In our data, *-ɛkɛn* is a general
past imperfective used for prolonged actions such as progressives,
states and iterative actions. In contrast, *-tad* appears to be restrict-
ed to single actions; we analyse it as a pure past tense, not as a per-
fective past tense.

This distinction in the active past seems to hold not only for Odi-
sha Turi but also for the data from Grierson for Ranchi, Jashpur and
Sarangarth Turi, at least tendentially. Compare the use of *-ɛkɛn* in
(17) from Odisha Turi for a prolonged action with that for *-tad* be-
low in (22), also from Odisha Turi, or in (18), from Jashpur Turi, from
Grierson.

(17) *duba*   *tɛŋrɛ*   *bandra=kun*   *ɲɛl-ɛkɛn=a=ku*
      tree   on.top.of   monkey=PL   see=PST.IPFV.ACT=FIN=3PL
      'On top of the tree the monkeys watched/were watching [him do this].'

(Jashpur Turi, Grierson 1906, 131)

(18) *mackam=ke*   *kara*   *kuca-tad=a=e.*
      *machkam.*flower=OBJ   hail   smash-PST.ACT=FIN=3SG.SUBJ
      The *machkam* was smashed by hail (lit.: hail smashed the *machkam* flower).'

Standard negation in Turi is marked by *ka*, which appears directly be-
fore the predicate. The subject index attaches directly to this mark-
er instead of appearing as the final enclitic on the predicate, where
it is found in non-negated clauses (cf. again Figures 1 and 2). When
followed by the 3rd person singular animate marker *=ɛ*, underlying
/kɑ/ is realized as [kə].

(19) a. *sɛn-tan=ɛŋ*   *ka=iŋ*   *sɛn-tan=a*
          go-MID.PRS=1SG   neg=1SG   go-MID.PRS=FIN
          'I am going'   'I am not going'

      b. *sɛn-tan=ə=ɛ*   *kə=ɛ*   *sɛn-tan=a*
          go-MID.PRS=FIN=3SG.ANIM   neg=3SG.ANIM   go-MID.PRS=FIN
          's/he is going'   's/he is going'

Only one object can be marked on the verb in Turi, and only if this
object is considered animate. Object indexing in Odisha Turi follows
a secundative alignment pattern, i.e., a distinction is made between
the indexing of primary objects (= Patients and Goals) and secondary

objects (= Themes), as opposed to indirective alignment, which distinguishes between direct and indirect objects.

A note is necessary here with respect to terminology. With (mono-)transitive verbs, the primary object is P (for 'Patient'), corresponding roughly to the direct object of English. Bitransitive predicates, on the other hand, have G (for the 'Goal' of the action) and T (for 'Theme'), i.e., that entity which 'moves' to the goal with verbs such as *give*. When P and T receive the same marking, e.g., the accusative case in Latin or Sanskrit, they together form the category of direct objects. G is then marked differently from P and T, e.g., the dative in Latin or Sanskrit, and is the indirect object.

In Odisha Turi, P is indexed on the predicate when it references an animate entity, as in (12)-(13) above, repeated here for convenience as (20)-(21).

(Jharkhandi Turi)

(20) hɔn ĩja didi ka=i dal-ɛd=i=a.
child$_j$ REFL elder.sister$_k$ NEG=3SG.ANIM.SUB$_j$ hit-PST.ACT=3SG.ANIM.OBJ$_k$=FIN
'The child did not hit his/her own sister.'

(21) iŋ uni kɔɽa=kɛ ɲɛl-tad=i=iŋ (< *ɲɛl-tad=i=a=iŋ)
1SG that boy=OBJ see-PST.ACT=3SG. (< *see-PST.ACT=3SG.
OBJ=1SG.SUBJ OBJ=FIN=1SG.SUBJ)
'I saw that boy.'

However, an animate G is also indexed on the predicate, as (22) shows, without any applicative to raise it to object status. In (22) the =*i* of the form ɛm-tad=i=a=ku refers to the man to whom the monkeys gave the hats (not explicitly mentioned in (22)). As P and G are similarly indexed at the same slot on the Turi verb but not T, Turi has a primary/secondary object distinction. The secondary object is then T, the NP *sɔbu ʈɔpi* 'all [the] hats', which is not indexed on the verb.

(22) sɔbu katha ajum=kɛtɛ bandra=kun sɔbu ʈɔpi ɛm-tad=i=a=ku
all story hear=CVB monkey=PL all hat give-PST.ACT=3ANIM.OBJ=FIN=3PL
'After hearing his whole story, the monkeys gave him all the hats.'

The gentive-marked possessor of G can also be indexed as the primary object, as with *ʧhava=kun* 'the children' in (23), who are the possessors of *lahiʔɲ=kɛ* 'bellies.'

(23) ʈɔpi   akriŋ=kɛtɛ   ʈhaʋa=kun=rɛn   lahiʔŋ=kɛ   dana   ɛm=ku=ɛŋ.
      hat   sell=CVB   child=PL=GEN   belly=OBJ   food   give=3PL.OBJ=1SG
      'I will sell the hats and give the children food (lit.: Having sold the hats, I will
      give them_i food, the children's_i bellies).'

As examples (24)-(25) from Grierson's data for Ranchi Turi and Sa-
rangarh Turi show, an applicative construction is used in other dia-
lects of Turi when G is indexed on the verb, i.e., G is raised here to
the status of an object. In contrast, in Odisha Turi this is the default
indexing pattern, and no applicative is found.

(Ranchi Turi, Grierson 1906, 130)

(24) oɽo   ac   khurji   haʈiŋ-ad=kin=a=i
      and   anaph   property   distribute-APPL.PST.ACT=3DU.OBJ=FIN=3SG.SUBJ
      'And his father gave them both (= the two sons) their property.'

(Sarangarh Turi, Grierson 1906, 133)

(25) aba   …   hukum   yem-ad=i-y=a=e…*
      father   order   give-APPL.PST=3.OBJ-y-FIN=3SG.SUBJ
      'The father … gave them the order …

*   Grierson (1906, 128-9) notes that in Sarangarh Turi the distinction between
    singular and plural is "often confounded", hence the =i in yem-ad=i-y=a=e refers to
    the sons (plural), despite the apparently singular form.

2.6.1   Copulas

Odiya Turi has two copulas, hɛn-, a locative copula, also used with
temporary states, and the enclitic identity copula =naŋ '=IDENT.COP'.
hɛn- appears to derive from the form tahɛn 'stay, remain', related
forms of which are also found as copulas in other Kherwarian lan-
guages, cf. Mundari tae but also Jharkhandi Turi tai and the Ranchi
dialect of Turi form tahi in Grierson's data, where it also still func-
tions as a full verb with the meaning 'stay, remain'. As a copula it is
only found in Ranchi Turi in the past tenses with the form tahi-, while
hen- is used in the present tense.
    In the Odiya and Jharkhandi Turi data in our corpus, hɛn- marks for
the person/number of the subject at the position which, with transi-
tive verbs, is used to index objects, and is followed by the finite mark-
er =a. We therefore gloss this index as OBJ to call attention to this.

2.6.2 Temporary/Locative Copula

(Jharkhandi Turi)

(26) *ʧɛja*   *lɛkha*   *hɛn=ku=a*   *ʧhuʋa puta?*
what   like   LOC.COP=3PL=FIN   children
'How (lit.: like what) are the children?'

(Jharkhandi Turi)

(27) *ʧɛja*   *lɛkha*   *hɛn=mɛ=a?*
what   like   LOC.COP=2SG.OBJ=FIN
'How are you?'

Exceptionally, with *hɛn-* inanimate subjects are indexed on the verb by *=aʔ*, whereas with other verbs inanimate subjects and objects are not indexed on the verb.

(Odiya Turi)

(28) *bahrirɛ*   *lim*   *duba*   *hɛn=aʔ=a.*
outside   Neem.tree   tree   LOC.COP=3SG.INAN.OBJ=FIN
'There is a Neem tree outside.'

In contrast, in Grierson's data for Turi in the Ranchi area the subject is marked in the usual subject position, verb-finally; cf. example (29).

(Ranchi Turi – Grierson 1906, 130)

(29) *miat'*   *[h]oɽ=ke*[*]   *baria*   *chaua*   *tahi=ken=a=kin.*
one   man=OBJ   two   child   COP=PST.MID=FIN=DU
'A man had two sons (lit.: 'to one man two children were').'
   *   The form is given as <noɽ> in Grierson (1906, 130), although <hoɽ> was surely intended. This may also be an artefact of the reprint which we consulted.

The locative copula in Odisha Turi has the suppletive form *kanɔʔɔ.*

(30) *ɲɛl-laʔ=ə=ɛ*   *ʤɛ*   *miaʔn*   *au*   *ʈopi*   *kanɔʔɔ.*
see-PST.PERF.ACT=FIN=3SG.ANIM   then   one   and   hat   NEG.PRS.COP
'He saw then that there was (= is) not one single hat.'

The temporary state/locative copula marked for an inanimate 3rd person singular subject (= default subject marking) can also be used with a verbal stem and a 'subject' marked by the objective case marker *=kɛ* to denote obligation.

(31) *iŋ=kɛ*   *ɔɽaʔa*   *sɛn*   *hɛn=aʔ=a.*
    1SG=OBJ   house   go   LOC.COP=3SG.INAN.OBJ=FIN
    'I have to go home.'

2.6.2.1   Identity Copula

In Odiya Turi the identity copula is marked by *=naŋ*, the etymology of which is unclear.

(32) *iŋ=aʔ*   *ɲumu*   *prʌbhat=naŋ.*
    1SG=gen   name   Prabhat=IDENT.COP
    'My name is Prabhat.'

(33) *am*   *oka*   *ɖihi=re=naŋ?*
    2SG   which   village=LOC=IDENT.COP
    'What village are you from?' (lit.: 'You are one in which village?')

In contrast, there is no present-tense identity copula in our Jharkhandi Turi data; instead, the two NPs are simply juxtapositioned, as in (34).

(Jharkhandi Turi)

(34) *ʌm=a*   *tʃeja*   *numu?*
    2SG=GEN   what   name
    'What is your name?'

2.6.3   Imperatives and Hortatives

There are a number of imperatives in our corpus, all marked for the second person singular and in the middle voice. The verb is marked by the middle voice marker *-ɔ* and the marker of the second person, singular, *-m*, resulting in *-ɔm* or the slightly irregular *=ɔʔb*. In all forms, the finite marker is lacking. See the examples in (35).

(35) a. *hiɟ-ɔm*   / *hiɟ-ɔʔb*   b. *sɛn-ɔm*   / *sɛn-ɔʔb*
       come-MID.2SG   come-MID.2SG   go-MID.2SG   go-MID.2SG
       'Come!'   'Go!'

A few exceptional forms are found in which neither voice nor person is marked, as these stems do not derive from Turi verbal roots but either from particles which have lexicalized in this function or they have been borrowed from Indo-Aryan, such as those in (36).

(36) *dan!*   *ɛla!*   *ɟu!*
      'give!' (< Indo-Aryan)   'come!'   'go!'

Imperatives are negated by the prohibitive marker *alu*, to which the enclitic subject marker attaches, as shown in (37). Note that the imperative form lacks all TAM marking.

(37) *am*   *alu=m*   *sɛn!*
      2SG   PROH=2SG   go
      'Don't go!'

Hortatives are formed the same way as future-tense verbs, i.e., in the active with zero marking for the future and in the middle voice with the middle-voice marker *-ɔ* attached directly to the verb stem. This is then followed by the finite marker *=a* and inclusive 1st person indexing; cf. example (38), and further forms from the texts in Appendix 2, given in (39).

(38) *ɛla*   *rɛ*   *tʃhaʋa=kun*   *sɛn=a=bu*   *iskul*
      come.IMP   VOC   child=PL   go=FIN=1PL.INCL   school
      'Come along, children! Let's go to school!'

(39) a. *paṛh=a=bu!*   b. *lɛkəɛ=a=bu!*   c. *itu-ɔ=a=bu!*
          learn=FIN=1PL.INCL   count=FIN=1PL.   learn-MID=FIN=1PL.
                                 INCL   INCL
          'Let's learn!'   'Let's count!'   'Let's learn!'

2.6.4   Converbs

There are two common non-finite forms contained in our data. These are:
  • the sequential converb, marked by *=kɛtɛ*, which directly follows the verb stem, as in (40).

(40) *miaʔn*   *pʰɛriʋala*   *ṭɔpi*   *idi=kɛtɛ*   *haṭ*   *sɛn-kɛn=ə=ɛ.*
      one   hawker   hat   take=SEQ   market   go-MID.PST=FIN=3SG.ANIM
      'A hawker took (= having taken) a hat [and] went to market.'

  • the imperfective converb, usually marked by a repetition of the verb stem followed by the marker *-tɛ* .

(41) *pʰɛriʋala*   *landa*   *landa-tɛ*   *haṭ*   *sɛn-lɛn=ə=ɛ.*
      hawker   laugh   laugh-SIM   market   go-MID.PST.PERF=FIN=3SG.ANIM
      'The hawker went to market, laughing all the way.'

Some monosyllabic roots partially reduplicate internally before -tɛ (42), although not all (43).

(42) sɛ-sɛn   sɛ-sɛn-tɛ   thəka-ɛn=ə=ɛ.
    RDP-go   RDP-go-SIM   become.tired-MID.PST=FIN=3SG.ANIM
    'Walking along he became tired.'

(43) dub   dub-tɛ   gitiʔ-ɛn=ə=ɛ.
    sit   sit-SIM   sleep-MID.PST=FIN=3SG.ANIM
   'While sitting there he fell asleep.'

This concludes our grammar sketch. In addition to this sketch, Appendix 1 contains a list of vocabulary items for Odisha Turi according to the list used in Kobayashi et al. (2003). This is followed in Appendix 2 by two short segmented, glossed and translated Turi texts.

3 Turi and Its Relation to Other Kherwarian (North Munda) Languages

In this section we deal with the position of Turi within the Kherwarian branch of North Munda. Due to the very limited data which until now has been available, most of it stemming from the short discussion in Grierson's *Linguistic Survey of India* from 1906, Turi's position within North Munda is still unclear. In the present section we therefore present the results of an automated comparison of Turi with other varieties of the Kherwarian group. To this end, we collected data for Odisha Turi for as many of the 274 lexemes as possible discussed in Kobayashi et al. (2003) for 12 other North Munda varieties and analysed these with the help of the program COG from the Summer Institute of Linguistics in order to determine Turi's position within North Munda.

3.1 The Position of Turi Within the Munda Family – Previous Discussions

The Munda languages, to which Turi belongs, form the western-most branch of the Austro-Asiatic language family, which stretches from Central India in the west to Vietnam in the east. Map 2, from Sidwell (2015, 144), provides an overview of the spread of this family and its main branches [map 2].
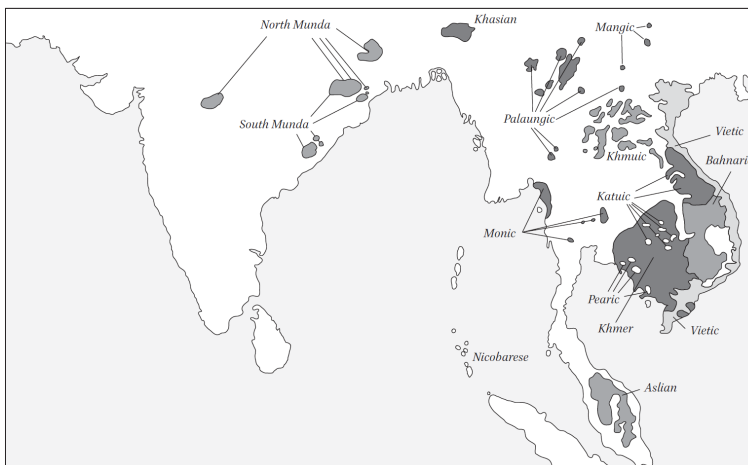
Figure 3 presents the traditional internal classification of Munda, from Zide (1969, 412) **[fig. 3]**. In this classification, the Munda group bifurcates into North Munda and South Munda branches. The northern branch then bifurcates into Korku, spoken in central India, and Kherwarian, spoken in eastern India. Only the Kherwarian branch will be discussed in the following.[10]
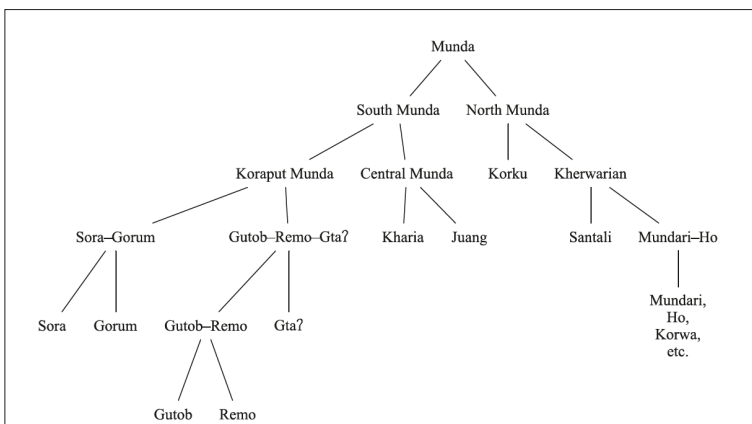

Figure 3   The Munda languages according to Zide (1969, 412)

---

[10]   For an overview of revisions of the classification of Munda languages, see Anderson 2015.

Turi belongs to the Kherwarian group and is largely mutually intelligible with the languages of both the Santali and the Mundari-Ho branches, but what has remained unclear until now is to which of these two branches it belongs, or whether it possibly constitutes a third, independent branch.

Grierson (1906, 128) writes the following of Turi: "The Birhâṛ dialect is closely related to Muṇḍārī, and the speech of the Tūrīs also agrees with that language in most essential points. In a few characteristics, however, it follows Santālī, as against Muṇḍārī". Further on the same page, he writes that "[i]n Sambalpur the Tūrī dialect is almost pure Muṇḍārī" but then goes on to note similarities with Santali, most notably with respect to phonology, which is central to the comparative method: "Forms such as *pēā*, three; *pūniā*, four, in Tūrī agree with Santālī, as does the phonology of the dialect in most points".

Munda (1968, 46-7) on the other hand notes a number of similarities between Turi and both Mundari and Santali.[11] This indeterminate status of Turi within Kherwarian persists until today, e.g. with the *Glottolog* (Hammarström et al. 2022) classifying Turi as a Mundaric language while the *Ethnologue* (Eberhard et al. 2023) groups it together with Santali.

3.2   The Data

In order to determine the position of Turi within the Kherwarian languages, the data for all 274 lexical items contained in Kobayashi et al. (2003, 347-67) for three dialects of Ho (Chaibasa, Goilkera and Ghatshila), four dialects of Mundari (Darigatu, Bandugara, Salgadih and Kera) as well as five dialects of Santali (Kadma, Heben, Tikahara, Hatsara and Simoldohi) were entered by Luna Hemmerling, a student aide at Kiel University, into a spreadsheet. Our own data on Odisha Turi were then added to this spreadsheet.[12]

---

11   Both Osada (1991, 175) and Anderson et al. (2008, 198) cite the following quote from Munda (1968, i-ii), which is even more explicit with respect to Munda's views on the place of Turi within Kherwarian: "The place of Turi was left undefined in Grierson's LSI but we feel that it – along with Asuri, Birhor and Korwa – is now more like Mundari than Santali. In certain respects (e. g., in sharing the same vowels in a few items and in dropping morpheme final vowels in certain forms), however, they look more like Santali than Mundari but they can be derived for the most part as simply from Pre-Mundari". Unfortunately, the first author of this study's own photocopy of Munda 1968, which does not appear to have been published, does not contain any pages with Roman numeration, so that we cannot confirm this quote, although it is largely in line with the discussion of Turi in that work on pp. 46-7.

12   Despite the wealth of data contained in Osada's (1991) publication of Father Ponette's field notes, these data were not entered into the above-mentioned list, primarily due to uncertainties with respect to the granularity of the transcriptions.

Our original aim was to enter detailed data on several dialects of Turi into this file for comparison with these other Kherwarian varieties. Unfortunately, this turned out to be more difficult than anticipated. First, we have not been able to locate any speakers in the state of Jharkhand under the age of 60, and many of the speakers that we have located are either semi-speakers or have not used the language in many years. The situation with COVID-19 further complicated our work, making it impossible to reach many villages during this time, and a number of these older speakers who we had contacted before the pandemic tragically succumbed to this illness during the second wave in 2021. We have however located a relatively large number of Turi speakers from northwestern Odisha, from communities in which Turi is still being learned by the younger generation. As Covid until very recently continued to make travel to these villages difficult, a group of five Turi adults, three male and two female, were invited to Ranchi to cooperate with the authors of this study during the course of five days to elicit stories and songs and to complete as much of this vocabulary list as possible.

As the members of this group came from different villages, there is unfortunately no one single local variety which we could record but five slightly varying varieties, although all speakers were from a relatively small geographical area and their data do not differ substantially. As a result, however, our data are not as fine-grained as that for the 12 Ho, Mundari and Santali dialects documented in Kobayashi et al. (2003). Our data for Odisha Turi were therefore entered as one form of Turi, although we have included all variant forms from these respondents in the spreadsheet. Altogether, data for 227 of the 274 items in the above-mentioned list were obtained in this manner. Data from the considerably more different Jharkhandi Turi were not included, as we still have no data for many lexical entries.[13]

The data were then cleaned to ensure maximal comparability. First, as it is imperative that only cognate forms are being compared, we removed all loan words from the 13 varieties in the data which we were able to determine, including very old loan words from Indo-Aryan such as *daru* 'tree' (no. 118) but of course also more recent loans, such as various words for 'lake' or 'sea' (nos. 140-1), cf. e.g. *sagar*, *samud* (from Sanskrit), *dɔrɛja* (cf. Persian *dærja*), etc. We also removed compounds from the list as the monomorphemic words for the relevant lexemes were already contained elsewhere in the list.

---

13 Unfortunately, no morphological data such as TAM markers, case markers, different forms of PNG markers, etc., were included in this list. Although this type of data is essential for a true comparison of such closely related language varieties, as two reviewers noted, this is presently not an option as these forms are not noted for the other Kherawarian variants in Kobayashi et al. (2003). Hence, there are no cognate forms in that list to compare with the Turi grammatical forms.

E.g., transparent compounds such as e.g. 'tear' (no. 6), consisting of morphemes for 'eye' and 'water', i.e., 'eye-water', were removed from the data, since these two lexemes are already contained in the list (no. 5 'eye' and no. 143 'water').

Complex forms where one part was cognate with forms in other dialects in the list were handled differently; here we retained the cognate lexeme from the respective complex form and removed the other element, which was either opaque, a grammatical morpheme (e.g. a nominalizer) or a lexical morpheme found elsewhere in the list. To give an example, no. 39 'sweat': In the Santali data we find the forms *ud'gɛr* (Heben) and *ud'gər* (Kadma, Tikahara and Simoldohi), but *ud'gər da·k`* for the Hatsara form. Here, *da·k`* is clearly the morpheme for 'water' and was thus removed, as it is found elsewhere in the list (no. 143), while the first element, *ud'gər*, was retained, as it is cognate with the other four entries.

Our goal was primarily to compare phonological developments in these languages in order to determine their genealogical relationships with one another, not lexical similarity per se. For this reason, we only consider phonetic similarity in the following. We also deleted a number of problematic entries, such as deictic units which consisted primarily of grammatical morphemes where it was not always clear whether the forms were cognate (e.g., nos. 257-63). These will have to be studied in more detail in a future study.

Although perhaps the most characteristic phonological trait of Munda languages, the preglottalization and non-audible release of syllable-final voiced plosives (e.g., [ʔbᵐ]) were removed from the data, as these forms are often realized in Turi and many other Munda languages both as preglottalized and as non-glottalized variants even by the same speaker in natural speech, and we felt that it could skew the data if e.g. in some languages both forms occurred but only one of the two happened to be documented. For this reason, only the non-preglottalized forms were used. Finally, as no studies have yet been carried out on lexical accent in Turi or how this is to be defined, all primary and secondary accents were removed from the data for all varieties before comparison. We then deleted all lexical entries for which less than five forms in total were present from the 13 different linguistic varieties.

Finally, all entries were removed which did not have a corresponding non-Indo-Aryan entry for Turi. Altogether, these measures combined to reduce the number of lexical entries considerably, from the original 274 items – of which we have 227 for Turi – to 95. It is this smaller group that forms the basis for our comparison.

We stress here that this is meant only as a preliminary attempt to determine Turi's position within Kherwarian, based on our current knowledge. As finer-grained data for Turi and other Kherwarian languages are added to this lexical database, our understanding of Turi

and the other languages in this group will continue to improve. But as we now have enough data for a first analysis of Turi, we believe that an introduction to the language such as the present one should at least offer a preliminary discussion of Turi's place in Kherwarian, however tentative that may be.

3.3 Results

The data were then analysed with the software COG by the Summer Institute of Linguistics. COG was chosen for a number of reasons: First, the software is openly available and easy to use, even when the respective researcher does not have extensive computational skills. As such, our results can be reduplicated by other researchers regardless of their computational training. Second, COG allows researchers to view the results in various formats, such as dendograms, (non-rooted) trees, networks, and also provides a similarity matrix for the different varieties. Third, COG allows the researcher to actively participate in the analysis, e.g., by correcting false analyses where the algorithm mistakenly views different forms as cognate which the researcher does not consider cognate, or conversely by marking these forms as cognate when the algorithm has not analysed them as such.[14]

Finally, COG provides analyses with respect to lexical and phonetic similarity in both a UPGMA and a Neighbour-Joining (NJ) analysis.[15] Unfortunately, these are the only two algorithms offered in COG, although in view of the preliminary nature of our investigation, this will be sufficient for an initial perspective. The results of this comparison are presented in Figures 4-5.

---

[14]   Due to the preliminary status of our study, this option was not used in the present analysis as this would involve pair-by-pair viewings of all language varieties with one another, based on detailed knowledge of historical Munda developments, which is beyond the expertise of the members of our research group.

[15]   For reasons of space, we will not deal here further with the different assumptions made in the UPGMA vs. Neighbour-Joining analyses. Further discussion of some aspects of these and similar approaches can be found e.g. in Nichols, Warnow 2008. For our purposes it will be enough to simply consider these to be two competing approaches to analysing the data. For information on how COG determines phonetic similarity between two different languages, see Kondrak 2000.
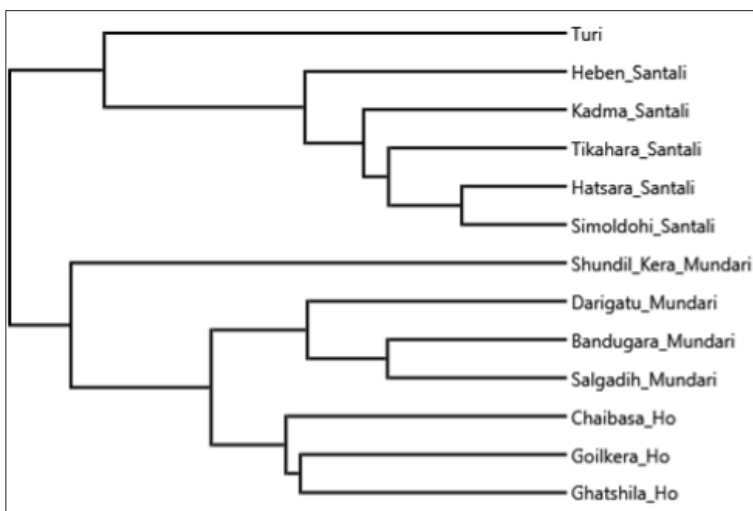
Figure 4   The genetic relationship of various Kherwarian languages and dialects – UPGMA, phonetic similarities
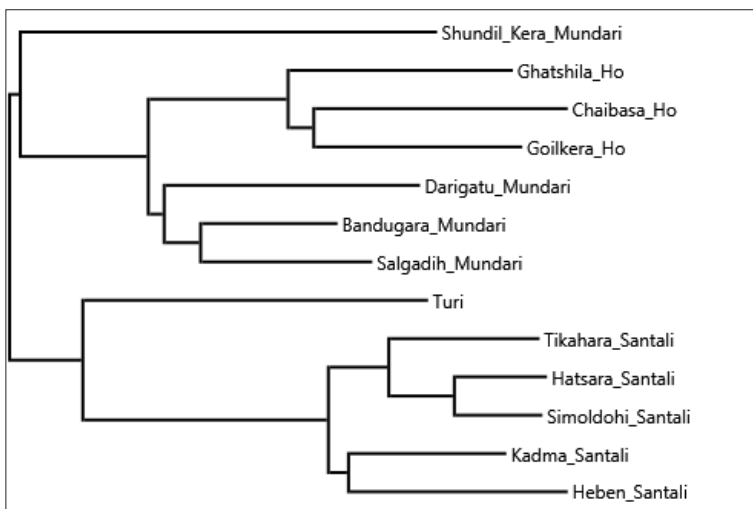
Figure 5   The genetic relationship of various Kherwarian languages and dialects – NJ, phonetic similarities

In both Figures 4 and 5, Santali dialects all group together, all Mundari dialects other than Kera Mundari (to which we will return) group together, and all Ho dialects group together. Furthermore, the Mundari and Ho branches – together with Kera Mundari – form a branch distinct from Santali **[figs 4-5]**. While this classification is not surprising, it does confirm that both analyses come to very similar

conclusions. In fact, they are identical with the exception of Kadma Santali and Heben Santali: in Figure 4, Heben forms the outermost branch of Santali, followed by Kadma and then the remaining three varieties (Tikahara, Hatsara and Simoldohi), while in Figure 5 Kadma and Heben together form a sister branch to the remaining three Santali dialects, which have the same internal relationships. Otherwise the two figures are identical with respect to their internal branching.

With respect to Turi, note that in both Figures 4 and 5 Turi is a sister language to the entire Santali branch, i.e., in neither analysis is Turi more closely related to Mundari-Ho than to Santali. This is more in line with the classification in the *Ethnologue* (Eberhard et al. 2023) than with the *Glottolog* (Hammarström et al. 2022), which classifies Turi as a member of the Mundaric branch.

Figures 4 and 5 also incidentally show that both algorithms classify Kera Mundari as a sister language to the entire Mundari-Ho branch, and not within the Mundari group. While this position of Kera Mundari may at first glance seem somewhat unexpected, it becomes more understandable when the history of this dialect is taken into consideration: Kera Mundari is known to be the result of a language shift by speakers of Kurux (Dravidian) to Mundari, as the result of which it "has unique characteristics and constitutes a distinct regional as well as ethnic dialect" (Kobayashi, Murmu 2008, 165). This shift provides a likely explanation for its relatively distant relation to the other members of this group, although it is considered a dialect of Mundari by the speakers themselves.

4 Discussion of the Results of the Comparison

With respect to the position of Turi within the Kherwarian branch, the results of both algorithms, i.e., UPGMA and NJ, come to the same conclusion, namely that Turi is a sister language to Santali, and that all of the Santali dialects cluster together as a sister branch to Turi, which joins the Santali group at a higher level. Otherwise the genealogical relationships are virtually the same in the two representations.

The question naturally arises why Turi, if it is indeed more closely related to Santali than Mundari-Ho, is spoken so far away from the Santali 'heartland', which today is considerably further to the east in eastern Jharkhand and beyond into West Bengal. We offer here the following tentative suggestion to account for this geographical separation.

Several Kherwarian-speaking ethnic groups such as the Santali, the Mundari and the Ho speak of the migration of their ancestors into their present homelands from the west, and it has recently been suggested that the Santal speakers now residing in eastern Jharkhand (and further still to the north and east) migrated there from western and central Jharkhand, perhaps from the fourteenth century onwards

(Das 2020, 1224-5). This could indicate that the break-up of this earlier, relatively homogeneous dialect group into the present distinct branches only goes back ca. 600 years, perhaps somewhat longer, which fits in well with the high degree of mutual intelligibility among the Kherwarian languages.

In this analysis, Turi and perhaps other smaller Kherwarian languages still spoken in western Jharkhand, eastern Chhattisgarh and northwestern Odisha may be remnants from the time before this eastward migration and before the differentiation of Kherwarian into distinct linguistic sub-groups. Assuming that the results of this admittedly preliminary study stand the test of time, we suggest the following scenario: Before the migration further east into eastern Jharkhand had begun, Kherwarian likely consisted of only very weakly differentiated dialects along a continuum, with at least two still very similar poles, one of which would go on to become the Santali-Turi group, the other the Mundari-Ho group. In this analysis, the ancestors of the present-day Turi would then have remained in western Jharkhand and Chhattisgarh or migrated southward into Odisha, unlike most other Kherwarian groups who migrated eastward.[16]

A word of caution is in order here, however: first, as noted in fn. 20, the data we analysed do not contain any grammatical marking such as TAM or verbal indexing, case or number, as the original list in Kobayashi et al. (2003) does not contain this information. Thus, we did not have such information from the 12 varieties in that list to compare with the Odisha Turi forms. Second, and equally important, is the small size of our database, so that even small changes in the data can lead to a slightly different analysis in either the NJ or the UPGMA analysis, or in both, with respect to the positions of Turi and Kera Mundari. When data are changed in the database for independent reasons, these two languages can either appear in their present positions in Figures 4 and 5, or they can appear together, as a sister branch to all other groups.

For example, when at a relatively late stage in our work a small number of Indo-Aryan loan words were discovered which had not previously been removed from the Santali, Mundari and Ho data from Kobayashi et al. (2003), this changed the position of Kera Mundari and Turi from where they are in Figures 4 and 5 above to clustering together as a separate branch of Kherwarian. Later, however, when a final few Indo-Aryan words were again discovered and removed, the present classification re-emerged.

This indicates to us not that there is a special relationship between Turi and Kera Mundari – there clearly is not. Rather, these

---

16   Presumably only much later did speakers of Turi, similar to other Munda groups, migrate to Assam (see Rau, Sidwell 2019, 36-7).

two languages do not fit well into either the Santali or the Mundari-Ho group. Kera Mundari is somewhat closer to Mundari-Ho than to Santali, and Turi is somewhat closer to Santali than to Mundari-Ho. However, the proximity to either group is apparently not strong enough to rule out another possibility: Whereas Kera Mundari's status as the result of language shift explains its special status quite well, Turi's special status may derive from it having broken off independently from the remaining Kherwarian groups at an earlier date, so that it may represent a third Kherwarian group, instead of descending from an earlier Santali-Turi group. As comparable data on different dialects of Turi and other Kherwarian varieties emerge, this question can hopefully be answered more clearly.

Based on the results of both algorithms, we can however state that Turi is not most closely related to Mundari-Ho. Whether it is a sister to the Santali branch or perhaps an independent branch of Kherwarian awaits further study.

5 Summary

In the present study we give a preliminary introduction to the Turi tribe and their traditional language. The Turi are an officially recognized Scheduled Caste residing in Jharkhand, Chhattisgarh, Odisha, Bihar, West Bengal and Assam. We also provide a skeleton grammar of Odisha Turi, including basic aspects of phonology, the nominal phrase and the verb system. Finally, the two appendices at the end of this study provide a basic vocabulary list for Turi, based on that given in Kobayashi et al. (2003), and two short texts in Turi. We also compare Turi with the 12 other North Munda varieties discussed in Kobayashi et al. (2003) to determine the position of Turi within Kherwarian using the program COG from the Summer Institute of Linguistics. The results suggest that Turi is closer to Santali than to the Mundari-Ho branch.

Our results also indicate that Kera Mundari, despite its name, is best considered not a Mundari dialect but rather a sister language to the Mundari-Ho branch from a phonological perspective, even if it is considered a Mundari dialect from a sociolinguistic perspective. This special status is no doubt due to the fact that the speakers of this language descend from earlier Kurux (Dravidian) speakers who switched to Mundari, leaving their own distinct imprint on the language in the process.

The fact that both algorithms used, i.e., UPGMA and Neighbour-Joining, assign Turi and Kera Mundari to the same respective positions vis-à-vis all other 12 Kherwarian varieties supports the results of our analysis. Nevertheless, further data are required before we can be sure of the internal structure of Kherwarian. We hope that

the present study will provide a foundation on which further studies can build.

Turi is now spoken by only very few people – almost certainly less than 2,000 and perhaps only a few hundred – but its speakers belong to an ethnic population of some 354,000. As our study suggests that Turi is a sister to Santali, which is primarily spoken considerably further to the east, this could mean that Turi emerged as a separate language when one group of speakers of Proto-Santali-Turi remained in eastern Chhattisgarh, western Jharkhand and northwestern Odisha during the general Kherwarian eastward migration. The language of those who remained in the west went on to become modern Turi, while that of the rest of this group, who continued eastwards, went on to become Santali.

There are still many open questions in this proposed development, questions that can only be answered through further fieldwork in the region. The past years have seen a number of important advances with respect to our understanding of the linguistic and ethnic history of this region, but it is clear that there is still much to be done.

While the Chotanagpur Plateau is often considered an accretion zone, following Nichols' (1992; 1997) conceptual categories,[17] it is becoming increasingly clear that Jharkhand is also a 'mini-spread zone'. That is, while the whole of the Chotanagpur Plateau can be considered an accretion zone with respect to the surrounding areas, especially the Gangetic Plain to the north, languages such as Santali, Mundari and Ho and Indo-Aryan languages such as Sadri/Nagpuri, Khortha and Kurmali have also spread throughout this zone at the expense of earlier indigenous languages. Traces of at least one of these earlier languages which are no longer spoken in the region, and which do not appear to belong to any known language family, have also recently been reported.[18] As new data from fieldwork emerge, much of what has long been considered conventional wisdom will likely give rise to new insights – and also to new questions.

---

[17] E.g., Ivani et al. 2021; Peterson 2017; 2022; forthcoming.

[18] Cf. e.g. the data on Kurmali in Paudyal, Peterson 2021, 296.

Abbreviations

1, 2, 3	1st, 2nd and 3rd persons
ACT	active voice
ANAPH	anaphoric pronoun
ANIM	animate
APPL	applicative
COP	copula
COUNT	counting morpheme
CVB	sequential converb
EXCL	exclusive
FIN	finite marker
GEN	genitive
IMP	imperative
IDENT	identity (copula)
INAN	inanimate
INCL	inclusive
IPFV	imperfective
LOC	locative
MID	middle voice
NEG	negative
OBJ	objective case
PROH	prohibitive
PST	past
PL	plural
RDP	reduplication
REFL	reflexive
SG	singular
SIM	simultaneous converb
VOC	vocative
-y-	hiatus-breaking element

Appendix 1: Odisha Turi basic vocabulary

The following list contains all Turi words elicited during our workshop together with speakers of Odisha Turi and follows the order used in Kobayashi et al. (2003). Words marked as '(IA)' for 'Indo-Aryan' were removed before entering the data into COG, as well as all compounds such as *mɛdaʔa* 'tear' (6), which likely consists of the lexemes *mɛʔn* 'eye' (5) and *daʔa* 'water' (143).

The use of '(IA)' does not necessarily imply that the respective Turi lexeme ultimately derives from Indo-Aryan but rather simply that the respective lexeme has the same or a highly similar form in an Indo-Aryan language of the region, from which it was likely borrowed. For example, *ɖihi* 'village' (247) has a very similar form in various Magadhan languages but may ultimately not be of Indo-Aryan origin, or *haphta* 'week' (165), ultimately from Persian but which has entered Turi through neighbouring Indo-Aryan languages. All loanwords which we could identify were removed from the list before comparison – both from Turi as well as from the other North Munda languages. '-' in the following list means that we were not yet able to elicit the respective form corresponding to the morpheme in the list in Kobayashi et al. 2003.

1.	head	*bɔhɔʔɔ*
2.	hair	*uʔm*
3.	forehead	*malaŋ*
4.	eyebrow	*bʰama (IA)*
5.	eye	*mɛʔn*
6.	tear	*mɛdaʔa*
7.	ear	*lutur*
8.	nose	*mũ*
9.	mouth	-
10.	lip	*limtir*
11.	tongue	*alaŋ*
12.	spit	-
13.	tooth	*ɖaʈa (IA)*
14.	chin	*tʰuɽʰi (IA)*
15.	cheek	*dʒɔha*
16.	moustache	*mɛtʃʰa (IA)*
17,	face	*mɛʔn muhaɽ*
18.	neck	*hɔʈɔʔɔ*
19.	throat	*saŋk*
20.	shoulder	*kʰand (IA)*
21.	back	*dɛja*

22.	waist	*majaŋ*
23.	buttock	*tʃutʰal (IA)*
24.	chest	*kɔɽɔm*
25.	breast	-
26.	belly	*lahiʔɲ*
27.	navel	*buʈi*
28.	arm	*tiʔi*
29.	elbow	*kuhuni (IA)*
30.	hand	*tiʔi*
31.	finger	*kaʈuʔ*
32.	nail	*ram*
33.	leg	*dʒaŋga (IA)*
34.	knee	*ʈʰɛuna (IA)*
35.	liver	*kaldʒa (IA)*
36.	heart	*dʒiu (IA)*
37.	guts	*pɔʈa*
38.	skin	*hartaʔa*
39.	sweat	*balbaldaʔa*
40.	filth	*pʰuhuɽi*
41.	pus	*sɔ̃dɔrɔ*
42.	hair	*uʔm*
43.	fat	*tʃarbi (IA)*
44.	blood	*majɔm*
45.	bone	*dʒaŋ (IA)*
46.	flesh	*dʒil*
47.	body	*hɔɽɔ*
48.	disease	*dʒar (IA)*
49.	wound	*gʱaɔ (IA)*
50.	medicine	*ran*
51.	rice	*tʃaʊli (IA)*
52.	powder	*gunɖa (IA)*
52.1	flour	*aʈa (IA)*
53.	salt	*buluŋ*
54.	oil	*sunum*
55.	liquor	*arkʰi (IA)*
56.	tobacco	*tambaku (IA)*
57.	taste	*sibil*
58.	flavour	*sɔ̃ʊ̃*
59.	food	*dʒɔmɛ*
60.	meat	*dʒil*
61.	egg	*bili*
62.	chicken	*sim*
63.	bird	*ɔɽɛ*

64.	wing	*pɑɛŋk (IA)*
65.	feather	*pudga (IA)*
66.	nest	*kʰɔta (IA)*
67.	beak	*tʰɔŋ (IA)*
68.	horn	*siŋ (IA)*
69.	cow	*uriʔ*
70.	knife	*kuntʃi (IA)*
71.	sword	*kʰaɳɖa*
72.	blade	*uhula*
73.	pole	*kuʈa (IA)*
74.	bow	*dʰanu (IA)*
75.	arrow	*tʃɛl (IA)*
76.	lance	-
77.	thread	*sutam (IA)*
78.	needle	*sudʒi (IA)*
79.	clothe	*lidʒaʔa*
80.	paper	*kagadʒ (IA)*
81.	thing	*dʒinis (IA)*
82.	snake	*biŋ*
83.	worm	*tidʒu*
84.	fly	*rɔ̃*
85.	mosquito	*sikiɽi*
86.	flea	-
87.	louse	*siku*
88.	ant	*muʔn*
89.	fish	*haku*
90.	shellfish	-
91.	animal	*dʒatu (IA)*
92.	hunting	*sikar (IA)*
93.	net	*dʒal (IA)*
94.	dog	*sɛta*
95.	rope	*dãɔ̃ra*
96.	string	*baɛr*
97.	sheep	*gaɽa*
98.	horse	*gʰɔɽa (IA)*
99.	pig	*sukiri*
100.	tail	*pɔtʃʰ (IA)*
101.	animal hair	*uʔm*
102.	fur	-
103.	sack	*basta (IA)*
104.	pan	*taʋa (IA)*
105.	kettle	*gaɲ, gã dʒ (IA)*
106.	jar₁	-

107.	jar₂	*gʰamala (IA)*
108.	roof	*tʃat (IA)*
109.	wall	*katʰi*
110.	window	*dʒʰarka (IA)*
111.	door	*siniŋ, siŋduar (duar is IA)*
112.	house	*ɔɽaʔa*
113.	vehicle	*dʒanbahan (IA)*
114.	vessel	*huɖiŋ gaɲɟ*
115.	well	*bauli (IA)*
116.	job	*kami (IA)*
117.	money	*ʈaka (IA), paisa (IA), kɛtʃa*
118.	tree	*duba*
119.	stem	-
120.	branch	*ɖahuɽa*
121.	grass	*gʰãs (IA)*
122.	stalk	*hapa*
123.	root	*dʒɛri (IA)*
124.	leaf	*sɛkam*
125.	flower	*baha*
126.	fruit	*pʰal (IA)*
127.	seed	*bihɔn*
128.	bark	*tʃʰali (IA)*
129.	rice-field	*bahal dɔɛn (dɔɛn IA)*
130.	groove	-
131.	forest	*buru*
132.	road	*hɔra*
133.	hole	*lata (IA)*
134.	bridge	*pulia (IA)*
135.	river	*nai (IA)*
136.	mountain	*buru*
137.	plain₁	*paɽia*
138.	plain₂	*saman (IA)*
139.	pond	*ban (IA)*
140.	lake	*sagar (IA)*
141.	sea	-
142.	island	*ʈapu (IA)*
143.	water	*daʔa*
144.	ice	*barap, barapʰ (IA)*
145.	stone	*diri*
146.	earth	*dʰarti (IA), ɔt*
147.	sand	*dʰuri (IA)*
148.	dust	*lukum dʰuri (dʰuri IA)*
149.	smoke	*sukul*

150.	ash	*tɔrɛʔ*
151.	fire	*sɛŋgɛl*
152.	wind	*dʰuka (IA)*
153.	cloud	*rimil*
154.	fog	*kuhuri (IA)*
155.	rain	*barsa (IA)*
156.	snow	-
157.	sky	*sarag (IA)*
158.	rainbow	*in, indrɔdʰanu (IA)*
159.	sun	*siŋgi*
160.	moon	*tʃaŋɖuʔu (IA)*
161.	shadow	*umbul*
162.	star	*ipil*
163.	day	*siŋgi*
164.	daily	*sɔbu (IA), hilaŋ*
165.	week	*hapʰta (IA)*
166.	month	*mɔhina (IA)*
167.	year	*batʃʰar, baras (both IA)*
168.	morning	*sɛtaʔa*
169.	noon	*tikin*
170.	evening	*ajuʔb*
171.	night	*ninda (IA)*
172.	yesterday	*hɔla*
173.	tomorrow	*gapa*
174.	today	*tisiŋ, tihiŋ*
175.	now	*nahaʔã*
176.	when	*ɔka hilaŋ*
177.	time	*bɛra (IA)*
178.	hour	-
179.	one	*miaʔn*
180.	two	*barɛa*
181.	three	*ρɛa, ρɛnɛŋ*
182.	four	*tʃar (IA)*
183.	five	*pantʃ (IA)*
184.	six	-
185.	seven	-
186.	eight	-
187.	nine	-
188.	ten	-
189.	twenty	-
190.	hundred	-
191.	how much	*cimin*
192.	how many	*cimin*

193.	half	ad^ha (IA)
194.	altogether	sɔbu, səb, dʒɛtɛ (all three IA)
195.	some	midʒaŋ
196.	number	-
197.	age	batʃʰa (IA)
198.	first time	pahila kɛtɛ (pahila IA)
199.	husband	hɛrɛl
200.	wife	laŋgi, lãɟi, lani
201.	marriage	ʋiha (IA)
202.	father	aba (IA)
203.	mother	ajɔ (IA)
204.	grandfather	ɔdʒa (IA)
205.	grandmother	ai (IA)
206.	son	bap (likely IA)
207.	daughter	mai (IA)
208.	child	hɔpɔn
209.	young	dʒaʋan (IA)
210.	grandchild	natija (IA)
211.	elder brother	dada (IA)
212.	elder sister	didi, bai (both IA)
213.	younger brother	bʰai (IA)
214.	younger sister	bahin (IA)
215.	sibling	-
216.	sister	-
217.	family	kuʈum (IA)
218.	friend	gati
219.	quarrel	ʤʰagaɽa (IA)
220.	force	daɽʰi
221.	dumb	kuhula
222.	deaf	bʰaɛ̃ra (IA)
223.	blind	andʰa (IA)
224.	man	hɔɽ
225.	woman	laŋgi, lãɟi, lani
226.	person	hɔɽ
227.	I	iŋ, in
228.	you	am
229.	he	uni
230.	she	uni
231.	we	alɛ
232.	you (pl.)	apɛ
233.	they (m.)	unku
234.	they (f.)	unku
235.	self	ĩja

236.	other	-
237.	who	-
238.	first name	*ɲumu*
239.	family name	-
240.	letter	*tʃiʈtʰi (IA)*
241.	voice	-
242.	sound	-
243.	language	-
244.	mind	*mɔn (IA)*
245.	god	-
246.	festival	-
247.	village	*ɖihi (IA, Magadhan)*
248.	town	-
249.	this	-
250.	it	-
251.	that	*hanɛ, han*
252.	which	*ɔka*
253.	what	*cɛnaʔa, cɛkan, cɛa*
234.	why	-
255.	this	-
256.	how	-
257.	here	-
258.	there	-
259.	that place	-
260.	where	-
261.	this way	-
262.	that way	-
263.	away	-
264.	which way	-
265.	place	*ʈʰə (IA)*
266.	left	-
267.	right	-
268.	front	-
269.	back	*dɛja*
270.	inside	*bʰitri (IA)*
271.	out	*bahrirɛ (IA)*
272.	space	-
273.	up	-
274.	down	-

Appendix 2: Two Short Turi Texts

The following two texts were both composed by Ms. Bishakha Mallik and were translated and analysed during our five-day workshop in Ranchi (see Section 1). We are grateful to her for her permission to publish these here.

Text A: A short Turi school song for young children:

(44) ɛla   rɛ   ʧhaʋa=kun, sɛn=a=bu   iskul.   (repeat once)
     come.IMP   voc   child=PL   go=FIN=1PL.INCL school
     'Come along, children! Let's go to school!'

(45) iskul=rɛ   abu   gɔɳitɔ   paɽh=a=bu.   (repeat once)
     school=LOC   1PL.INCL   math   learn=FIN=IPL.INCL
     'Let's learn math at school!'

(46) alɛ   itu-ɔ=a=bu   lɛkəɛ=a=bu.
     1PL.EXCL   learn-MID=FIN=1PL.INCL   count=FIN=1PL.INCL
     'Let's learn, let's count!'

(47) miaʔn barɛa pɛa   ʧar=gɔɽ   panʧ=gɔɽ   lɛkəɛ=a=bu.
     one   two   three four=COUNT five=COUNT count=FIN=1PL.INCL
     'Let's count 1 – 2 – 3 – 4 – 5!'

Text B: The monkeys' hats - translation from the popular story in Hindi used in schools throughout India.

(48) bandra   ʈɔpi
     monkey   hat
     'The monkeys' hats'

(49) miaʔn phɛriʋala ʈɔpi idi=kɛtɛ haʈ   sɛn-kɛn=ə=ɛ.
     one   hawker   hat   take=CVB market go-PST.MID=FIN=3SG.ANIM
     'A hawker took some hats to market (= having taken hats, went to market).'

(50) *sɛ-sɛn*　　　　*sɛ-sɛn=tɛ*　　　　*thəka-ɛn=ə=ɛ.*
RDP-go　　　　　　RDP-go=SIM　　　　become.tired-PST.MID=FIN=3SG.ANIM
'Walking along he became tired.'

(51) *miaʔn*　*duba*　*pɛndarɛ*　*dubʔm-ɛn=ə=ɛ.*
one　　tree　　under　　　sit.down-PST.MID=FIN=3SG.ANIM
'He sat down under a tree.'

(52) *dub*　　*dub=tɛ*　　　*gitiʔ-ɛn=ə=ɛ.*
sit.down　sit.down=SIM　　sleep-PST.MID=FIN=3SG.ANIM
'While sitting there he fell asleep.'

(53) *duba*　*tɛŋrɛ*　　　*bandra=kun*　*ɲɛl-ɛkɛn=a=ku.*
tree　　on.top.of　　monkey=PL　　see-PST.IPFV.ACT=FIN=3PL
'At the top of the tree the monkeys watched/were watching.'

(54) *phɛrivala*　*gitiʔ-ɛn=ə=ɛ*　*mɛnthã*　*hiʔc=kɛtɛ*　　　*dʒɛtɛ*　*tɔpi=kɛ*
hawker　　　sleep-PST.　　then　　come=CVB　　　all　　hat=OBJ
　　　　　　MID=3SG.ANIM
pindh=kɛtɛ　*dubɑ=rɛ*　　*rakaʔm-ɛn=a=ku.*
put.on=CVB　　tree=LOC　　climb-PST.MID=FIN=3PL
'The hawker fell asleep then, having come, they [= the monkeys] put on all the hats and climbed up (= in) the tree.'

(55) *phɛrivala=rɛn*　　　　　*durum*　　　*bhaŋa-ɛn=a*
hawker=GEN　　　　　　　sleep(n.)　　open(itr.)-PST.MID=FIN
'The hawker woke up (= the hawker's sleep opened).'

(56) *ɲɛl-laʔ=ə=ɛ*　　　　　　　　*dʒɛ*　*miaʔn*　*au*　*tɔpi*　*kanɔʔɔ.*
see-PST.PERF.ACT=FIN=3SG.ANIM　then　one　　and　hat　　NEG.PRS.COP
'He saw then that there was (= is) not one single hat.'

(57) *mɔn*　　*dukh=kɛtɛ*　*duba*　*tɛŋkɛ*　*saŋgil-laʔ=ə=ɛ.*
mind(n.)　be.sad=CVB　tree　　on.top.of　look.upwards-PST.PERF.
　　　　　　　　　　　　　　　　　　ACT=FIN=3SG.ANIM
'Feeling sad he looked up at the tree.'

(58) *ɲɛl-laʔ=ə=ɛ*　　　　　　　　*dʒɛ*　*dʒɛtɛ*　*tɔpi=kɛ*　*bandra=kun*
see-PST.PERF.ACT=FIN=3SG.ANIM　then　all　　hat=OBJ　　monkey=PL
pindh-akan=a=ku.
put.on-PERF.MID=FIN=3PL
'Then he saw that all the monkeys were wearing (= have put on) all the hats.'

(59) *bandra=kun=kɛ* *katha-laʔ=ə=ɛ*   *"dan*   *rɛ*   *bandra=kun*
      monkey=PL=OBJ say-PST.PERF.ACT=FIN=3SG.ANIM give.IMP(IA) VOC monkey=PL
      ʈɔpi=kɛ   *haʈ*   *sɛn=ɛŋ.*
      hat=OBJ   market   go=1SG
      'He said to the monkeys "Give the hats [to me], oh monkeys! I will go to market.'

(60) *ʈɔpi*   *akriŋ=kɛtɛ*   *ʧhaʋa=kun=rɛn*   *lahiʔŋ=kɛ*   *dana*   *ɛm=ku=ɛŋ".*
      hat   sell=CVB   child=PL=GEN   belly=OBJ   food   give=3PL.OBJ=1SG
      'I will sell (= having sold) the hats [and] give the children food".'

(61) *sɔbu*   *katha*   *ajum=kɛtɛ*   *bandra=kun*   *sɔbu*   *ʈɔpi*   *ɛm-tad=i=a=ku*
      all   story   hear=CVB   monkey=PL   all   hat   give-PST.ACT=3SG.
                                                                ANIM.OBJ=FIN=3PL
      'After hearing his whole story, the monkeys gave him all the hats.'

(62) *phɛriʋala*   *landa*   *landa=tɛ*   *haʈ*   *sɛn-lɛn=ə=ɛ.*
      hawker   laugh   laugh=SIM   market   go-PST.PERF.MID=FIN=3SG.ANIM
      'The hawker went off happily (= laughing, went) to market.'

Bibliography

Anderson, D.S. (ed.) (2008). *The Munda Languages*. London; New York: Routledge.
https://doi.org/10.4324/9781315822433

Anderson, G.D.S. (2015). "Overview of the Munda languages". Jenny, M.; Sidwell, P. (eds), *The Handbook of Austroasiatic Languages* [Grammars and Language Sketches of the World's Languages, Mainland and Insular Southeast Asia]. Leiden; Boston: Brill, 364-414.

Anderson, G.D.S.; Osada, T.; Harrison, K.D. (2008). "Ho and the other Kherwarian Languages". Anderson 2008, 195-255.

Das, N.J. (2020). "History of Origin of the Santals of India". *Journal of Xi'an University of Architecture & Technology*, 12(5), 1222-6.
https://xajzkjdx.cn/volume-xii-issue-5-may-2020/

Eberhard, D.M.; Simons, G.F.; Fennig, Ch.D. (eds) (2024). *Ethnologue: Languages of the World. Twenty-Seventh edition*. Dallas (TX): SIL International.
http://www.ethnologue.com.

Grierson, G.A. (1906). *Linguistic Survey of India. Vol. IV. Muṇḍā and Dravidian languages* [Reprint]. Delhi; Varanasi; Patna: Motilal Banarsidass.

Hammarström, H.; Forkel, R.; Haspelmath, M.; Bank, S. (eds) (2022). *Glottolog 4.6.*
https://glottolog.org/

Ivani, J.K.; Paudyal, N.; Peterson, J. (2021). "A House Divided? Evidence for the East-west Indo-aryan Divide and its Significance for the Study of Northern South Asia". Ivani, J.K.; Peterson, J. (eds), Special issue, *Journal of South Asian Languages and Linguistics*, 7(2), 235-74.
https://doi.org/10.1515/jsall-2021-2029

Kobayashi, M.; Murmu, G. (2008). "Keraʔ Mundari". Anderson 2008, 165-94.

Kobayashi, M.; Murmu, G.; Osada, T. (2003). "Report on a Preliminary Survey of the Dialects of Kherwarian Languages". *Journal of Asian and African Studies*, 66, 331-64.

Kondrak, G. (2000). "A New Algorithm for the Alignment of Phonetic Sequences". *Proceedings of the First Meeting of the North American Chapter of the Association for Computational Linguistics* (ANLP-NAACL 2000) (Seattle, April 2000). Stroudsburg (PA): Association for Computational Linguistics, 288-95.
https://aclanthology.org/A00-2038

Munda, R.D. (1968). *Proto-Kherwarian Phonemic System.* Chicago: University of Chicago.

Nichols, J. (1992). *Linguistic Diversity in Space and Time*. Chicago: Chicago University Press.

Nichols, J. (1997). "Modelling Ancient Population Structures and Movement in Linguistics". *Annual Review of Anthropology* 26, 359-84.
https://www.jstor.org/stable/2952527

Nichols, J.; Warnow, T. (2008). "Tutorial on Computational Linguistic Phylogeny". *Language and Linguistics Compass*, 2(5), 760-820.
https://doi.org/10.1111/j.1749-818X.2008.00082.x

Osada, T. (1991). "Father Ponette's field note on Turi". *Journal of Asian and African Studies,* 42, 175-89.

Paudyal, N.P.; Peterson, J. (2021). "How One Language Became Four: The Impact of Different Contact-Scenarios Between 'Sadani' and the Tribal Languages of Jharkhand". Ivani, J.K.; Peterson, J. (eds), Special issue, *Journal of South Asian Languages and Linguistics*, 7(2), 275-306.
https://doi.org/10.1515/jsall-2021-2028

Peterson, J. (2017). "Fitting the Pieces Together. Towards a Linguistic Prehistory of Eastern-Central South Asia (and Beyond)". *Journal of South Asian Languages and Linguistics*, 4(2), 211-57.
https://doi.org/10.1515/jsall-2017-0008

Peterson, J. (2022). "A Sociolinguistic-Typological Approach to the Linguistic Prehistory of South Asia – Two Case Studies". Language Dynamics and Change, 12(2), 224-73.
https://doi.org/10.1163/22105832-bja10018

Peterson, J. (forthcoming). "The Spread of Munda in Prehistoric South Asia – the View from Areal Typology". *Bulletin of Deccan College Postgraduate and Research Institute*, vol. 81.

Peterson, J.; Minz, A.S. (2021). मेरी पहली तुरी चित्रमाला (My First Turi Picture Book). Ranchi: Dr. Shyama Prasad Mukherjee University.

Rau, F.; Sidwel, P. (2019). "The Munda Maritime Hypothesis". *Journal of the Southeast Asian Linguistics Society*, 12(2), 35-57.
http://hdl.handle.net/10524/52454

Sidwell, P. (2015). "Austroasiatic Classification". Jenny, M.; Sidwell, P. (eds), *The Handbook of Austroasiatic Languages* (Grammars and Language Sketches of the World's Languages, Mainland and Insular Southeast Asia). Leiden; Boston: Brill, 144-220.
https://doi.org/10.1163/9789004283572_004

Zide, N.H. (1969). "Munda and Non-Munda Austroasiatic Languages". Sebeok, Th. (ed.), *Current Trends in Linguistics*, vol. 5. The Hague: Mouton, 411-30.