



Santali language

Santali (ᱥᱟᱱᱛᱟᱲᱤ, Pronounced: [santaɾi], ᱥᱟᱱᱛᱟᱲᱤ, ᱥᱟᱱᱛᱟᱲᱤ, ᱥᱟᱱᱛᱟᱲᱤ) is a Kherwarian Munda language spoken natively by the Santal people of South Asia. It is the most widely-spoken language of the Munda subfamily of the Austroasiatic languages, related to Ho and Mundari, spoken mainly in the Indian states of Assam, Bihar, Jharkhand, Mizoram, Odisha, Tripura and West Bengal.^[5] It is a recognised regional language of India as per the Eighth Schedule of the Indian Constitution.^[6] It is spoken by around 7.6 million people in India, Bangladesh, Bhutan and Nepal, making it the third most-spoken Austroasiatic language after Vietnamese and Khmer.^[5]

Santali was a mainly oral language until developments were made by European missionaries to write it in Bengali, Odia and Roman scripts. Eventually, the Ol Chiki script was developed by Raghunath Murmu in 1925. Ol Chiki is alphabetic, sharing none of the syllabic properties of the other Indic scripts, and is now widely used to write Santali in India.

History

According to linguist Paul Sidwell, Munda languages probably arrived on the coast of Odisha from Indochina about 4000–3500 years ago, and spread after the Indo-Aryan migration to Odisha.^[7]

Until the nineteenth century, Santali had no written language and all shared knowledge was transmitted by word of mouth from generation to generation. European interest in the study of the languages of India led to the first efforts at documenting the Santali language. Bengali, Odia and Roman scripts were first used to write Santali before the 1860s by European anthropologists, folklorists and missionaries including A. R. Campbell, Lars Skrefsrud and Paul Bodding. Their efforts resulted in Santali dictionaries, versions of folk tales, and the study of the morphology, syntax and phonetic structure of the language.

Santali

ᱥᱟᱱᱛᱟᱲᱤ

ᱥᱟᱱᱛᱟᱲᱤ

The word *Santali* in Ol Chiki script

Native to	<u>India</u> , <u>Bangladesh</u> , <u>Nepal</u>
Ethnicity	<u>Santal</u>
Native speakers	7.6 million (2011 census ^[1] ^[2])
Language family	<u>Austroasiatic</u> <ul style="list-style-type: none"><u>Munda</u><ul style="list-style-type: none"><u>North Munda</u><ul style="list-style-type: none"><u>Kherwarian</u><ul style="list-style-type: none"><u>Santal</u><ul style="list-style-type: none">Santali
Dialects	<u>Mahali (Mahili)</u> <u>Kamari-Santali</u> <u>Khole</u> <u>Lohari-Santali</u> <u>Manjhi</u> <u>Paharia</u>

The Ol Chiki script was created for Santali by Mayurbhanj poet Raghunath Murmu in 1925 and first publicised in 1939.^[8]

Ol Chiki as a Santali script is widely accepted among Santal communities. Presently in West Bengal, Odisha, and Jharkhand, Ol Chiki is the official script for Santali literature & language.^{[9][10]} However, users from Bangladesh use Bengali script instead.

Santali was honoured in December 2013 when the University Grants Commission of India decided to introduce the language in the National Eligibility Test to allow lecturers to use the language in colleges and universities.^[11]

Geographic distribution

Santali is spoken by over seven million people across India, Bangladesh, Bhutan, and Nepal, with India being its native country and having the largest number of speakers.^[5] According to 2011 census, India has a total of 7,368,192 Santali speakers (including 358,579 Karmali, 26,399 Mahli).^{[13][14]} State wise distribution is Jharkhand (2.75 million), West Bengal (2.43 million), Odisha (0.86 million), Bihar (0.46 million), Assam (0.21 million) and a few thousand in each of Chhattisgarh, and in north-eastern states Tripura, Arunachal Pradesh, Mizoram.^[15]

The highest concentrations of Santali language speakers are in Santhal Pargana division, as well as East Singhbhum and Seraikela Kharsawan districts of Jharkhand, the Jangalmahals region of West Bengal (Jhargram, Bankura and Purulia districts) and Mayurbhanj district of Odisha.

Smaller pockets of Santali language speakers are found in the northern Chota Nagpur plateau (Hazaribagh, Giridih, Ramgarh, Bokaro and Dhanbad districts), Balesore and Kendujhar districts of Odisha, and throughout western and northern West Bengal (Birbhum, Paschim Medinipur, Hooghly, Paschim Bardhaman, Purba Bardhaman, Malda, Dakshin Dinajpur, Uttar Dinajpur, Jalpaiguri and Darjeeling districts), Banka district and Purnia division of Bihar (Araria, Katihar, Purnia and Kishanganj districts), and tea-garden regions of Assam (Kokrajhar, Sonitpur, Chirang and Udalguri districts). Outside India, the language is spoken in pockets of Rangpur and Rajshahi divisions of northern Bangladesh as well as the Morang and Jhapa districts in the Terai of Koshi Province in Nepal.^{[16][17]}

Writing system

- Ol Chiki script (official)^[3]
- Bengali-Assamese script^[4]
- Odia script
- Devanagari
- Roman script

Official status

Official language in

- India
 - Jharkhand (additional)
 - West Bengal (additional)

Language codes

ISO 639-2

sat (https://www.loc.gov/standard/s/iso639-2/php/languagecodes_name.php?code_ID=388)

ISO 639-3

Either:
sat – Santali
mjx – Mahali

Glottolog

sant1410 (<http://glottolog.org/resource/language/id/sant1410>)
Santali
maha1291 (<http://glottolog.org/resource/language/id/maha1291>)
Mahali

Official status

Santali is one of India's 22 scheduled languages.^[6] It is also recognised as the additional official language of the states of Jharkhand and West Bengal.^{[18][19]}

Dialects

Dialects of Santali include Kamari-Santali, Khole, Lohari-Santali, Mahali, Manjhi, Paharia.^{[5][20][21]}

Phonology

Consonants

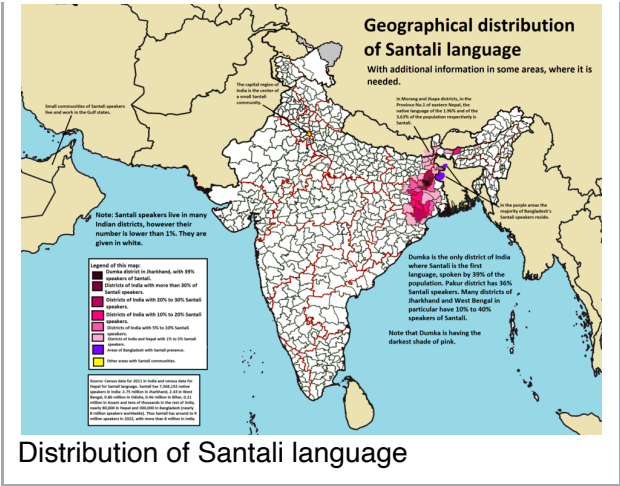
Santali has 21 consonants, not counting the 10 aspirated stops which occur primarily, but not exclusively, in Indo-Aryan loanwords and are given in parentheses in the table below.^[22]

		<u>Bilabial</u>	<u>Alveolar</u>	<u>Retroflex</u>	<u>Palatal</u>	<u>Velar</u>	<u>Glottal</u>
<u>Nasal</u>		<u>m</u>	<u>n</u>	(<u>ɳ</u>)*	<u>ɲ</u>	<u>ŋ</u>	
<u>Stop</u>	<u>voiceless</u>	<u>p</u> (<u>pʰ</u>)	<u>t</u> (<u>tʰ</u>)	<u>ʈ</u> (<u>tʰ</u>)	<u>t͡ʃ</u> (<u>tʃʰ</u>)	<u>k</u> (<u>kʰ</u>)	<u>ʔ</u>
	<u>voiced</u>	<u>b</u> (<u>bʱ</u>)	<u>d</u> (<u>dʱ</u>)	<u>ɖ</u> (<u>dʱ</u>)	<u>d͡ʒ</u> (<u>dʒʱ</u>)	<u>ɡ</u> (<u>ɡʱ</u>)	
<u>Fricative</u>			<u>s</u>				<u>h</u>
<u>Trill/Flap</u>			<u>r</u>	<u>ɽ</u>			
<u>Approximant</u>			<u>l</u>		<u>j</u>	<u>w</u>	

*ɳ only appears as an allophone of /n/ before /d/.

In native words, the opposition between voiceless and voiced stops is neutralised in word-final position. A typical Munda feature is that word-final stops are "checked", i. e. glottalised and unreleased.

Bodding (1929) noted that in the vowel space between an open syllable and a syllable that starts with a vowel, if both vowels are of the same height, approximant /w/ is inserted in between cues of two low vowels, and /y/ for mid-high and high vowels.



A girl speaking Santali.

Vowels

Santali has eight oral and six nasal vowel phonemes. With the exception of /e o/, all oral vowels have a nasalized counterpart.

	Front	Central	Back
High	i ĩ		u ũ
Mid-high	e	ə ə̃	o
Mid-low	ɛ ẽ		ɔ õ
Low		a ã	

There are numerous diphthongs.

Word prominence

Santali prosody exhibits iambic patterns with stress is always released in the second syllable in most disyllabic words, excepting loan words from Hindi, Bihari, Bengali and Assamese.^{[23][24]}

Vowel harmony

Like all Kherwarian languages, vowel harmony in Santali is a morphological triggered process.^[25]

Morphology

Santali, like all Munda languages, is a suffixing agglutinating language.

Nouns

Nouns are inflected for number and case.^[26]

Number

Three numbers are distinguished: singular, dual and plural.^[27]



Santali books in Mayurbhanj Book Fair

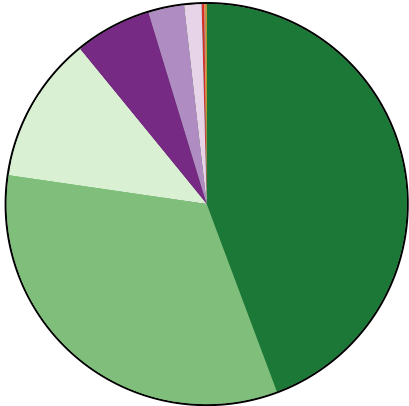
Singular	୧୦୦୩ (seta)	'dog'
Dual	୧୦୦୩-ବୢ୧୧ (seta-ken)	'two dogs'
Plural	୧୦୦୩-ବୢ୩ (seta-ko)	'dogs'

Case

The case suffix follows the number suffix. The following cases are distinguished:^[28]

Case	Marker	Function
<u>Nominative</u>	-Ø	Subject and object
<u>Genitive</u>	-୩୧୧ (animate) -୩୧, -୩୧୧୧୧ (inanimate)	Possessor
<u>Comitative</u>	-୩୩୧୧/ -୩୩୧୧	Goal, place
<u>Instrumental-Locative</u>	-୦୧	Instrument, cause, motion
<u>Sociative</u>	-୧୩୩	Association
<u>Allative</u>	-୧୧୧/-୧୧୧	Direction
<u>Ablative</u>	-ବ୩୩୧୧/-ବ୩୩୧୧	Source, origin
<u>Locative</u>	-୩୧	Spatio-temporal location

Transcript version:



Distribution of Santali language in the states of India^[12]

- Jharkhand (44.4%)
- West Bengal (33%)
- Odisha (11.7%)
- Bihar (6.2%)
- Assam (2.9%)
- Maharashtra (1.4%)
- Chhattisgarh (0.2%)
- Tripura (0.1%)
- Other states (0.10%)

Case	Marker	Function
<u>Nominative</u>	-Ø	Subject and object
<u>Genitive</u>	- <i>rɛn</i> (animate) - <i>ak'</i> , - <i>rɛak'</i> (inanimate)	Possessor
Comitative	- <i>ʈhɛn/ʈhɛc'</i>	Goal, place
Instrumental-Locative	- <i>tɛ</i>	Instrument, cause, motion
<u>Sociative</u>	- <i>são</i>	Association
<u>Allative</u>	- <i>sɛn/-sɛc'</i>	Direction
<u>Ablative</u>	- <i>khɔn/-khɔc'</i>	Source, origin
Locative	- <i>rɛ</i>	Spatio-temporal location

Possession

Santali has possessive suffixes which are only used with kinship terms: 1st person *-n*, 2nd person *-m*, 3rd person *-t*. The suffixes do not distinguish possessor number.^[29]

Pronouns

The personal pronouns in Santali distinguish inclusive and exclusive first person and anaphoric and demonstrative third person.^[30]

Personal pronouns

		Singular	Dual	Plural
1st person	exclusive	ꠘꠑ	ꠘꠤꠞꠑ	ꠘꠞꠤ
	inclusive		ꠘꠞꠘꠤ	ꠘꠞꠘꠤ
2nd person		ꠘꠘ	ꠘꠞꠤꠤ	ꠘꠞꠤꠤ
3rd person	Anaphoric	ꠘꠞ	ꠘꠤꠞꠤ	ꠘꠞꠤ
	Demonstrative	ꠘꠤꠞ	ꠘꠤꠞꠤꠤ	ꠘꠤꠞꠤꠤ

Transcript version:

		Singular	Dual	Plural
1st person	exclusive	iŋ	əliŋ	alɛ
	inclusive		alaŋ	abo
2nd person		am	aben	apɛ
3rd person	Anaphoric	ac'	əkin	ako
	Demonstrative	uni	unkin	onko

The interrogative pronouns have different forms for animate ('who?') and inanimate ('what?'), and referential ('which?') vs. non-referential.^[31]

Interrogative pronouns

	Animate	Inanimate
Referential	ᱫᱟᱜᱟᱴ	ᱫᱷᱟᱥᱱ
Non-referential	ᱫᱷᱟᱴᱥ	ᱫᱷᱟᱴᱥ

Transcript version:

	Animate	Inanimate
Referential	ɔkoɐ	oka
Non-referential	cele	cet'

The indefinite pronouns are:^[32]

Indefinite pronouns

	Animate	Inanimate
'any'	ᱫᱷᱟᱴᱥᱫᱟᱜᱟᱴ	ᱫᱷᱟᱴᱥ
'some'	ᱫᱷᱟᱴᱥ	ᱫᱷᱟᱴᱥᱫᱟᱜᱟᱴ
'another'	ᱫᱷᱟᱴᱥᱫᱟᱜᱟᱴ	ᱫᱷᱟᱴᱥᱫᱟᱜᱟᱴ

Transcript version:

	Animate	Inanimate
'any'	<i>jāheā</i>	<i>jāhā</i>
'some'	<i>adom</i>	<i>adomak</i>
'another'	<i>ɛʃak'ic'</i>	<i>ɛʃak'ak'</i>

The demonstratives distinguish three degrees of deixis (proximate, distal, remote) and simple ('this', 'that', etc.) and particular ('just this', 'just that') forms.^[33]

Demonstratives

		Simple		Particular	
		Animate	Inanimate	Animate	Inanimate
Proximate	Singular	୧ବଢ	୧ଦୱନ୍	୧ଟ	୧ଟଢ.
	Dual	୧ବବଟ	୧ଦୱନ୍ବଟ	୧ଟବଟ	୧ଟଢ.ବଟ
	Plural	୧ଡବଡ/୧ବବବ	୧ଦୱନ୍ବଡ	୧ଢ.ବଡ/୧ବବବ	୧ଟଢ.ବଡ
Distal	Singular	ବ୧ଟ	ଡ୧ନ୍	ଟ୧ଟ	ଟ୧ନ୍.
	Dual	ଡ୧ବଟ	ଡ୧ନ୍ବଟ	ଟ୧ବଟ	ଟ୧ନ୍.ବଟ
	Plural	ଡ୧ବଡ/ବ୧ବବ	ଡ୧ନ୍ବଡ	ଢ.୧ବଡ/ଟ୧ବବ	ଟ୧ନ୍.ବଡ
Remote	Singular	ଓନ୍.୧ଟ	ଓନ୍୧ନ୍		
	Dual	ଓନ୍.୧ବଟ	ଓନ୍୧ନ୍ବଟ		
	Plural	ଓନ୍.୧ବଡ	ଓନ୍୧ନ୍ବଡ		

Transcript version:

		Simple		Particular	
		Animate	Inanimate	Animate	Inanimate
Proximate	Singular	<i>nui</i>	<i>noa</i>	<i>nii</i>	<i>niə</i>
	Dual	<i>nukin</i>	<i>noakin</i>	<i>nikin</i>	<i>niəkin</i>
	Plural	<i>noko/nuku</i>	<i>noako</i>	<i>neko/niku</i>	<i>niəko</i>
Distal	Singular	<i>uni</i>	<i>ona</i>	<i>ini</i>	<i>inə</i>
	Dual	<i>onkin</i>	<i>onakin</i>	<i>inkin</i>	<i>inəkin</i>
	Plural	<i>onko/unku</i>	<i>onako</i>	<i>enko/inku</i>	<i>inəko</i>
Remote	Singular	<i>həni</i>	<i>hana</i>		
	Dual	<i>hənkin</i>	<i>hanakin</i>		
	Plural	<i>hanko</i>	<i>hanako</i>		

Numerals

The basic cardinal numbers (transcribed into Latin script IPA)^[34] are:

1	ᱠᱟᱹ	<i>mit'</i>
2	ᱡᱷᱟᱨᱠᱷᱚ	<i>bar</i>
3	ᱵᱚ	<i>pε</i>
4	ᱵᱤᱨ	<i>pon</i>
5	ᱠᱚᱴᱚᱨ	<i>mɔ̃ɽɛ̃</i>
6	ᱢᱚᱴᱚᱨ	<i>turui</i>
7	ᱵᱤᱨᱵᱚᱨ	<i>ɛyae</i>
8	ᱠᱚᱨᱚᱰ	<i>irəl</i>
9	ᱠᱚᱨᱚ	<i>arɛ</i>
10	ᱪᱚᱰ	<i>gel</i>
20	ᱠᱚᱵᱚᱨ	<i>-isi</i>
100	ᱚᱨᱚᱨ	<i>-sae</i>

The numerals are used with numeral classifiers. Distributive numerals are formed by reduplicating the first consonant and vowel, e.g. *babar* 'two each'.

Numbers basically follow a base-10 pattern. Numbers from 11 to 19 are formed by addition, "gel" ('10') followed by the single-digit number (1 through 9). Multiples of ten are formed by multiplication: the single-digit number (2 through 9) is followed by "gel" ('10'). Some numbers are part of a base-20 number system. 20 can be "bar gel" or "isi".

ᱵᱚ
pe
(3

×

ᱪᱚᱰ
gel
(10)

or
or
or

(ᱠᱟᱹ)
(mit')

ᱠᱚᱵᱚᱨ
isi

ᱪᱚᱰ
gel

((1)

×

20

+

10)

30

Derivation

To derive new nominals, the stems of lexical verbs, adjectives, and other nouns can employ many different methods, including affixation, reduplication, and compounding.

Suffixation: Two nominalizing suffixes *-ic* for animate, and *-ak* for inanimate noun class, are used to form referential nominals.^[35]

Verbs → nouns: *jɔm* ('eat') > *jɔmak* ('food')

adjectives → nouns: *nɔtɛ* ('this side') > *nɔtɛn* ('belonging to this side') > *nɔtɛnak* ('thing of this side') / *nɔtɛnic* ('one of this side')

ponɖ ('white') > *ponɖak* ('white thing') / *ponɖic* ('white one')

suffixes → nouns: *ɔl-tɛ* (write-INS) > *ɔltɛak* ('that with which is written(pen)')

Infixation is the most productive derivation method in Santali. Infixes *-tV-*, *-nV-*, *-mV-*, *-ɭV-*, and *-pV-* are often inserted into nouns, verbs, adjectives to derive new words.^[36]

ɛhɔp ('begin') > *ɛtɔhɔp* ('beginning')

rakap ('rise', 'ascend') > *ranakap* ('development')

Prefixation in North Munda has been reduced to a very few restricted exceptions.^[37]

cɛt ('teach') > *macɛt* ('teacher')

Verbs

Verbs in Santali inflect for tense, aspect and mood, voice and the person and number of the subject and sometimes of the object.^[38]

Santali TAM

There are no specific markers for the Imperative.^[39]

Santali verb paradigm	Transitive	Intransitive
Future/Present	-Ø	-ok
Present Progressive	-et (-kan)	-ok-kan
Simple Past	-ket	-en
Pluperfect	-let	-len
Perfect	-akat	-akan
Past perfect	-akat-tahẽkan	-akan-tahẽkan
Past progressive	-et-tahẽkan	-ok-kan-tahẽkan
Optative	-ke	-k-ok
Irrealis	-le	-len
Conditional	-khan	

Subject markers

		singular	dual	plural
1st person	exclusive	-ŋ(iŋ)	-liŋ	-le
	inclusive		-laŋ	-bon
2nd person		-m	-ben	-pe
3rd person		-e	-kin	-ko

Object markers

Transitive verbs with pronominal objects take infixed object markers.

		singular	dual	plural
1st person	exclusive	-iŋ-	-liŋ-	-le-
	inclusive		-laŋ-	-bon-
2nd person		-me-	-ben-	-pe-
3rd person		-e-	-kin-	-ko-

In applicative constructions, inanimate objects are marked with a pronominal suffix, a checked *-k*.

Indirect object indexing

Transitive verbs may form agreements with non-arguments/outside/indirect objects. To denote inalienable possession of the concerned indirect object, prefix *-t-* is attached to the applicative forms of the pronouns; otherwise it is marked in the noun phrase and functions as an attribute.

ako-ge=ko

3PL-EMPH=3PL.SUBJ

idi-ke-t-ko-tako-a

take-AOR-TR-3PL.OBJ-3PL.POSS-FIN

'They took theirs away themselves.'

Limitations of Santali indexation

In Santali as well as Kherwarian languages, the pronominal subject markers are highly mobile clitics that tend to attach to the word preceding the verb stem. If the subjects and objects are not considered topical to the discourse and dropped, then the subject marker may appear at the end of the verbal complex. Subbarao & Everaert (2021) and Koshy (2021) categorized Kherwarian subject markers as phrasal clitics, as their functions encompass the entire verbal construction rather than affecting single individual verb.

hɔpɔn-iŋ

son-1SG

dɔ

TOP

khube-ge-y=e

very.much-FOC-y=3SG.SUBJ

kəmi-ke-t-tae-tiŋ-a

work-AOR-TR-3SG.POSS-1SG.POSS-FIN/IND

'This son of mine worked very much and well for him.'

Indexing arguments in Santali is essentially intertwined with the distinction of animacy of arguments. Distinction between animate/inanimate is not marked on nouns at all, but is conveyed through morphosyntax, such as in genitive and locative cases and verbal agreement. That is, if an argument of the verb does not belong to the animate noun class, the verb will not index that argument. Inanimate entities such as flower, tree,

rice, book, food,... and objects that cannot move by themselves like vehicles (eg. motorbike, Tesla, airplane) are never indexed by the verb. The following examples illustrate the distinction between inanimate and animate versions of the same noun in two different contexts.

<i>Kanehito</i>	<i>Yamada</i>	<i>Frieren=e</i>	<i>ɔl-ed-a</i>
Kanehito	Yamada	Frieren=3SG.SUBJ	write-PROG.A-FIN
'Kanehito Yamada is writing <u>Frieren</u> (manga).'			
<i>Frieren</i>	<i>Aura</i>	<i>a-jɔ-aka-d-e-a=e</i>	
Frieren	Aura	CAUS-eat-PRE-TR-3SG.OBJ-FIN=3SG.SUBJ	
'Frieren has given Aura to eat (it).'			

However, there are some notable exceptions of inanimate objects that are significant ('sun', 'moon', 'star') or culturally important ('doll') are considered animate in Santali. Nouns like 'Government' is also considered a single body of animate entities and is marked with third person singular. Even mushroom, thorn being pricked, puff-ball, earwax are perceived as animate and are indexed by pronominal markers as such, showing the unpredictability of the Santali animacy-based indexation system.^[40]

Imperative

As described by Ghosh (2008), there are no specific markers for the imperative series. However, in the affirmative imperative, the indicative/finite marker *-a* is replaced by second person markers. In the negative imperative, verb (TAM/person-syntagma) takes *-a* while the imperative subject marker moves to the enclitic position behind the negative particle, right before the verb (See ##Negation).

<i>daya-kate</i>	<i>ma-ge</i>	<i>oko-baŋcao-ka-ŋ-tabon-pe</i>
show.mercy-CONV	MOD-FOG	hide-save-ACT.RES-1SG.OBJ-1PL.INCL.POSS-2PL.SUBJ.IMP
'Please show kindness and hide and save me (for the sake of us)'		

Finiteness

Any finite verbs will attach *-a*, except the imperative and in the subordinate clause.^[41]

<i>noa-rɛak</i>	<i>mit</i>	<i>ŋaŋ</i>	<i>kəhəni</i>	<i>lɔi-ad-iŋ-a=e</i>
this-GEN	one	CLF	story	tell-APPL.PST.A-1SG.OBJ-FIN=3SG.SUBJ
'He told me a story about this.'				

Causative

There are two causative markers: *a-* and *-oco*. *-oco* is attached on every type of verb stems, and *a-* is restricted to two transitive verbs *jɔm* ('eat') and *ɲu* ('drink').^[42]

sɛn-otɕo-dare-a-e-a=ɲ

GO-CAUS-ABIL-BEN-3SG.OBJ-FIN=1SG.SUBJ

'I let/made him come.'

Reciprocal

Infix *-pV-* turns transitive and ditransitive verb roots into reciprocal meaning, but in many verbs it also conveys that the action is done together by two participants.^[43]

dal ('beat') > *dapal* ('beat each other')

landa ('laugh') > *lapanda* ('laugh together')

Benefactive

The benefactive for transitive and ditransitive stems is *-ka* in Northern Santali dialect and *-ka-k* in Southern Santali. In Southern Santali, if the object is animate, the last *-k* will be replaced by pronominal clitics.

tɔl ('bind') > *tɔlka* ('to bind for somebody')

Medio-passive

Transitive verbs and a limited number of intransitive and intransitive-transitive verb roots will take *-jɔn* to form the Medio-passive voice.^[44]

Passive and Reflexive

Transitive roots, transitive-intransitive roots, and causative stems will take *-ok* to derive passive stems. In the transitive-intransitive roots, it denotes the prominence of transitivity. Attaching it to transitive verbs will create reflexivity.^[44]

ɲɛl ('see') > *ɲɛlok* ('be seen') (passive)

ranoco ('cause to medicate') > *ranocok* ('be caused to medicate') (causative > passive)

mak ('cut') > *makok* ('cut oneself') (reflexive)

Nominal verbalisation

In daily speeches, nominal roots can be found functioning as verbs with appropriate inflection. The verbalization of nominals extends to interrogatives and indefinites. Adjectives that are derived from nominals can take inflection as well as person indexation, too.

ɔjɔn-ad-e-a=ŋ

medicine-APPL.PST.A-3SG.OBJ-FIN=1SG.SUBJ

'I gave him medicine.'

uni

He

dɔ

TOP

am-ak'-kan-a-e

2SG-GEN-COP-FIN-3SG.SUBJ

'He is yours.'

hana

that.far.JNAN

dare

tree

noa

this.JNAN

dare-ko-khɔn

tree-PL-ABL

dɔ

TOP

sẽra-gɛ-a

big-FOC-FIN

'That tree is bigger than this tree.'

Serial verb constructions

Two or more verbs and modifiers can combine together to derive a compound verb. Normally they are combinations of two transitive verbs or two intransitive verbs and limited numbers of transitive+intransitive and intransitive+transitive combinations.^[45]

ŋɛlɔam-led-e-tahẽkan-a=ko

see.find-PLUP.A-3SG.OBJ-COP-FIN=3PL.SUBJ

'They had seen and found him.'

Auxiliary verb constructions

Complex predicates are pervasive in Munda clause structure. In Santali, there are univerbated auxiliary constructions to mark many functions, such as denoting a quick, sudden, or intense action.^[46] Santali AVCs exhibit split-doubled pattern: the lexical verb may index the object argument, and the auxiliary verb may index the subject argument.

ɲɛl-ɡɔt-ke-d-e-a=pɛ

see-AUX-AOR-TR-3SG.OBJ-FIN=2PL.SUBJ

'You guys suddenly caught sight of him' or 'You guys saw him off/said good-bye to him.'

Negation

There are three particles in Santali used to express negation: *baɲ*, *ɔhɔ* and *alo*. *baɲ* and *ba* (shortened form) are the negatives for interrogative and declarative sentences; *ɔhɔ* is the emphatic negative of declarative sentences; *alo* is the prohibitive negative in the imperative. These negation particles will take away the subject marker from the verb.^[47]

ba=ko

NEG=3PL.SUBJ

'They did not catch him.'

sap-le-d-e-a

catch-ANT-TR-3SG.OBJ-FIN

alo=m

PROH=2SG.SUBJ

'Don't tell him.'

lɔi-Ø-a-e-a

tell-PRES-BEN-3SG.OBJ-FIN

Syntax

The unmarked word order of Santali is SOV, though topics can be fronted.^[48] The subject marker may appear enclitic to the verb itself if there is no preceding word.

ir-ke-t-ta-e-a=ko

cut-ASP-TR-POSS-3SG.OBJ-FIN=3PL.SUBJ

'They have cut his (paddy).'

Vocabulary

In daily conversations, Santali speakers generally employ high percentages of words of native Austroasiatic/Munda/Santali origins, compared to other Munda languages such as Kharia and Juang. Loan words, mostly borrowed from Hindi, Assamese, Bengali, Nepali, Oriya and even English may account for almost 20% of the lexemes of daily needs. Younger generation who have opportunities to engage in higher education tend to be more accustomed with lexical influence from neighboring languages as well as English.^[49]

See also

- Languages of India
- Languages with official status in India
- List of Indian languages by total speakers
- National Translation Mission
- Santali Wikipedia
- OI Chiki script

References

1. "Statement 1: Abstract of speakers' strength of languages and mother tongues – 2011" (http://www.censusindia.gov.in/2011Census/Language_MTs.html). *www.censusindia.gov.in*. Office of the Registrar General & Census Commissioner, India. Archived (https://web.archive.org/web/20190716072837/http://censusindia.gov.in/2011Census/Language_MTs.html) from the original on 16 July 2019. Retrieved 7 July 2018.
2. Santali (<https://www.ethnologue.com/21/language/sat>) at *Ethnologue* (21st ed., 2018) Mahali (<https://www.ethnologue.com/21/language/mjx>) at *Ethnologue* (21st ed., 2018)
3. "P and AR & e-Governance Dept" (<http://wbpar.gov.in/writereaddata/Issue%20of%20Gazette%20Notification%20issued%20on%2027.02.2014.pdf>) (PDF). *wbpar.gov.in*. Retrieved 10 January 2021.
4. "Redirected" (<http://www.ethnologue.com/21/language/sat/>). 19 November 2019. Archived (<https://web.archive.org/web/20190509014159/http://www.ethnologue.com/21/language/sat/>) from the original on 9 May 2019. Retrieved 9 May 2019.
5. Santali (<https://www.ethnologue.com/18/language/sat/>) at *Ethnologue* (18th ed., 2015) (subscription required) Mahali (<https://www.ethnologue.com/18/language/mjx/>) at *Ethnologue* (18th ed., 2015) (subscription required)

6. "Distribution of the 22 Scheduled Languages" (http://censusindia.gov.in/Census_Data_2001/Census_Data_Online/Language/parta.htm). *censusindia.gov.in*. Census of India. 20 May 2013. Archived (https://web.archive.org/web/20130207163538/http://www.censusindia.gov.in/Census_Data_2001/Census_Data_Online/Language/parta.htm) from the original on 7 February 2013. Retrieved 26 February 2018.
7. Sidwell, Paul. 2018. *Austroasiatic Studies: state of the art in 2018* (https://www.academia.edu/36689736/Austroasiatic_Studies_state_of_the_art_in_2018). Archived (https://web.archive.org/web/20180522203539/https://www.academia.edu/36689736/Austroasiatic_Studies_state_of_the_art_in_2018) 22 May 2018 at the Wayback Machine Presentation at the Graduate Institute of Linguistics, National Tsing Hua University, Taiwan, 22 May 2018.
8. Hembram, Phatik Chandra (2002). *Santhali, a Natural Language* (<https://books.google.com/books?id=XIkAAAAMAAJ>). U. Hembram. p. 165.
9. "Ol Chiki (Ol Cemet', Ol, Santali)" (http://scriptsource.org/cms/script_s/page.php?item_id=script_detail&key=Olck). *Scriptsource.org*. Archived (https://web.archive.org/web/20151127020321/http://scriptsource.org/cms/scripts/page.php?item_id=script_detail&key=Olck) from the original on 27 November 2015. Retrieved 19 March 2015.
10. "Santali Localization" (<http://www.andovar.com/resource-center/languages/santali-localization/>). *Andovar.com*. Archived (<https://web.archive.org/web/20160317004624/http://www.andovar.com/resource-center/languages/santali-localization/>) from the original on 17 March 2016. Retrieved 19 March 2015.
11. "Syllabus for UGC NET Santali, Dec 2013" (http://www.ugcnetonline.in/Subject_Code_95.pdf) (PDF). Archived (https://web.archive.org/web/20181106223817/https://ugcnetonline.in/Subject_Code_95.pdf) (PDF) from the original on 6 November 2018. Retrieved 4 January 2020.
12. "C-16: Population by mother tongue, India - 2011" (<https://censusindia.gov.in/nada/index.php/catalog/10191/download/13303/DDW-C16-STMT-MDDS-0000.XLSX>). Office of the Registrar General & Census Commissioner, India.
13. "SCHEDULED LANGUAGES IN DESCENDING ORDER OF SPEAKERS' STRENGTH - 2011" (<http://censusindia.gov.in/2011Census/Language-2011/Statement-4.pdf>) (PDF). *census.gov.in*. Archived (<https://ghostarchive.org/archive/20221009/http://censusindia.gov.in/2011census/Language-2011/Statement-4.pdf>) (PDF) from the original on 9 October 2022. Retrieved 17 December 2019.
14. "ABSTRACT OF SPEAKERS' STRENGTH OF LANGUAGES AND MOTHER TONGUES - 2011" (http://censusindia.gov.in/2011Census/C-16_25062018_NEW.pdf) (PDF). *census.gov.in*. Archived (https://web.archive.org/web/20181114073412/http://www.censusindia.gov.in/2011Census/C-16_25062018_NEW.pdf) (PDF) from the original on 14 November 2018. Retrieved 17 December 2019.
15. "PART-A: DISTRIBUTION OF THE 22 SCHEDULED LANGUAGES-INDIA/STATES/UNION TERRITORIES - 2011 CENSUS" (<http://censusindia.gov.in/2011Census/Language-2011/Part-A.pdf>) (PDF). *census.gov.in*. Archived (<https://web.archive.org/web/20220415174433/https://censusindia.gov.in/2011Census/Language-2011/Part-A.pdf>) (PDF) from the original on 15 April 2022. Retrieved 17 December 2019.
16. "Santhali" (<https://www.ethnologue.com/language/sat>). *Ethnologue*. Archived (<https://web.archive.org/web/20200525220301/https://www.ethnologue.com/language/sat>) from the original on 25 May 2020. Retrieved 4 January 2020.
17. "Santhali becomes India's first tribal language to get own Wikipedia edition" (<https://www.hindustantimes.com/india-news/santhali-becomes-india-s-first-tribal-language-to-get-own-wikipedia-edition/story-fUP5LRZIOIEBIVGNCjW5cK.html>). *Hindustan Times*. 9 August 2018. Archived (<https://web.archive.org/web/20190222174300/https://www.hindustantimes.com/india-news/santhali-becomes-india-s-first-tribal-language-to-get-own-wikipedia-edition/story-fUP5LRZIOIEBIVGNCjW5cK.html>) from the original on 22 February 2019. Retrieved 22 February 2019.

18. "Second language" (<https://www.indiatoday.in/india/north/story/second-official-language-of-jharkhand-governor-syed-ahmad-143959-2011-10-22>). *India Today*. 22 October 2011. Archived (<https://web.archive.org/web/20220214083406/https://www.indiatoday.in/india/north/story/second-official-language-of-jharkhand-governor-syed-ahmad-143959-2011-10-22>) from the original on 14 February 2022. Retrieved 5 November 2019.
19. Roy, Anirban (27 May 2011). "West Bengal to have six more languages for official use" (<https://www.indiatoday.in/india/east/story/west-bengal-mamata-banerjee-recognizes-six-non-bengali-languages-134507-2011-05-27>). *India Today*. Archived (<https://web.archive.org/web/20230306100207/https://www.indiatoday.in/india/east/story/west-bengal-mamata-banerjee-recognizes-six-non-bengali-languages-134507-2011-05-26>) from the original on 6 March 2023. Retrieved 5 November 2019.
20. "Glottolog 3.2 – Santali" (<http://glottolog.org/resource/languoid/id/sant1410>). *glottolog.org*. Archived (<https://web.archive.org/web/20180709154251/http://glottolog.org/resource/languoid/id/sant1410>) from the original on 9 July 2018. Retrieved 26 February 2018.
21. "Santali: Pahari language" (<http://globalrecordings.net/en/language/16318>). *Global recordings network*. Archived (<https://web.archive.org/web/20181203202736/http://globalrecordings.net/en/language/16318>) from the original on 3 December 2018. Retrieved 26 February 2018.
22. Anderson, Gregory D.S. (2007). *The Munda verb: typological perspectives*. Berlin: Mouton de Gruyter.
23. Ghosh (2008), p. 30.
24. Hildebrandt, Kristine; Anderson, Gregory D. S. (2023). "Word Prominence in Languages of Southern Asia". In Hulst, Harry van der; Bogomolets, Ksenia (eds.). *Word Prominence in Languages with Complex Morphologies*. Oxford University Press. pp. 520–564. doi:10.1093/oso/9780198840589.003.0017 (<https://doi.org/10.1093/oso/9780198840589.003.0017>). ISBN 978-0-19-884058-9.
25. Horo, Luke; Anderson, Gregory D. S.; Harrison, K. David (2024). "Vowel Harmony in the Munda Languages". In Hulst, Harry van der; Ritter, Nancy A. (eds.). *The Oxford Handbook of Vowel Harmony*. Oxford University Press. pp. 723–728. doi:10.1093/oxfordhb/9780198826804.013.57 (<https://doi.org/10.1093/oxfordhb/9780198826804.013.57>). ISBN 978-0-19-882680-4.
26. Ghosh (2008), p. 32.
27. Ghosh (2008), pp. 32–33.
28. Ghosh (2008), pp. 34–38.
29. Ghosh (2008), p. 38.
30. Ghosh (2008), p. 41.
31. Ghosh (2008), p. 43.
32. Ghosh (2008), p. 44.
33. Ghosh (2008), p. 45.
34. "Santali" (<https://web.archive.org/web/20171201040627/https://mpi-lingweb.shh.mpg.de/numeral/Santali.htm>). The Department of Linguistics, Max Planck Institute (Leipzig, Germany). 2001. Archived from the original (<https://mpi-lingweb.shh.mpg.de/numeral/Santali.htm>) on 1 December 2017. Retrieved 27 November 2017.
35. Ghosh (2008), p. 50.
36. Ghosh (2008), p. 51.
37. Ghosh (2008), p. 52.
38. Ghosh (2008), p. 53ff..
39. Ghosh (2008), p. 64.
40. Ghosh (2008), p. 85.
41. Ghosh (2008), p. 66.
42. Ghosh (2008), p. 68.
43. Ghosh (2008), p. 69.
44. Ghosh (2008), p. 70.
45. Ghosh (2008), p. 72.
46. Ghosh (2008), p. 73.
47. Ghosh (2008), p. 66-67.
48. Ghosh (2008), p. 74.

49. Ghosh (2008), p. 88.

Works cited

- Ghosh, Arun (2008). "Santali". In Anderson, Gregory D.S. (ed.). *The Munda Languages*. London: Routledge. pp. 11–98.

Further reading

- Byomkes Chakrabarti (1992). *A comparative study of Santali and Bengali*. Calcutta: K.P. Bagchi & Co. ISBN 81-7074-128-9
- Hansda, Kali Charan (2015). *Fundamental of Santhal Language*. Sambalpur.
- Hembram, P. C. (2002). *Santali, a natural language*. New Delhi: U. Hembram.
- Newberry, J. (2000). *North Munda dialects: Mundari, Santali, Bhumia*. Victoria, B.C.: J. Newberry. ISBN 0-921599-68-4
- Mitra, P. C. (1988). *Santali, the base of world languages*. Calcutta: Firma KLM.
- Зограф Г. А. (1960/1990). Языки Южной Азии. М.: Наука (1-е изд., 1960).
- Лекомцев, Ю. К. (1968). Некоторые характерные черты сантальского предложения // Языки Индии, Пакистана, Непала и Цейлона: материалы научной конференции. М: Наука, 311–321.
- Grierson, George A. (1906). *Linguistic Survey of India* (<http://dsal.uchicago.edu/books/lsi/lsi.php?volume=4&pages=701>). Vol. IV, *Mundā and Dravidian languages*. Calcutta: Office of the Superintendent of Government Printing, India.
- Maspero, Henri. (1952). *Les langues mounda*. Meillet A., Cohen M. (dir.), Les langues du monde, P.: CNRS.
- Neukom, Lukas. (2001). *Santali*. München: LINCOM Europa.
- Pinnow, Heinz-Jürgen. (1966). *A comparative study of the verb in the Munda languages*. Zide, Norman H. (ed.) Studies in comparative Austroasiatic linguistics. London—The Hague—Paris: Mouton, 96–193.
- Sakuntala De. (2011). *Santali : a linguistic study*. Memoir (Anthropological Survey of India). Kolkata: Anthropological Survey of India, Govt. of India.
- Vermeer, Hans J. (1969). *Untersuchungen zum Bau zentral-süd-asiatischer Sprachen (ein Beitrag zur Sprachbundfrage)*. Heidelberg: J. Groos.
- 2006-d. Santali. In E. K. Brown (ed.) *Encyclopedia of Languages and Linguistics*. Oxford: Elsevier Press.

Dictionaries

- Bodding, Paul O. (1929). *A Santal dictionary*. Oslo: J. Dybwad.
- A. R. Campbell (1899). *A Santali-English dictionary* (<https://archive.org/details/asantalienglish00campgoog>). Santal Mission Press.

- English-Santali/Santali-English dictionaries (<http://wesanthals.tripod.com/id54.html>)
- Macphail, R. M. (1964). *An Introduction to Santali*, Parts I & II. Benagaria: The Santali Literature Board, Santali Christian Council.
- Minegishi, M., & Murmu, G. (2001). *Santali basic lexicon with grammatical notes*. Tōkyō: Institute for the Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies. [ISBN 4-87297-791-2](#)

Grammars and primers

- Bodding, Paul O. 1929/1952. *A Santal Grammar for the Beginners*, Benagaria: Santal Mission of the Northern Churches (1st edition, 1929).
- Cole, F. T. (1896). *Santali primer* (<https://archive.org/details/santliprimer00colegoog>). Manbhum: Santal Mission Press.
- Macphail, R. M. (1953) *An Introduction to Santali*. Firma KLM Private Ltd.
- Muscat, George. (1989) *Santali: A New Approach*. Sahibganj, Bihar : Santali Book Depot.
- Skrefsrud, Lars Olsen (1873). *A Grammar of the Santhal Language* (<https://archive.org/details/aau0434.0001.001.umich.edu>). Benares: Medical Hall Press.
- Saren, Jagneswar "Ranakap Santali Ronor" (Progressive Santali Grammar), 1st edition, 2012.

Literature

- Pandit Raghunath Murmu (1925) ronor : Mayurbhanj, Odisha Publisher ASECA, Mayurbhanj
- Bodding, Paul O., (ed.) (1923–1929) *Santali Folk Tales*. Oslo: Institutet for sammenlingenden kulturforskning, Publikationen. Vol. I—III.
- Campbell, A. (1891). *Santal folk tales* (<https://archive.org/details/cu31924024159489>). Pokhuria, India: Santal Mission Press.
- Murmu, G., & Das, A. K. (1998). *Bibliography, Santali literature*. Calcutta: Biswajnan. [ISBN 81-7525-080-1](#)
- *Santali Genesis Translation* (https://archive.org/details/rosettaproject_sat_gen-1).
- *The Dishom Beura*, India's First Santali Daily News Paper. Publisher, Managobinda Beshra, National Correspondent: Mr. Somenath Patnaik

External links

-
-
- National Translation Mission's (NTM) Santali Pages (http://www.ntm.org.in/languages/Santali/default_Santali.asp)
 - OLAC resources in and about the Santali language (<http://www.language-archives.org/language/sat>)
 - OLAC resources in and about the Mahali language (<http://www.language-archives.org/language/mjx>)
 - RWAAI Repository and Workspace for Austroasiatic Intangible Heritage (<http://projekt.ht.lu.se/rwaai>)
 - Santali language in RWAAI Digital Archive (<http://hdl.handle.net/10050/00-0000-0000-0003-A6AF-2@view>)

Retrieved from "https://en.wikipedia.org/w/index.php?title=Santali_language&oldid=1289972808"