Finding Remo: A Preliminary Phonetic Analysis of the Language*

Andrew Cheng, Swarthmore College
acheng6@swarthmore.edu
December 2011

Abstract

The Remo language is a Munda language in the Austro-asiatic family that is spoken by the Remo tribe of eastern India. With fewer than nine thousand native speakers, Remo is one of the highlights of what the Living Tongues Institute has classified as the Greater South Asia Language Hotspot.

Between 1950, when Remo was first extensively written about, and the present day, very little attention has been paid to this community. In 2005, the Living Tongues Institute for Endangered Languages began a project to document Remo, to increase both awareness of and access to Remo on the Internet. As a part of this project, I was drawn to studying the unique aspects of this language.

There is a paucity of published sources and recent linguistic research on the language of Remo. What it really needs is a complete and up-to-date reference grammar and lexicon, which will move it beyond the status of an under-documented language and into the common cognizance of the linguistic community. This thesis makes progress toward this goal by filling in gaps left by previous work on Remo in the areas of its phonetics and phonology. My hope is that others may use it in the future for the purposes of investigating Munda phonology, writing the aforementioned Remo reference grammar, or even learning how to speak this hidden treasure of a language.

1 Introduction

Of the world's approximately seven thousand languages, more than half are spoken by only 0.2 percent of the global population (Harrison 2007:13). These languages are spoken in small communities that tend to be native tribes, or remnants thereof, and, for the most part, have been largely ignored by their local governing authorities. The expansion of urbanization leads speakers of these languages to abandon their mother tongue in favor of

* Many thanks are in order for the many that have helped me in the creation of this thesis: Jeremy Fahringer and Jacob Phillips for their company and assistance on the Remo team during summer research at the Endangered Languages Laboratory of Swarthmore College; Louise St. Amour and Allyson Bunch for their helpful, encouraging critiques; Professor K. David Harrison for essentially handing me a thesis topic along with a summer job and remaining more confident in my abilities than I myself am; Professor Nathan Sanders for unflinching cheerfulness and indispensable advising; and finally, Opino, Sonia, Sukari, Sudarshan, Losmi, and the rest of the Remo language community for making this project possible – I may never meet them, but their voices will remain with me for a very long time. I am solely responsible for any and all errors, linguistic or otherwise, in this thesis.

a more dominant language or a lingua franca that may be more economically advantageous. Thus, the shrinking populations of tribal communities is directly tied to the diminishing use of their languages. Unfortunately, yet unsurprisingly, language endangerment can lead to language extinction, which is marked by the death of the last speaker of a moribund language and is usually irreversible. Linguists from around the world are racing to document and preserve endangered languages before this happens.

In 2010, the Living Tongues Institute for Endangered Languages carried out fieldwork and research on Remo, an endangered language of eastern India. A significant corpus of data was compiled, and during the summer of 2011, I worked with the data on a lexicographic project, creating a "talking dictionary" of the Remo language. This dictionary is available online (Anderson and Harrison, 2005/2011) and has audio recordings available for many entries that allow users to hear the words for themselves (allowing them to make their own phonetic judgments). Even with the dictionary, however, the linguistic research on Remo is sparse, and it can be considered under-documented at best.

1.1 Overview of thesis

This thesis is an effort to synthesize the research that I have done on Remo and put to use the online dictionary that I helped create. I begin with an ethnographic sketch of the Remo tribe in Section 2. In Section 3, I explain the phenomenon of language endagerment and map out Remo's status as an endangered langauge, followed by a survey of the past linguistic work done with the Remo language in Section 4. The bulk of the linguistic analysis is presented in Section 5: basic phonetic and phonological features

of Remo, as well as particular features I find interesting, including an alveolo-palatal affricate, retroflexes, and the Remo vowel space. I conclude in Section 6 with a discussion of some aspects of Remo that may prove useful for the scientific community in the fields of linguistics and anthropology.

2 Ethnographic sketch

2.1 The Remo tribe

The Remo tribe lives in the lush tropical Jeypore Hills (also called the Bonda Hills, the Bondo Highlands, or any variation of these) in the deep southeastern part of the Koraput District, Orissa, India (Anderson and Harrison 2008). In this district, the tribal populations constitute almost 55% of the total population (Nanda 1994:191), with Remo numbers falling between 6,500 (Anderson and Harrison 2007/2011) and 9,000 (Lewis 2009), distributed among at least twenty-three small villages. The Remo lead a traditional lifestyle, practicing horticulture and hunting for food, with limited contact with the outside world. There are two main clans, or moieties, of the Remo people: the Tiger clan (*Killo bonso*) and the Cobra clan (*Ontal bonso*). The Cobra clan is the larger of the two, its members numbering about twice as much as members of the Tiger clan. The Remo tribes are unique in their bisection into just two moeities; other surrounding tribes generally have several moities (Bear, Fish, Monkey, etc.). But the Remo peoples' "real dual organization" (Elwin 1950:34) is quite unique; Elwin thinks that the Remo tribe's power must have once been vested in their tribe's two-party system, as it was the basis for their "territorial exogamy" (Elwin 1950:34) – that is, marrying outside of the clan for political purposes.

2.2 The Remo language

Remo is called *remosam* 'language of man' by its speakers. In earlier documents published on the Remo tribe and language (Verrier 1950, Bhattacharya 1968), it has been referred to as Bonda or Bondo, and on occasion still is today: Ethnologue, for example, has a corresponding entry on Bondo (ISO 639-3: bfw), but its entry on Remo describes an extinct Panoan language of Peru (Lewis 2009). I refer to the language of interest using the form closest to the people's autoglottonym for their tongue: Remo.

Remo is in the Munda family, which includes other small languages geographically concentrated in eastern India, such as Korku, Santali, Ho, Gtaʔ, and Sora (Fig. 1). Of these, Remo is most closely related to Gutob, with a 45-51% lexical similarity (Lewis 2009). Munda languages together are a part of the Austro-asiatic language family, which also includes Khmer, Vietnamese, and Mon. However, Munda languages as a whole exhibit typological patterns that differ substantially from other Austro-asiatic languages (Donegan and Stampe 2004).

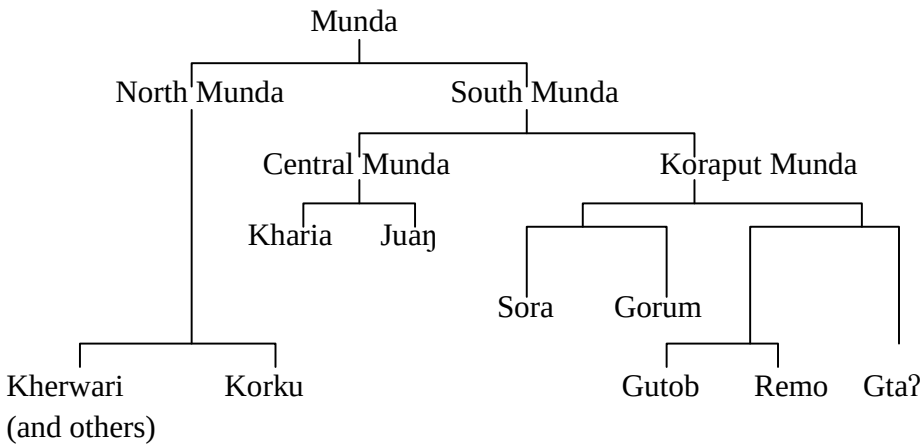

Fig. 1. Munda language family tree, replicated from Fernandez (1968).

There are two general dialects of Remo: Hill Remo (or Upper Remo), which is spoken in the villages on the Jeypore Hills, and Plains Remo (or Lower Remo), which is spoken in the Malkangiri Plain between the Jeypore Hills and the main road into the region. Fernandez claims that "there is evidence of… two [dialects] in the hills" (1967:2), although he does not go into further detail. The data collected by Anderson and Harrison between 2005 and 2011 were taken from speakers of Hill Remo. Fernandez's (1963) work on Remo focuses on the dialect spoken in and around the village of Mundlipada, which he locates in the highland region. However, most of the data compiled in Bhattacharya's 1968 glossary of Remo comes from a Plains Remo informant. (In the summer of 1962, Bhattacharya and Fernandez did work together in the field, under the tutelage of Norman Zide, and Bhattacharya (1968:xvii) claims their attempts to contact the Hill Remo were unsuccessful.) Bhattacharya says that the two dialects do not show "any very great difference" (1968:xvi).

With no standard written form, second language learners of Remo can write the language phonetically using an adaptation of the writing system of Desiya (a dialect of Oriya), an Indo-Aryan language and the dominant language of Orissa (Lewis 2009). In fact, Desiya Oriya is "fairly known and frequently used as the second language by the grown-up Bondas" (Bhattacharya 1968:xviii). However, there is no strong push for Remo to adopt a writing system (David Harrison, p.c., June 2011), and in the present time, the literacy rate among the Remo tribe is the lowest of all tribes in the Koraput district (Nanda 1994:191).

3 On Language Endangerment

Harrison (2007) has stated that hundreds of languages spoken on our planet today are facing extinction, in a way that is closely analogous to species extinction, within the next fifty years.

In order to place language endangerment in context, we must first define language extinction. A language will become extinct when its last native speaker dies (Harrison 2007). Therefore, languages in danger of extinction are those with a dwindling population of native speakers. This can happen naturally with time, as younger generations of a population make the decision (or are pushed by outside influences) to speak a different language from the one their parents' generation grew up speaking. Languages are not passed on genetically like hereditary traits; they must be acquired socially (Harrison 2007). Unfortunately, in the case of many indigenous peoples who speak tribal languages, the social capital of their native language may appear of less value compared to a more common or dominant language. There is often a societally-ingrained prejudice against indigenous peoples that may lead to shame and the refusal to linguistically associate oneself with a tribe by speaking the tribal language. As time goes on, then, the number of speakers of an endangered language can easily shrink from stable thousands, to delicate hundreds, to mere dozens.

Areas of the world where many endangered languages can be found in high densities are called "language hotspots" (Harrison and Anderson 2006). Language hotspots tend to be geographically rural areas where, despite a richness of language diversity, there is little scientific documentation of the many languages found in them.

Remo is one of many languages that make up the Greater South Asia language hotspot. The Living Tongues Institute currently does research specifically on languages of the Munda family, which constitutes the majority of this language hotspot (Fig. 2).

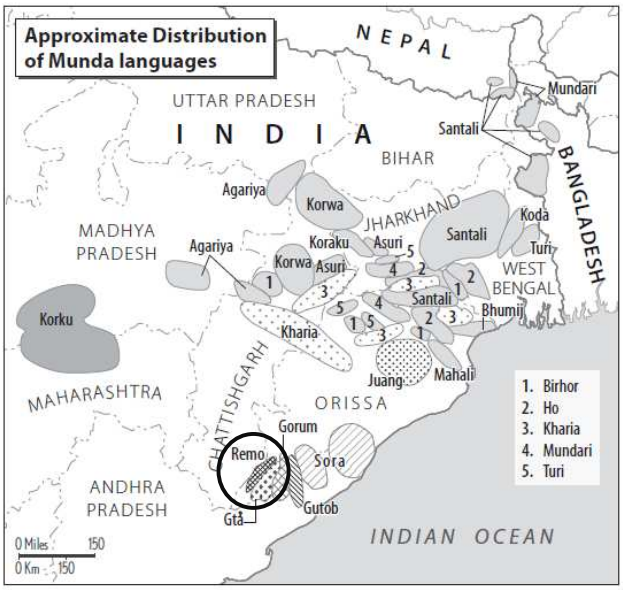

Fig. 2. Munda Languages Hotspot, with Remo territory circled. (Anderson 2007)

The Remo tribe has been the victim of a social prejudice that brands them as primitive or even violent people: "By plainsmen and officials the Bondos are regarded as entirely savage… their violent homocidal ways, their unfamiliar tongue… the inaccessibility of their abode, separate them from the rest of Orissa" (Elwin 1950:1). They are also the targets of "deliberately destructive propaganda, aimed at eliminating everything that is distinctively 'Bondo' in their lives and reducing them to the drab mediocrity of the Hindi untouchable" (Elwin 1950:7). Even their government records perpetuate the prejudice against this "backward", "ex-criminal" (Nanda 1994:10), and "useless" (May 1873) group.

According to Elwin, the Remo tribe is uniquely resistant to external pressure on

their culture or language; at least the surrounding tribes have had "singularly little influence" (1950:21) on daily life. But mainstream Hindu modes of thought are creeping – or even marching – in. Elwin describes the unexpected effect of imprisonment by local authorities on the Remo man. Occasionally, it is beneficial.

> "[He] returns home and is received … with open arms. His social position is now assured [because] he has contacted civilization and survived. He has learnt another, perhaps more than one, language and now can hold his own with police and revenue officials. … Moreover, they possess the advantage, which to the visitor is of incomparable merit, of talking a non-Austro-asiatic language. They can be understood." (Elwin 1950:22-23)

Here lies the epitome of urbanization's influence on small tribal cultures. If, by 1950, the impetus for Remo speakers to learn other languages was already present in the form of prison sentences, how much greater would that pressure be today, with all that modern life in Orissa has to offer beyond punitive governmental institutions!

The population of the Remo tribe is stable, but it is not very large, especially relative to some of its fellow Munda language-speaking tribes, such as Bhumij (100,000), Sora (300,000), and Ho (1.5 million) (Anderson and Harrison 2007/2011). The language is unwritten, which, while not a pre-requisite for language stability, does restrict greater access to the Remo language through writing, such as transcribing folklore and traditional knowledge for future generations. Also, without a writing system, the Remo language is effectively barred from many modern technological advancements offered by computers and the Internet.

In primary schools, to which Remo parents are already reluctant to send their children, the textbooks are all written in classic Oriya (Nanda 1994), a language more familiar to the teachers than to the pupils, and they depict stories from Hindu mythology

rather than traditional Remo folklore. To add insult to injury, some of the stories in the children's primer explicitly call the Bonda lifestyle a curse and patronize the entire tribe to an almost painful extent. In a learning environment like this, "it is not without reason that the most successful in the school are disenchanted with their trival setting and seek to emulate status symbols prevalent in the dominant culture" (Nanda 1994:175), including mastery of Oriya instead of Remo.

All of these factors combined lay significant pressure on the Remo language, causing it to be what Anderson and Harrison (2007/2011) have called “likely endangered”. So, in the face of an endangered language, linguists have very little time – a few decades at most – to “document the lion’s share of linguistic diversity before it vanishes forever” (Harrison 2007:206). Harrison provides a colorful metaphor to explain why it would not do to study only the most widespread languages in an effort to understand human beings’ communication systems fully:

> “Imagine a zoologist describing mammals by looking only at the top hundred most common ones. It would be easier to examine dogs and cats and cows and rabbits… But if we did, we would never know that a mammal could swim (whales), fly (bats), lay eggs (echidna), use tools (sea otters and orangutans), or have an inflatable balloon growing from its head (male hooded seal).” (Harrison 2007:211)

Likewise, the study of all languages, common or not, endangered or otherwise, is fundamentally necessary in order to contribute to our larger picture of what language looks like, of how humans communicate. “Humans rely first and foremost on language because it is the most compact and efficient channel for transmitting ideas.” (Harrison 2007:221)

4 Previously published literature

In light of the pressure to accurately document languages such as Remo, we must first trace its already-recorded history. In this section I introduce previously published literature on Remo, with an emphasis on particular milestones such as ethnographies, dictionaries, and grammars.

One of the earliest monographs on the Remo tribe, published in 1950, is *Bondo Highlander*, an ethnography written by Verrier Elwin based on his visit to the Bondo Hills. It is a colorfully-written work that describes the daily life, customs, and traditional folklore of the Remo people. Any mentions of the Remo tribe in literature before this are scant: J.A. May's "Notes on the Bhondas of Jaypur", published in *The Indian Antiquary* in 1873; some notes in the 1911 *Census of India*; and R.C.S. Bell's *Koraput District Gazetteer* in 1945, which takes much of its material from information published in older gazetteers. Elwin states that "none of these accounts do more than stress the unusual appearance of Bondo women and the remarkable organization of the pit-dormitory" (Verrier 1950:vi). However, C. Furer-Haimendorf's "Megalithic Ritual among the Gadabas and Bondos of Orissa", and "Avenues to Marriage among the Bondos of Orissa", which were published in 1943, are what first sparked Elwin's interest in the Remo people. From his anthropological perspective, he goes into great detail about rituals such as the complex process of marriage or the famous dormitory-style living quarters of Remo youth.

Bikram Narayan Nanda's *Contours of Continuigy and Change: The Story of the Bonda Highlanders* (1994), is another ethnographic work which explores the ways in which outside cultural influences have shaped the Remo community in the twentieth

century. Verrier sets the groundwork by extensively documenting the social framework of daily life in the highlands, and Nanda puts forth a social history that traces how their lives – in all respects, from agricultural techniques and festivals down to the use of umbrellas over *siali*-leaf hats (Nanda 1994:142) – have changed. However, there is very little analysis of the language itself in *Contours of Continuity and Change* and *Bondo Highlander*.

The Munda Language Research Project of the Census of India has compiled many disparate articles and unpublished works that mentioned Remo in the mid-twentieth century, but the majority of linguistic research focused on Remo was done by Frank Fernandez in 1963 and Norman Zide in 1965. Already, Remo was known to be in the Munda family of languages, and its close relationship to Gutob was clear. It also had limited references in issues of the *International Journal of American Linguistics* (1965) and the *Linguistic Survey of India* (1906 and 1927).

Sudhibhushan Bhattacharya's *A Bonda Dictionary* (1968) was the first published dictionary of Remo. While Anderson and Harrison (2008) claim that this is "still the best source on Remo vocabulary", more than one-third of the total Plains Remo words in this dictionary are loanwords from other languages, most often Desiya Oriya. Bhattacharya reasons that the admission of many loanwords into the dictionary "may help… to determine the degree and the different stages of [linguistic] acculturation" (1968:xviii).

Another available lexicon is the Munda Lexical Archive (Donegan and Stampe 2004), which Patricia Donegan and David Stampe of the University of Hawaii have been working on over the past decade or so. The Munda Lexical Archive includes a large database of Remo words; it is a compilation of data from Elwin (1950), Fernandez

(1963), and Bhattacharya (1968), along with Donegan and Stampe's original research. More recently, the organization I worked with in the summer of 2011, The Living Tongues Institute for Endangered Languages, created and maintains a searchable online "Talking Dictionary" of Remo from a corpus of elicitations collected over the past six years (Anderson and Harrison 2005/2011).

Finally, Gregory Anderson's *The Munda Languages* (2008) includes a chapter on Remo (Anderson and Harrison 2008) that thoroughly covers everything in previously published literature and includes research conducted in 2005 and 2007. Remo is also a language undergoing research in the Munda Languages Project of the Living Tongues Institute, which is a part of National Geographic's "Enduring Voices" Project. This research includes the previously mentioned online talking dictionary and brief portions of an unpublished sketch grammar.

5 Phonetics and Phonology

5.1 Methodology

The data and elicitations consulted in this section are part of the fieldwork performed by David Harrison and Gregory Anderson of the Living Tongues Institute for Endangered Languages, with support from Swarthmore College and a grant from the National Science Foundation. I, along with other student researchers at Swarthmore College, analyzed the data and processed it into dictionary format.

The consultant in the data sets from September 2010 and February 2011 is Sonia Dangada-Majhi. He is a native speaker of the Hill dialect of Remo and can also communicate in Desiya Oriya and, to a limited degree, English. Interpretation was

provided by Opino Gomango, a local project coordinator for Living Tongues who speaks Sora, Desiya Oriya, and English.

From the large pool of Remo data, I chose sound files that were the clearest and most similar to each other in terms of phoneme environment. For basic consonants, I analyzed bilabial voiced and voiceless stops, alveolar voiced and voiceless stops, retroflex voiced and voiceless stops, and velar voiced and voiceless stops in short, one- or two-syllable words where the consonant was immediately followed by a vowel. I measured release burst intensity and closure duration for all stops and voice onset time (VOT) for voiceless stops.

To determine the vowel space, I primarily looked at words that had only one vowel and took measurements from both instances of the vowel if it occurred twice in the word. I then measured the first and second formants (F1 and F2) for each vowel. For sounds of comparison languages (German, Swedish, and Hindi), I retrieved sound files from the database of Peter Ladefoged's *Vowels and Consonants* course website (2005), Urban Sikeborg's *Introduction to Swedish* course website (1997/1998), and the German Phonetic Alphabet website (Battenburg and Swanson 1999/2000). All measurements were made using Praat software (Boersma and Weenink 1992/2011), and all images of spectrograms, waveforms, and spectral slices are screenshots from Praat software.

5.2 Consonants

5.2.1 Consonant Inventory of Remo

Based on the existing literature (Bhattacharya 1968, Fernandez 1967, Anderson 2008), Table 5.1 below is the most recent published consonant inventory of Remo. In the

following sub-sections, I challenge some of these purported phonemes and propose alternatives or additions to the inventory.

Table 5.1. Consonant inventory of Remo, adapted from Anderson and Harrison (2008) and Fernandez (1968).

	Labial	Dental[1]	Alveolar	Retroflex	Fronto-palatal	Velar	Glottal
Stops	p b	t d		ʈ ɖ		k g	ʔ
Affricates					ǰ[2]		
Fricatives			s z				
Nasals	m	n		ɳ	ɲ	ŋ	
Laterals		l					
Trills		r					
Approximants	w				y[3]		

5.2.2 Basic measurements

Table 5.2 below shows the release burst intensity, closure duration, and voice onset times of the labial, alveolar, and velar stops in Remo[4]. A complete table that includes all raw measurements, rounded means, and rounded standard errors can be found in Appendix A.

Table 5.2. Remo plosives' burst intensity, closure duration, and voice onset time, in rounded means.

Phoneme	Release burst intensity (dB)	closure duration (sec)	VOT (sec)
/b/	79	0.088	-
/p/	76.4	0.075	0.029
/d/	75	0.07	-
/t/	77.3	0.062	0.025
/g/	78	0.035	-
/k/	73	0.066	0.036

[1] Fernandez (1968) identifies /n/, /l/, /r/, /d/, and /t/ as dental, but Anderson and Harrison (2008) place them in the same (unlabeled) column as /s/ and /z/. I treat all seven of these phonemes as "dento-alveolar", accepting that I am uncertain as to their actual place of articulation.

[2] Fernandez (1968) and Anderson and Harrison (2008) use the symbols ‹ǰ› and ‹j›, respectively, to denote a voiced fronto-palatal affricate, not to be confused with the IPA symbol /j/ used to denote the (fronto-/pre-) palatal approximant (as in the English /ju/ "you"). I explore this phoneme more thoroughly in Section 5.2.3.

[3] Fernandez (1968) and Anderson and Harrison (2008) use the symbol ‹y› to denote the (fronto-) palatal approximant that IPA denotes with /j/, not to be confused with the IPA symbol /y/ used to denote the close front rounded vowel (as in the French /ty/ "you"). I use the current IPA symbol /j/ to denote the (fronto-/pre-) palatal approximant, as will be reflected in the revised consonant inventory in Section 5.2.5.

[4] I address the retroflex voiced and voiceless stops in a different manner in section 5.2.4.

5.2.3 Mystery voiceless affricate in Hill Remo

Anderson and Harrison (2008) do not indicate any affricates in their consonant inventory of Hill Remo; their inventory is adapted from Fernandez (1968), who does identify a voiced "lax fronto-palatal affricate" (Fernandez 1968:9) that likely corresponds to IPA [d͡ʒ] or [d͡ʑ]. Anderson and Harrison (2008) identify this sound with the symbol ‹j›, as well, but give no further information about it. Notably, Fernandez and Anderson do not identify a voiceless counterpart for this affricate in Hill Remo. However, Bhattacharya briefly mentions a "voiceless unaspirated palatal affricate" found in a few loanwords from Desiya Oriya, which "in most cases … is changed to *s* in [Remo]" (Bhattacharya 1968:xxvi). He uses ‹c› to denote this sound[5]. Some words from Bhattacharya's dictionary that contain this so-called affricate include:

(1) cinta 'thought'
ciʈka 'garden lizard'
cer 'root'
capoʈ[6] 'slap'
konca 'piece'

As it so happens, these are *all* marked as loanwords from Desiya Oriya, and thus we cannot be sure if this affricate occurs naturally in the language. Elicitations from the 2005-2011 corpus of data do not include any of these words, and in only a few cases have I been able to find other, similar voiceless affricates (2)[7].

[5] Note that the IPA symbol /c/ is used to represent a voiceless palatal stop, not an affricate.
[6] The question of the Remo retroflex trill is addressed in section 5.2.4.
[7] For reasons of user-friendliness and simplicity, the Remo Online Talking Dictionary uses the affricate /t͡ʃ/ to represent this still-unknown phoneme. The actual phoneme may have a different place of articulation, but / t͡ʃ/ is nevertheless a good approximation.

(2) ʔ t͡ʃaˈʈa 'four'
ʔ t͡ʃoɖa 'six'
ʔ t͡ʃatətə 'seven'
ʔ t͡ʃimta 'sad'
ʔ t͡ʃire kerõ 'white paddy'
ʔ t͡ʃovoʔniŋ 'catch'
t͡ʃini 'sugar'

The word *t͡ʃini* 'sugar' is the only one of these words that has a direct parallel in Bhattacharya 1968: the entry *sini*, with a note that it is a loan from Desiya Oriyia's *cini*. Most of the remaining words appear in Bhattacharya 1968 in apparently unrelated forms and in the Munda Lexical Archive in potentially related forms, albeit still lacking the affricate phoneme. Table 5.3 below shows a comparison.

Table 5.3. Comparison of several words from the elicited corpus of Hill Remo, Bhattacharya's 1968 dictionary, and the Munda Lexical Archive (Donegan and Stampe 2004), with entries in the latter that are marked as being loanwords in a separate column.

English	2005-2011 Elicitations	Bhattacharya 1968	MLA 2004	MLA 2004 (marked loan)
four	t͡ʃaˈʈa	uʔuŋ	uʔuñ	sar
six	t͡ʃoɖa	tiʔri	tiʔiri	sogota
seven	t͡ʃatətə[8]	giʔ	giʔ	satgota
sad	t͡ʃimta	-	-	-
white paddy	t͡ʃire kerõ	-	-	-
catch	t͡ʃovoʔniŋ[9]	sop	sop	-
sugar	t͡ʃini	sini	nabaed	sini

There are two things about these parallel data that I find highly unusual. The first is that Hill Remo would have the voiced affricate occurring in abundance while its voiceless counterpart almost never occurs, and even then, only in loanwords. The second is that the words from the 2005-2011 data (which are elicitations of Hill Remo) that

[8] In the 2005-2011 elicitations, satətə also occurs.
[9] In the 2005-2011 elicitations, sovoʔniŋ also occurs.

include a voiceless affricate are basic words such as numbers and emotions that are unlikely to have been borrowed from other languages and so thoroughly integrated into native speakers' vocabulary in a mere fifty years[10].

In the following, I offer a phonetic analysis of this mystery sound in order identify it. I will henceforth refer to it as Phoneme X, found in the Remo words 'four' and 'six' as spoken by Sonia Dangada-Majhi.

The duration measurements for Phoneme X (Table 5.4) clearly indicate that Phoneme X is an affricate when compared to a Remo stop and a Remo fricative. Phoneme X has a fricated stop release with a duration of only 0.053 sec, which is much shorter than the duration of the Remo fricative /s/. The closure duration of the stop portion of Phoneme X is also shorter than a typical Remo stop.

Table 5.4. Comparison of stop closure duration and release duration of Remo coronal stops, fricatives, and Phoneme X. A complete data table can be found in Appendix B.

	/t/	/s/	Phoneme X
Stop closure duration (sec)	0.0714	-	0.0176
Fricative Duration (sec)	-	0.180	0.053

A look at the spectrogram for Phoneme X (Fig. 3) also provides evidence that this sound is an affricate. It is longer in duration than a stop and also has the diffuse white noise typical of sibilants, complete with dark bands at specific frequences, which I use to identify the place of articulation in a later sub-section.

[10] I am holding out on the possibility that the dialects of Hill Remo and Plains Remo may be different enough that even words for basic numbers sound unrelated due to what Bhattacharya states about the two dialects not showing "any very great difference" (1968).

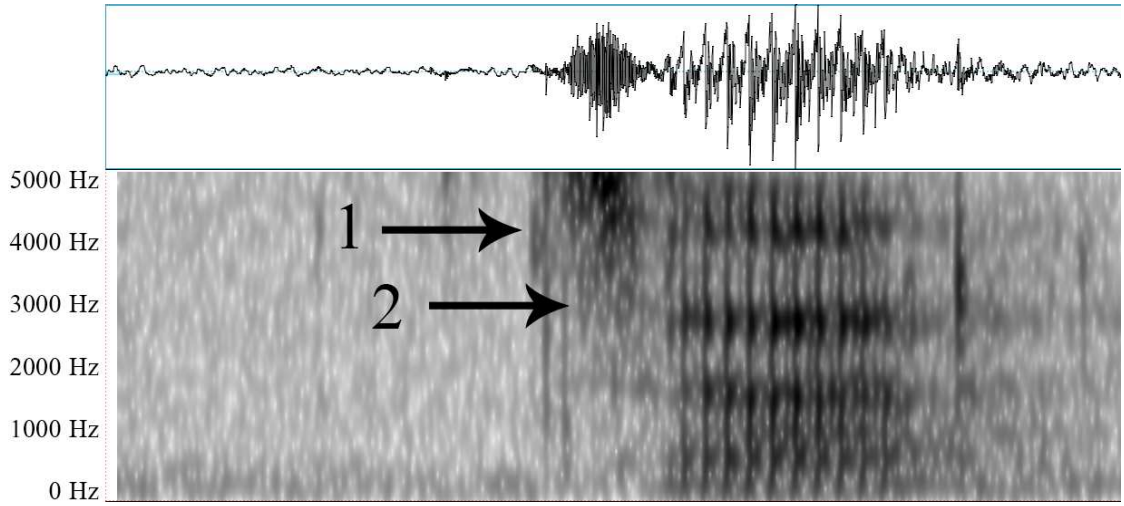

Fig. 3. Spectrogram of [Xa] from Remo *Xaʈa* 'four'. Phoneme X is an affricate with a stop portion (arrow 1) and a fricated release portion (arrow 2).

To test for voicing, I compare Phoneme X to both the voiced palato-alveolar affricate [d͡ʒ] and the voiceless palato-alveolar affricate [t͡ʃ] in Hindi. The waveform and spectrogram image for the Hindi word /d͡ʒal/ ‘net’ (Fig. 4), we can see the pitch track of the spectrogram (automatically generated by Praat with the dotted blue line) and the periodic wave during the closure, indicative of voicing in the waveform on the top half. When this is compared to the waveform and spectrogram image for the Hindi word /t͡ʃal/ ‘turn’, we see that there is clearly no voicing in /t͡ʃal/. While the closure seems to have a voicing bar in the spectrogram in Fig. 5, Praat calculates no pitch track, so there appears to be just background noise, and not voicing; this is especially obvious in the waveform for /t͡ʃal/: it is an essentially flat line during the closure of this sound, indicating almost no noise at all, let alone voicing.

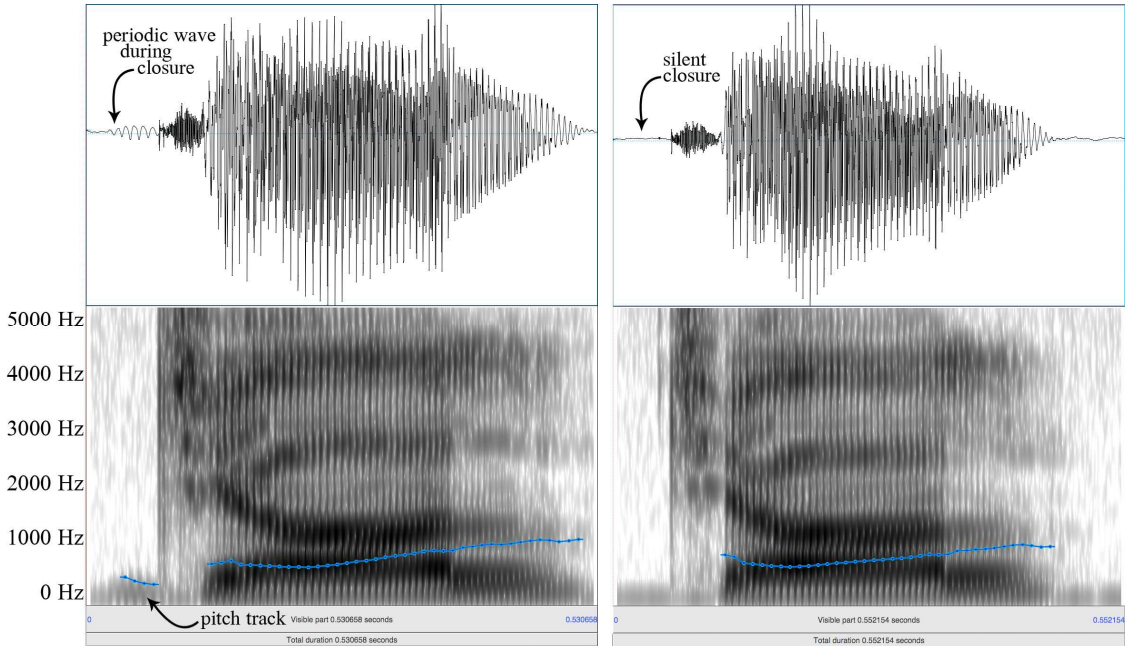

Fig. 4. Hindi [d͡ʒal] 'net'

Fig. 5. Hindi [t͡ʃal] 'turn'

Next, I compare Phoneme X with the voiced palato-alveolar affricate in Remo /d͡ʒ/. Fig. 6 shows that /d͡ʒ/ is voiced, indicated both by the darker voicing bar around 500 Hz in the spectrogram and the blue pitch track calculated by Praat, as well as by the periodicity during the closure of the waveform. Fig. 7 has a dark band around 500 Hz as well, but this is clearly caused by the background noise during the elicitation, because in the waveform we can see softer, aperiodic white noise rather than louder, periodic repetition during the closure. In comparison, it is clear that Phoneme X is unvoiced.

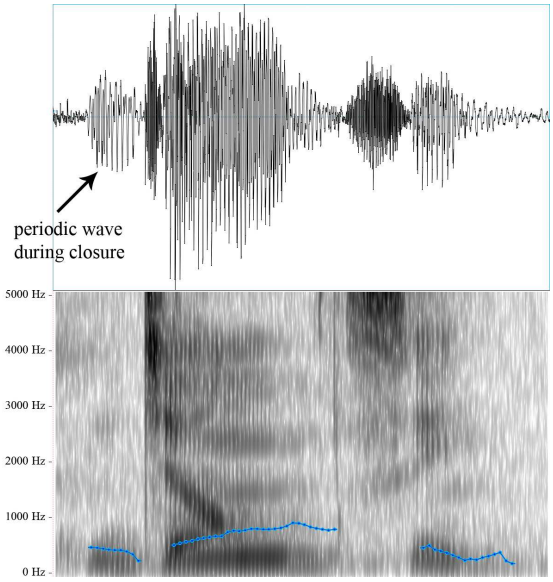

Fig. 6. Remo [d͡ʒoŋsiŋ] 'hen'

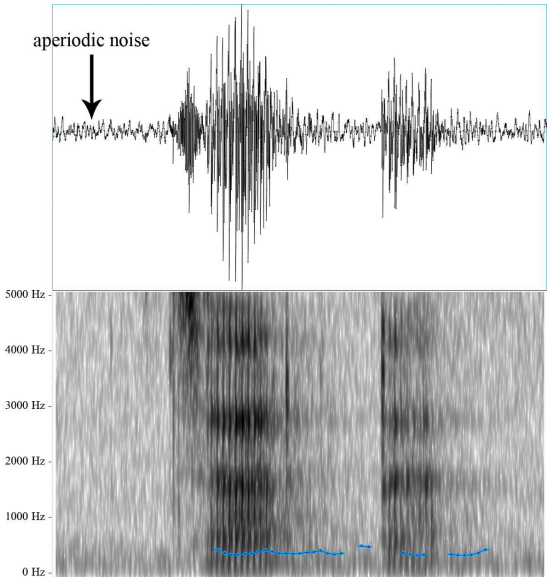

Fig. 7. Remo [Xaʈa] 'four'

In order to definitely determine where in the oral cavity Phoneme X is articulated, it would be helpful to take video recordings of native speakers, particularly their tongue and mouth area, and actually look. However, it is also possible to glean some information from Praat and a spectrogram of Phoneme X in order to determine where it is articulated. From Ladefoged's (2005) database, Sikeborg's *Introduction to Swedish* (1997/1998), the German Phonetic Alphabet website (Battenburg and Swanson, 1999/2000), the Remo Online Talking Dictionary (Anderson and Harrison 2005/2011), and my own elicitations, I obtained examples of fricatives at various places of articulation: alveolar /s/, post-alveolar /ʃ/, alveolo-palatal /ɕ/, and palatal /ç/.

For the centers of gravity, I used Praat to splice fricatives from different words together to create one large "synthesized fricative". For example, when eliciting the alveolar fricative /s/ in English, I recorded myself eliciting five words: say, see, sigh, so, and sue, and then spliced together a narrow portion of the /s/ in each word to create the

synthesized fricative. Praat automatically calculated the centers of gravity for each synthesized fricative in Hertz. These values are given below in Table 5.5.

Table 5.5. Center of gravity over fricative at power 2.0

Phoneme	/s/	/ʃ/	/ɕ/	/ç/	Phoneme X
comparison language	English: say, see, sigh, so, sue	English: shea, she, shy, show, shoe	Swedish[11]	German[12]	Remo[13]
center of gravity (Hz)	5347.13	3629.98	4062.16	4802.29	4902.83

At 4902.83 Hz, Phoneme X has a higher center of gravity than almost all of the other fricatives. It appears to be closest to the palatal fricative /ç/ and the alveolar fricative /s/, and somewhat close to the palato-alveolar fricative /ɕ/. However, Phoneme X could not be /ts/ because it sounds nothing like the alveolar affricate. We can use spectral slices to further compare fricatives visually; figures 8-12 are the spectral slices from individual examples of each fricative and the fricative portion of Phoneme X[14]. The spectral slice of the fricative release of Phoneme X (Fig. 8) is closest in form to /s/ (Fig. 9) and /ɕ/ (Fig. 11): both LPC-smoothed spectral slices have a sharp dip on the left side and descend with few peaks to the bottom right corner. Having already ruled out /s/, /ɕ/ is a reasonable match for Phoneme X.

[11] The value for the center of gravity for the Swedish /ɕ/ was calculated by creating a synthesized fricative from the words *kjol* 'skirt', *kilo* 'kilo', *check* 'check', and *tjúgo* 'twenty' (Sikeborg 1997/1998).
[12] The value for the center of gravity for the German /ç/ was calculated by creating a synthesized fricative from portions of fricatives from Battenburg and Swanson 1999/2000: the sound *ch* itself and *dich* 'you'.
[13] The value for the center of gravity for Phoneme X was calculated by averaging the centers of gravity for Remo 'four' and 'six'.
[14] I obtained spectral slices for individual words rather than the synthesized fricatives for greater precision and used LPC smoothing to clarify the overall shape and distribution of peaks.

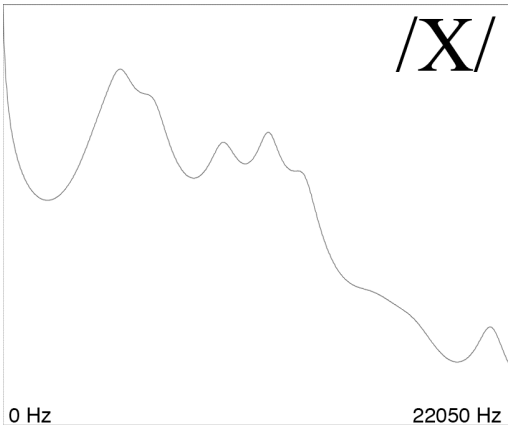

Fig. 8. Spectral slice of Phoneme X.

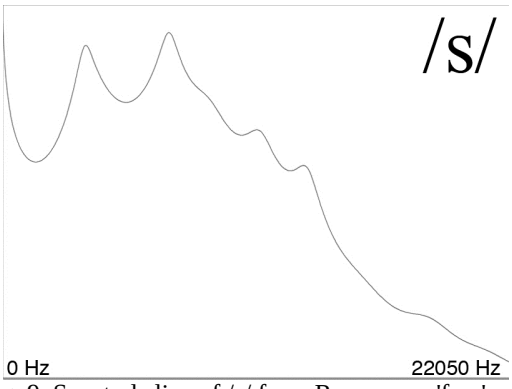

Fig. 9. Spectral slice of /s/ from Remo *soro* 'free'

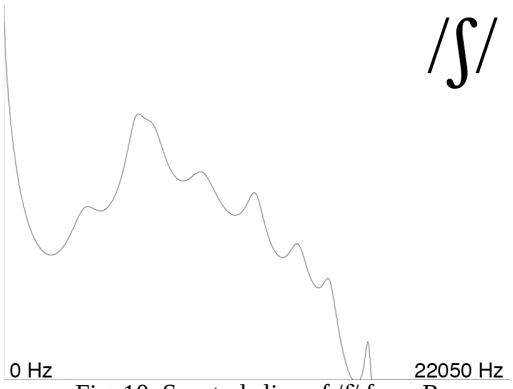

Fig. 10. Spectral slice of /ʃ/ from Remo *logɻai ʔi ʔditupa* '(you-DL) were falling down'

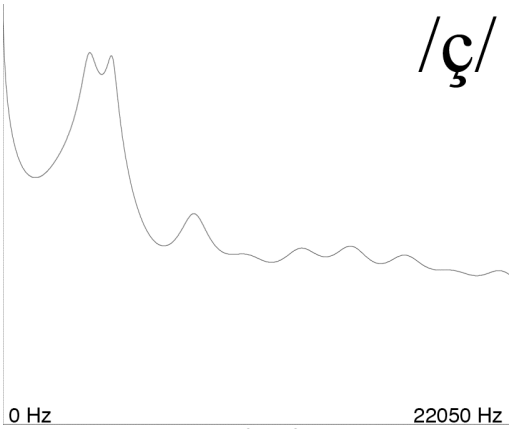

Fig. 11. Spectral slice of /ç/ from Greek /çɛɾi/ 'hand'

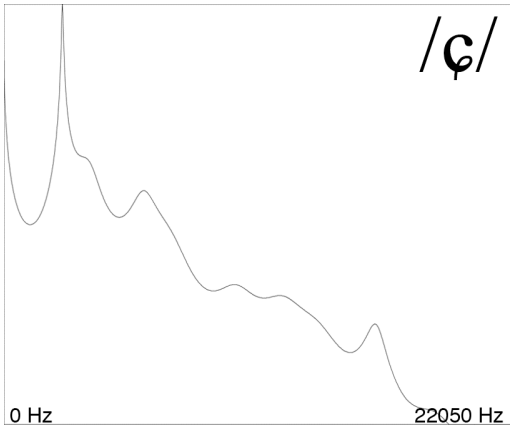

Fig. 12. Spectral slice of /ɕ/ from Swedish /ɕol/ 'skirt'

In conclusion, I believe Phoneme X is the voiceless alveolo-palatal (also called pre-palatal or fronto-palatal) affricate: /t͡ɕ/. If this is truly the case, then it makes more

sense for the voiced post-alveolar affricate that I have been citing in Remo words such as *d͡ʒoŋsiŋ* 'hen' to actually be the voiced alveolo-palatal affricate /d͡ʑ/, which creates a more standard voiced-voiceless pair.

Neither in Bhattacharya's (1968) dictionary nor in the Living Tongues Institute's Online Talking Dictionary (Anderson and Harrison 2005/2011) is this phoneme accounted for. It remains a question for historical linguists where exactly these curiously affricated words: *t͡ɕaʈa* 'four' and *t͡ɕoda* 'six' came from and when. There are no cognates for these two numbers in other Munda languages, including Gtaʔ, Ho, Bhumij, or Sora; however, Desiya Oriya's *chaari* 'four' and *cha'a* 'six' (Beaumont 1997/2011) are likely candidates.

5.2.4 Remo retroflexes

It is somewhat a curiosity of the Munda languages that they should have retroflex consonants, although this is actually unsurprising. As an Austro-asiatic language, Remo should not have any inherited retroflexes. However, although only 11% of the world's languages have retroflex stops (Ladefoged and Bhaskararao 1983:292), all of the Indo-Aryan languages spoken in the areas surrounding the Remo tribes (Oriya and Hindi in particular) use retroflex (or retracted) stops, nasals, and liquids, so the incorporation of the retroflex place of articulation into Remo seems reasonable.

Bhattacharya's dictionary (1968) adopts an alphabet based on an Indian system used at the time, which includes symbols for the retroflex stops /ɖ ʈ/, the retroflex nasal /ɳ/, and a retroflex flap /ɽ/. As a native speaker of a language that does not distinguish

between retroflex sounds and other coronal places of articulation, I found it difficult to hear if Remo really does have these retroflex consonants, or if it is partly due to some unconscious bias toward the exotic that I hear a retroflex when there is none.

Bhattacharya states that the retroflex nasal /ɳ/ only occurs in a few words loaned from Desiya Oriya. Fernandez (1967) also indicates that there is a retroflex nasal, but, oddly enough, there are almost no examples of it given in either of these sources. The sole instance of a retroflex nasal in Bhattacharya (1968) is the word *roɳa* ‘a man of the Raɳa community/caste’, which is a loanword from Desiya Oriya. Furthermore, the sole instance of a retroflex nasal in Fernandez (1967) is *ɲoɳɖa* 'honey', which when elicited by Sonia Dangada-Majhi in 2010, sounded closer to [ɲõːɖaʔ], without a second nasal stop. A closer look at the spectrogram (Fig. 13) reveals that Sonia's 'honey' does not have a retroflex nasal between the vowel and the oral stop – indeed, it appears not to have any sound at all.

While this is not conclusive evidence against the presence of a retroflex nasal in Remo, I feel that there is a substantial-enough lack of occurrences of this phoneme that I drop it from the rest of my analysis of Remo retroflexes.

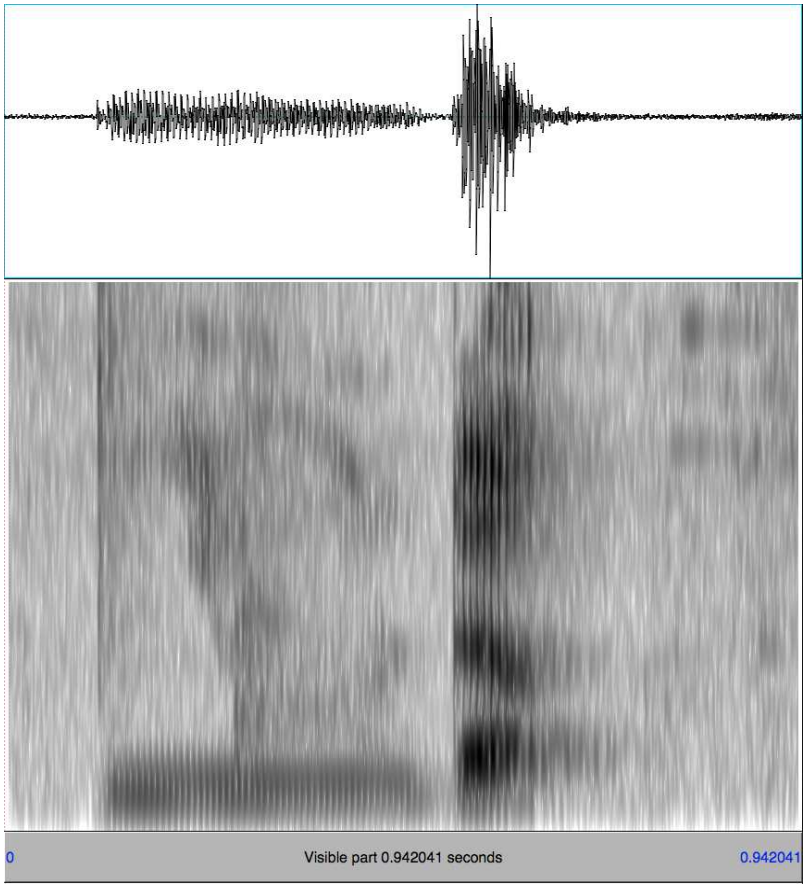

Fig. 13. Remo [ɲõːɖaʔ] 'honey'

From Bhattacharya's dictionary's corpus (1968), there are a few minimal pairs that distinguish retroflex stops from alveolar stops:

(3) dug 'be dirty'
ɖug 'flavor (food)'

(4) duma 'deer species'
ɖuma 'ancestor'

(5) tai 'dig up'
ʈai 'creeper species'

(6) tip 'drip'
ʈip 'topmost part'

The existence of minimal pairs indicates that the retroflex position is contrastive in Remo. However, there are nearly four times as many entries under *ḍ* (Bhattacharya's symbol for the voiced retroflex stop [ɖ]) than under *d*, which raises suspicions that /d/ and /ɖ/ may actually be allophonic. Another point of inquiry is the fact that most of the words

that begin with /d/ in the dictionary are labeled as loanwords from Desiya Oriya. For the voiceless versions of these stops, Bhattacharya lists almost seven times as many words beginning with *t* than with *ṭ* [ʈ], and notes that the latter also "occurs mostly in the Desiya [Oriya] loanwords" (1968:xxvi). While I expected a more uniform distribution of these phonemes in the dictionary, Zide (1970) has shown from a reconstruction of the proto-Munda consonant system that a paired alveolar (or dental) stop /t/ and retroflex voiced stop /ɖ/ is not only possible but likely. "The asymmetric points of articulation of *t* and *ḍ* are preserved in most of the modern languages, where they correspond in loan phonology to the dental *t* and retroflex *ḍ* of Indo-Aryan" (Zide 1970:414), he writes of the plausible origins of this unusual pairing.

Although there are no recordings from Bhattacharya's corpus to analyze in order to get a clearer picture of the places of articulation, the recordings from the 2005-2011 elicitations are readily accessible online. Unfortunately, none of the Remo words from (3-6) were elicited in the 2005-2011 data. (In two instances, a different Remo word than is attested in Bhattacharya was given for the English: the word *inostabaj* was given for 'dirty', and *gugu?* for 'dig'.) To measure retroflex stops, I chose instead several other words that were listed in Bhattacharya (1968) and also elicited by Sonia Dangada-Majhi in 2010. I used Praat to measure the second and third formant values during the release burst of the stop. I also calculated the average burst intensity in decibels during the center of the release burst. Then, I calculated the rounded means and standard errors for the four sounds, as recorded in Table 5.6 below.

Table 5.6. Formant measurements and burst intensity in rounded mean for several purported retroflex stops. Full data table can be found in Appendix C.

Phoneme	F2 (Hz)	F3 (Hz)	Burst intensity (dB)
/t/	1740	2900	75
/ʈ/	2200	3040	64
/d/	1600	2770	78
/ɖ/	1920	2900	73

Hamann (2003) says that all retroflexes share three articulatory characteristics: "raising of the tongue tip (apicality), a sublingual cavity, and retraction of the tongue body." Not being able to see the vocal tracts of the elicitors at the present moment, however, I look instead for acoustic cues in order to find the retroflex of Remo. Hamann shows that when studying spectral information about a retroflex, sound duration, voice onset time (VOT), formant levels, and burst intensity can all give important information about the articulation. I chose to measure formant values and burst intensity.

Retroflexes show a convergence of F2 and F3 in all vocalic environments. Retroflexes have a very low F3, but F2 is largely context-dependent. A low F3, in fact, “is sufficient enough to distinguish retroflexes from other coronals." (Hamann 2003:63) There will also be asymmetry in the transitions in and out of the segment, with vowel-to-consonant transitions especially distinct. In the word *aɖoŋ* 'fish' (Fig. 14), the asymmetry of the formants is clearly visible on both sides of the closure, especially in the vowel [a] before the closure; the disruption of the preceding vowel is a trademark of retroflex stops.

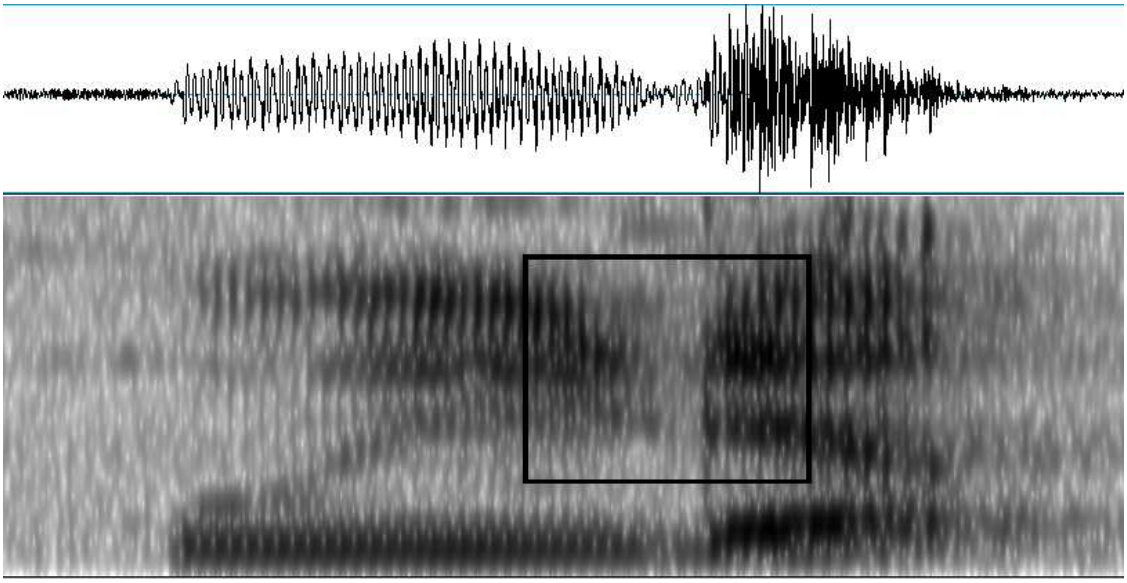

Fig. 14. Remo [aɖoŋ] 'fish', with the formant asymmetry indicated within the box.

Judging from the numerical data in Table 5.6, the F2 values for the purported voiced and voiceless retroflex stops are greater than the F2 values for the voiced and voiceless alveolar stops, but the F3 values also appear to be slightly greater, as well. There does not seem to be a discernible pattern for the formant “pinch” that F2 and F3 are expected to display. This is potentially a problem, as it may indicate one of several things. There may be gross errors in measurement that have caused the values to be different from what was expected. It is also possible that the errors are not in the measurement but rather in the classification of the words I measured and their groupings. Since all the values are so close together, it is conceivable that some retroflex stops were measured as alveolar stops and vice versa, and this purely by accident, not ideal but also not surprising for someone whose ear cannot yet perfectly distinguish retroflex sounds from other, phonemically similar ones.

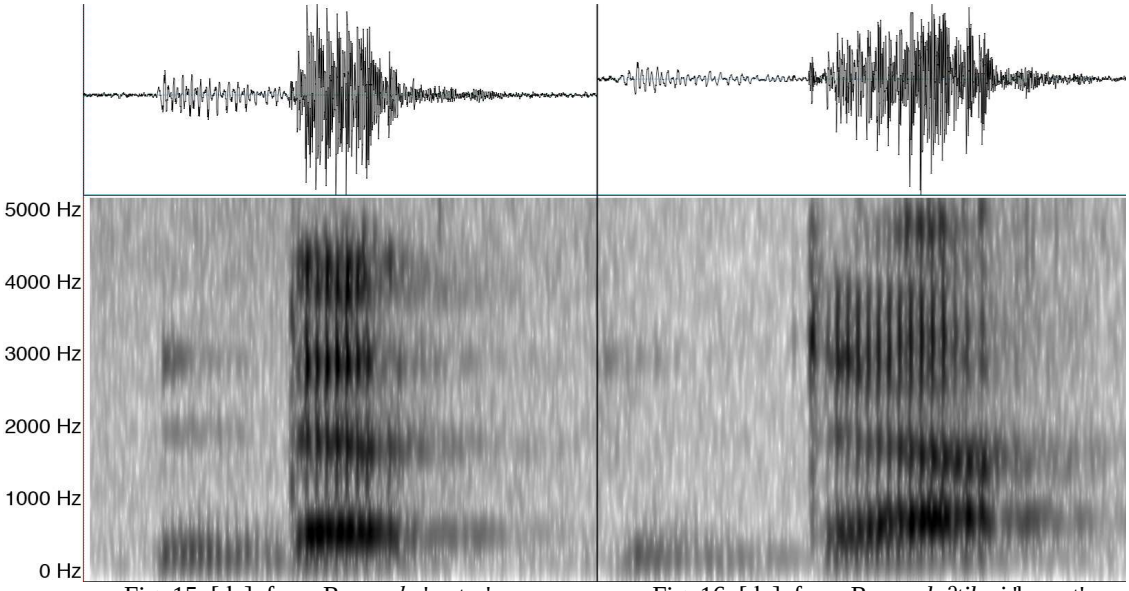

Fig. 15. [ɖa], from Remo *ɖa* 'water'

Fig. 16. [ɖa], from Remo *daʔtikwi* 'breast'

Figures 15 and 16 juxtapose a purported voiced retroflex stop and a purported voiced alveolar stop for strict comparison. The retroflex stop does show a slight pinch of F2 and F3 coming together, which is not quite there on the alveolar stop. With the evidence inconclusive, I end this sub-section with a call to embark on more research on the differentiation between voiced and voiceless retroflex and alveolar sounds in Remo. With the available data in the corpus a bit lacking, it would be highly useful to look for more examples of retroflex-alveolar minimal pairs and document them.

Lastly, Bhattacharya calls the retroflex flap /ɽ/ “cacuminal”, which is closer to the Hindi retroflex – which is apical, meaning only the tongue tip is involved in articulation – than to the Tamil retroflex, in which the entire bottom of the tongue rests against the palate. Fernandez (1968) suggests that the Remo /ɽ/ is an allophone of the retroflex stop /ɖ/, occurring intervocalically and following /ʔ b g m/, and describes the Remo /r/ as an "apico-alveolar trill". Subsequently, Anderson and Harrison (2008) do not discuss a

retroflex trill or flap of any sort but do list /ɽ/ in a table of possible consonant clusters (2008:561). I would like to think that my ear has been trained enough to be able to distinguish the seemingly common laterals in the Remo elicitations, but there is no guarantee of this. Due to this difficulty, I have not attempted to search for the minimal pairs that would distinguish between /r/ and /ɖ/ (much less between /r/ and /ɽ/) in the 2005-2011 recorded data, but I settle on Fernandez's identification of /r/ as a discrete phoneme and /ɽ/ as an allophone of /ɖ/. Consequently, further and more careful research than I have had the time or resources to carry out is suggested.

5.2.5 Revised Remo consonant table

Based on conclusions drawn from previous sections, I now offer an updated consonant inventory of Remo (Table 5.7). This table is adapted from Anderson and Harrison (2008), which was in turn adapted from Fernandez 1968. I have shown that Remo has two affricates, the voiced /d͡ʑ/ and the voiceless /t͡ɕ/, which are alveolo-palatal (pre-/fronto-palatal), and have correspondingly put them in their own column, moving the palatal nasal and palatal approximant (which I have symbolized with the official IPA ‹j›, instead of the English-friendly ‹y›) to a Palatal column. I have also discussed the lack of evidence for a retroflex nasal /ɳ/ in Remo and have thus removed it, but have added the retroflex flap /ɽ/ in parentheses to indicate that it is an allophone. Lastly, I have moved /r/ from the Dental column to the Alveolar column, in accordance with current IPA standards.

Table 5.7. Revised donsonant phonemes of Hill Remo, adapted from Anderson and Harrison (2008) and Fernandez (1968).

	Labial	Coronal			Dorsal			Glottal
	Bilabial	Dental	Alveolar	Retroflex	Alveolo-palatal[15]	Palatal	Velar	Glottal
Stops	p b	t d		ʈ ɖ			k g	ʔ
Affricates					t͡ɕ d͡ʑ			
Fricatives			s z					
Nasals	m	N				ɲ	ŋ	
Laterals		L						
Trills			r					
Flaps				(ɽ)				
Approximants	w					j		

5.3 Vowels

I used Praat software to analyze Remo words and determine what the exact vowel space of Remo is. It was fairly straightforward to find good examples of all vowels from one speaker; I choose elicitations of the best quality and of as many words as possible in which the vowel occurred more than once. After selecting the center of the vowel in Praat, the software automatically calculated its first formant and second formant values (in Hertz). I took the average value (calculated mean with standard error) for each vowel and then plotted them on a F1xF2 chart.

5.3.1 Vowel Inventory of Remo

The vowel inventory of Remo is very straightforward with five basic vowels. There is also a non-phonemic schwa that appears most often following an unstressed /e/ (Table 5.8). These vowels can be combined into a variety of diphthongs, however, as

[15] Pre-palatal/fronto-palatal

with most Munda languages. In Remo, vowel nasalization is contrastive, albeit limitedly so (there are only a few accounted-for minimal pairs) (Anderson 2008:558).

Table 5.8. Vowel phonemes of Remo (Anderon 2008)

	Front unrounded	Central unrounded	Back rounded
High/close	i		u
Mid	e	(ə)	o
Low/open		a	

5.3.2 Formants for Remo Vowels

In Table 5.9 below is the rounded mean and standard errors of first and second formant values for the vowels of adult male Remo speaker Sonia Dangada-Majhi. These values are plotted onto the F1xF2 chart in Fig. 17.

Table 5.9. Mean F1 and F2 values for Remo vowels. A complete data table can be found in Appendix D.

Vowel	F1 (Hz)	F2 (Hz)
/i/	330	2440
/u/	380	1100
/e/	480	2150
/a/	770	1580
/o/	540	1130

Future research possibilities include measuring vowel diphthongs; while they are not too common, many combinations have been attested, including /aɪ/, /oɪ/, /au/, /iu/, and /io/ (Anderson 2008). Also unexplored is the role of nasalization in its "limited" (Anderson 2008) occurrence as a phonemic characteristic. Remo is on the whole a very nasal language, and it would be worth exploring this aspect of it in greater detail.

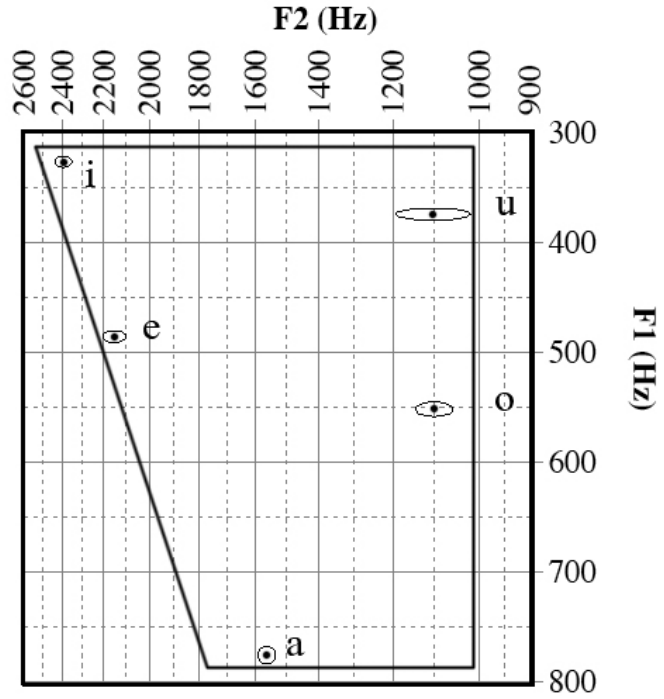

Fig. 17. F1xF2 chart of Remo vowel space.

6 Conclusion

It is undeniably important for people to study languages, especially endangered languages, in order to save irreplaceable cultural knowledge and living artifacts of humanity (that is, the phenomenon of language) before it is too late. I have looked at Remo consonants and determined the existence of a phoneme previously unaccounted-for in the literature. I have also shed light on the details of other aspects of the language, such as acoustic measurements of some consonants and vowels that are surprisingly different from what one might expect. All of this goes to show that Remo is not just another language, but one that likely offers more to the scientific community than meets the eye. Fortunately, there is much more than can be explored in this language and in the other Munda languages. With the Remo Online Talking Dictionary freely available as a continually-updated resource, it is entirely possible for anyone interested in taking a look

at the language to access most of the same raw data that I worked with, and I encourage all efforts to increase its use and widen the public's exposure to the work that Living Tongues and other bodies of dedicated linguists are doing for the benefit of the globe's quickly-disappearing languages. This thesis has been a very brief, short look at Remo, but my hope is that this is just the starting point of much more that is to be discovered from within the Bondo Hills.

Appendices

Appendix A: Remo consonant measurements

/b/	release burst intensity (dB)	closure duration (sec)
tulibaiʔbieʔ 'white ant'	76.13168588	0.095115
gubu 'pig'	83.49831355	0.084685
gubu 'pig'	80.19963267	0.082393
gubu 'pig'	79.23580756	0.089957
gubu 'pig'	75.26102472	0.087605
rounded mean	**79**	**0.088**
rounded s.e.	**1**	**0.002**

/p/	release burst intensity (dB)	closure duration (sec)	VOT (sec)
awitupe '(you) aren't going'	76.82565504	0.071075	0.031389
awitupe '(you) aren't going'	77.9341301	0.08891	0.018734
wipe 'you all went'	76.23650931	0.077218	0.036481
oi ponte 'yet'	74.73315568	0.064556	0.030632
rounded mean	**76.4**	**0.075**	**0.029**
rounded s.e.	**0.7**	**0.005**	**0.004**

/d/	release burst intensity (dB)	closure duration (sec)
pada 'ground'	78.02262306	0.080615
udoi 'south'	70.84896242	0.040414
dadi 'grandfather'	70.904639	0.084484
kulidos 'thirty'	80.71807532	0.059458
rounded mean	**75**	**0.07**
rounded s.e.	**3**	**0.01**

/t/	release burst intensity (dB)	closure duration (sec)	VOT (sec)
wita '(they) go'	78.72138937	0.067185	0.024701
witiŋ '(I) go'	79.25279812	0.054965	0.023348
wita '(he will) go'	78.01413482	0.067501	0.028883
ati 'elephant'	75.34197681	0.065171	0.021816
atad͡ʒa '(you) go there'	75.12741689	0.05348	0.026561
rounded mean	**77.3**	**0.062**	**0.025**
rounded s.e.	**0.9**	**0.003**	**0.001**

/g/	release burst intensity (dB)	closure duration (sec)
gigap 'crow'	73.41553357	0.036452
goiga '(they two) die'	78.51303169	0.036328
ogeptiŋ 'burned'	79.25380824	0.022552
agoitiŋ '(I won't) die'	79.00122339	0.034601
awiga '(you don't) go'	79.56931548	0.044607
rounded mean	**78**	**0.035**
rounded s.e.	**1**	**0.004**

/k/	release burst intensity	closure duration	VOT
sunukuti 'elbow'	64.80565736	0.064465	0.043876
kukusap 'lion'	79.16167232	0.079316	0.02456
druka 'tiger'	75.37072074	0.050934	0.028229
druka 'tiger'	73.90856601	0.07027	0.040638
druka 'tiger'	74.05841594	0.066611	0.043297
rounded mean	**73**	**0.066**	**0.036**
rounded s.e.	**2**	**0.005**	**0.004**

Appendix B: Comparison of durations for Phoneme X and selected stops and fricatives

Phoneme X	stop closure duration (sec)	fricative duration (sec)
t͡ɕaʈa 'four'	0.018236	0.04398
t͡ɕoda 'six'	0.016985	0.06281
rounded mean	**0.0176**	**0.053**
rounded s.e.	**0.0006**	**0.009**

/s/	stop closure duration (sec)	fricative duration (sec)
semoʔ 'tree'	-	0.178753
semaʔ 'bedbug'	-	0.16268
suruʔpi 'cockroach'	-	0.166985
soro 'free'	-	0.174091
long'	-	0.215785
rounded mean	**-**	**0.180**
rounded s.e.	**-**	**0.009**

/t/	stop closure duration (sec)	fricative duration (sec)
titi 'arm'	0.078426	-
goiʔtaŋ 'cow'	0.088522	-
ati 'elephant'	0.065171	-
atad͡ʑa '(you) go there'	0.05348	-
rounded mean	**0.0714**	-
rounded s.e.	**0.008**	-

Appendix C: Alveolar and retroflex stop formant values and burst intensity

/t/	F2 (Hz)	F3 (Hz)	Burst intensity (dB)
goiʔtaŋ 'cow'	1782.473397	2848.268642	80.7025406
taraga 'morning'	1663.094883	3079.304829	70.62363498
latabuʔ 'wildboar'	1778.450922	2653.232725	72.70296114
rounded mean	**1740**	**2900**	**75**
rounded s.e.	**40**	**100**	**3**

/t͡ʂ/	F2 (Hz)	F3 (Hz)	Burst intensity (dB)
t͡ʂaʈa 'four'	2068.16	3148.50	60.9003542
bolʈabaj 'beautiful'	1916.42	2777.18	66.98244508
bolʈaɖita 'clean'	2257.24	2831.51	76.9971721
laʈa 'jungle'	1776.05	3018.88	69.0396816
rounded mean	**2200**	**3040**	**64**
rounded s.e.	**200**	**170**	**3**

/d/	F2 (Hz)	F3 (Hz)	Burst intensity (dB)
datikwi 'breast'	1774.608839	2812.560696	77.20631947
duŋgaliʔwe 'thief'	1467.369766	2725.415749	79.12326527
rounded mean	**1600**	**2770**	**68**
rounded s.e.	**200**	**40**	**3**

/ɖ/	F2 (Hz)	F3 (Hz)	Burst intensity (dB)
ɖena 'wing'	1706.80	2930.33	76.13542844
uiɖaʔ 'air'	1956.37	2831.80	71.85099893
ɖem 'bridge'	2217.95	3174.60	68.01991893
aʔɖoŋ 'fish'	1769.56	2588.71	74.97438377
giɖap 'mouse'	1936.80	2760.33	75.26654948
rounded mean	**1920**	**2900**	**73**
rounded s.e.	**90**	**100**	**1**

Appendix D: Remo vowel formant values

/i/	F1 (Hz)	F2 (Hz)
kirime 'fingernail'	370.08442886390253	2298.175070618524
kirime 'fingernail'	349.94294046688975	2421.3667122425963
titim 'tamarind'	375.0272914	2468.271945
titim 'tamarind'	397.9883113	2427.063719
piriʔ 'bird'	298.2438044249737	2536.0139959849134
piriʔ 'bird'	303.5865875685255	2545.5232354042037
kiri 'fat, grease'	325.98356333772347	2430.4752547588455
kiri 'fat, grease'	324.3164501148754	2428.9786685149925
bimbi 'four-winged butterfly'	293.4714505	2419.818666
bimbi 'four-winged butterfly'	295.8656064	2411.32897
rounded mean	**330**	**2440**
rounded standard error	**10**	**20**

/u/	F1 (Hz)	F2 (Hz)
nkuʔ 'uncooked rice'	447.4444850966464	695.2259931087087
ntuŋ 'body lice'	349.3710896830557	1307.2241578470414
buŋte 'buffalo'	336.117780864156	946.9820332808442
gusoʔ 'dog'	329.6020653816558	1189.767089731631
suluŋ 'high'	387.3381022609346	1386.9719107053406
suluŋ 'high'	411.7717329523997	1124.1173797315491
rounded mean	**380**	**1100**
rounded standard error	**20**	**100**

/e/	F1 (Hz)	F2 (Hz)
leʔ 'bamboo shoot'	545.8600758091493	2421.0307565638536
bileʔsiŋ 'feather'	498.75212202767847	2158.0586571615036
ŋgereʔ 'boy, son'	419.1907943825548	2141.4159882519075
ŋgereʔ 'boy, son'	466.15712453667135	1936.851015760116
ɖem 'bridge'	530.506770256179	2105.586726518709
pineʔ 'flute'	445.7764199755088	2146.2716060589464
rounded mean	**480**	**2150**
rounded standard error	**20**	**60**

/a/	F1 (Hz)	F2 (Hz)
aʔa 'bamboo'	841.079603271729	1458.349203631743
aʔa 'bamboo'	873.8699327658925	1513.5022677855918
laglaʔk 'cave'	702.85537738494	1640.2881010157198
laglaʔk 'cave'	777.2395935901488	1565.5984189858284
laɽaʔ 'jungle'	678.5670587368701	1633.568396441149
laɽaʔ 'jungle'	809.945465426316	1686.741905063811
ɖaʔ 'water'	757.8391446169456	1654.506518677797
ba 'father'	729.7102879366695	1503.0785797473586

rounded mean	**770**	**1580**
rounded standard error	**20**	**30**

/o/	F1 (Hz)	F2 (Hz)
ngo̲ʔ 'crab'	454.491266212095	1029.98167910578
bolt̪ad̪ita 'clean'	554.337791126145	1112.43317302116
nto̲ŋboʔ 'brain'	623.0739270753166	994.399283864939
ntoŋbo̲ʔ 'brain'	541.5982693685925	1014.094106001591
so̲ro 'free'	566.8561604935402	1244.203381940274
soro̲ 'free'	585.5373010238833	1191.0809326706042
d͡ʑoŋ 'mother'	452.5932635523343	1307.0173372321037
rounded mean	**540**	**1130**
rounded standard error	**20**	**50**

Bibliography

Anderson, Gregory D.S. 2007. *The Munda Verb: Typological Perspectives*. Berlin: Mouton de Gruyter. 7.

Anderson, Gregory D.S. and Harrison, K. David. 2008. Remo (Bonda). In Gregory D.S. Anderson, ed. *The Munda Languages*. New York: Routledge. 557-632.

----. 2005-2011. Remo-English, English-Remo Online Talking Dictionary, v.1.0beta. Available online at < http://remo.swarthmore.edu>.

----. 2007-2011. LivingTongues.org. retrieved December 2011. <http://livingtongues.org>.

Battenburg, John D. and Swanson, Judy. 1999-2000. German Phonetic Alphabet. Available online at <ipap.calpoly.edu/gpa/index.html>.

Beaumont. 1997-2011. Freelang.net Oriya-English online dictionary. Available online at <http://www.freelang.net/online/oriya.php>.

Bell, R.C.S. 1945. *Koraput District Gazetteer*. Cuttack: Government Press.

Bhattacharya, Sudhibhushan. 1968. *A Bonda Dictionary*. Poona: Deccan College.

Boersma, Paul and Weenink, David. 1992-2011. Praat: Doing Phonetics by Computer [computer software]. v.5.3.03. Available online at <http://www.fon.hum.uva.nl/praat>.

Donegan, Patricia and Stampe, David. 2004a. Rhythm and the Synthetic Drift of Munda, in Rajendra Singh, ed. *The Yearbook of South Asian Languages and Linguistics.* vol. 7. Berlin: Mouton de Gruyter. 3-36.

----. 2004b. Munda Lexical Archive. Available online at <www.ling.hawaii.edu/austroasiatic>.

Elwin, Verrier. 1950. *Bondo Highlander*. Bombay, New York: Oxford University Press.

Fernandez, F. 1968. *A Grammatical Sketch of Remo: A Munda Language*. PhD dissertation, University of North Carolina.

Furer-Haimendorf, C. von. 1943. Avenues to Marriage among the Bondos of Orissa. *Man in India.* vol 23. New Delhi: Serials Publications. 158-172

----. 1945. Megalithic Ritual among the Gadabas and Bondos of Orissa. *Journal of the Royal Asiatic Society of Bengal*. vol. 9. Calcutta. 149-178.

Hamann, Silke. 2003. *The Phonetics and Phonology of Retroflexes*. Utrecht: LOT.

Harrison, K. David. 2007. *When Languages Die: The Extinction of the World's Languages and the Erosion of Human Knowledge*. New York: Oxford University Press.

Konow, Sten. 1906. Section on the Munda and Dravidian languages, in Grierson, G.A., ed. 1927. *Linguistic Survey of India*. Delhi: Motilal Banarsidass.

Ladefoged, Peter. 2005. *Vowels and Consonants*. Wiley-Blackwell. Supplementary course data. retrieved December 2011. <http://phonetics.ucla.edu/index.html>.

Ladefoged, Peter and Peri Bhaskararao. 1983. Non-quantal Aspects of Consonant Production: A Study of Retroflex Consonants. *Journal of Phonetics* 1985. Elsevier. 267-278.

Lewis, M. Paul, ed. 2009. *Ethnologue: Languages of the World*. 16th ed. Dallas: SIL International. retrieved December 2011. <http://www.ethnologue.com>.

May, J.A. 1873. Notes on the Bhondas of Jaypur. *The Indian Antiquary* 1873. Bombay: Popular Prakashan. 236-238.

Molony, J.C. 1911. *Census of India*. Part I. Calcutta: Superintendent Government Printing.

Nanda, Bikram N. 1994. *Contours of Continuity and Change: The Story of the Bonda Highlanders*. New Delhi: Sage Publications.

Sikeborg, Urban. 1997-1998. *Introduction to Swedish*. Stockholm: Stockholm School of Economics. retrieved December 2011. <www2.hhs.se/isa/Swedish/chap9.htm>.

Stampe, David L., ed. 1965. Recent Work in Munda Linguistics I. *International Journal of American Linguistics*. Chicago: University of Chicago Press. 31.

Zide, Norman. 1970. Munda and Non-Munda Austroasiatic Languages, in Thomas A. Sebeok, ed. *Current Trends in Linguistics*. vol. 1. Berlin: Mouton de Gruyter. 411-430.