From Transcript to "Trans-Script": Romanized Santali across Semiotic Media

Nishaant Choksi, *Indian Institute of Technology Gandhinagar, India*

ABSTRACT

Santali is an Austro-Asiatic language spoken throughout eastern India, Nepal, and Bangladesh. It is currently written in multiple scripts, including a Roman script devised by missionaries in the late nineteenth century, various Indic scripts, and an independently derived script, Ol-Chiki. Each of these script systems entails different sound-to-script relationships, especially for phones such as the word-final glottalized consonants, which are not present in the dominant Indo-European vernaculars. This article traces the historical transformations of sound-to-script relations in the various scripts of Santali and tracks in particular a Romanized Santali transcription orthography that developed as a way to mediate between different scripts. The Romanized Santali form assumed a particular importance as Santali speakers started using Santali in digital and online spaces due to software limitations. However, the differing use of variants within the script to represent sounds such as word-final glottal consonants shows that what appears to be a novel orthography is in fact a "trans-script," rhematizing the historical and ideological trajectories of the various script systems already in use in nondigital domains. The article claims that the Romanized "trans-script," though internally diverse, has been deployed to further the standardization project and cultural politics associated with the Ol-Chiki script.

Attention to the distribution of distinct graphic repertoires (or scripts) both within and across denotational linguistic codes has taken on increasing importance as scholars turn toward the study of digital communication. To account for the widespread use of "computer-mediated digraphia"

Contact Nishaant Choksi at Indian Institute of Tech Gandhinagar, Social Science, AB 3/333B Palaj, Gandhinagar, Gujarat 382355, India (nishaant.choksi@iitgn.ac.in).

This article was previously presented as part of organized panels at ICAAL-6 (Siem Reap), JASAS-29 (Kobe), and CALA-1 (Siem Reap). Thanks to Rika Yamashita, who coorganized two of these sessions, and to P. Kerim Friedman, who offered discussant comments. Thanks also to the issue editors Christina Davis and Chaise LaDousa for inviting me to contribute and to journal editor Asif Agha and the peer reviewers for their suggestions. Funding for this research was provided by Fulbright-Hays/DDRA, the America-Scandinavia Foundation, and the Japanese Society for the Promotion of Science (P16744).

Signs and Society, vol. 8, no. 1 (Winter 2020). 2326-4489/2020/0801-0004$10.00

in which participants write in one linguistic code by using the orthography associated with another one, Androutsopoulos proposes the term *trans-scripting* (2016, 291). Refitting the sociolinguistic notion of "trans-languaging," which challenges conventional models of code switching (García and Li 2014), from a "graphocentric" standpoint, trans-scripting aims to theorize the "networked writing" found in the digital sphere arising from the availability of multiple scripts to a community of language users (Androutsopoulos 2016, 290). Androutsopoulos focuses on digital communication in European languages and scripts, but this concept has been applied to other regions as well, such as in Wei and Zhu's recent study of Roman/Chinese "trans-scripting" by users on the digital WeChat platform (Wei and Zhu 2019). Yet the work both of Androutsopoulos and of Wei and Zhu is predicated on an assumption of a monographic stability between language and script, as well as the unique potential of the online space to disrupt this stability. This understanding of "trans-scripting," I suggest, does not adequately address the sociolinguistic realities in highly diverse graphic environments such as South Asia, where the isomorphism between language and script is frequently contested and multiscriptality is ubiquitous even outside the sphere of digital media (Singh 2001; LaDousa 2002; Sohoni and Brandt 2018).

This article focuses on the graphic practices of speakers of Santali, an indigenous, Austro-Asiatic minority language spoken by over 7.5 million speakers across a wide expanse covering the eastern Indian states of Orissa, Jharkhand, West Bengal, and Assam, as well as parts of Nepal and Bangladesh.[1] The study draws on a survey of print and digital sources spanning the last decade, as well as eighteen months of ethnographic fieldwork conducted in the eastern Indian state of West Bengal in 2009–11, in order to analyze the variety of graphic practice and the network of script systems in use by Santali speakers across digital and nondigital domains. For Santals, whose language has been subordinated in state policy to dominant Indo-European vernaculars, writing in Santali has been central to assertions of political and cultural autonomy. However, because written Santali operates within several overlapping multilingual and multiscriptal milieus and is not formally connected with state power, Santali has not historically been subject to institutionalized standardization. Santals continue to write their language in several scripts, including a Roman script devised by missionaries, the various Indic scripts of the regions where they reside, and an independent script created for Santali known as Ol-Chiki 'writing symbol'.

1. According to 2011 census data, available at Ethnologue https://www.ethnologue.com/language/sat/22.

The use of different scripts for a particular code does not exhaust the diversity of graphic practice. There are also graphic repertoires that cut across script systems, creating a densely layered trans-script in which the prevailing multiscriptal situation is indexically invoked without any formal mixing or substitution of graphemes. This article traces the multiple lives of one such repertoire, "Romanized" Santali, across different semiotic media, ranging from magazines, posters, and signboards to social media platforms and digital hardware. I will chart the emergence of this repertoire from the first Santali-language printing presses of the late nineteenth century to its subsequent transformations as Santali entered the digital sphere in the twenty-first century, and I will examine it in relation to the several different scripts used to write the Santali language. I focus particularly on the way that the different manifestations of Roman script encode the word-final glottalized stops (or "checked" consonants) and on how the different orthographic variants construct an "axis of differentiation" (Gal 2016) through which indexes of a phonetic feature, as well as social affiliation, come to be understood semiotically as an intrinsic feature of one of the available script systems present within the graphic ecology. Consequently, although the variants all derive from the same graphic system (Roman), they still come to serve as indexical icons of a larger multigraphic differentiation and are mobilized in a politics of autonomy in which language and script play a crucial role.

The first section discusses the development of the Santali Roman script in the early nineteenth century by Scandinavian missionaries. At the time, missionaries argued that the Roman script was better suited for Santali than the Indic scripts used to write neighboring Indo-European languages like Hindi or Bengali, since the Roman script was easily able to adapt diacritics to represent unique sounds such as the word-final glottalized (checked) consonants. This system of "transcription," originally meant to teach missionaries and non-native speakers how to properly pronounce the Santali language, eventually became a marker of "native"-ness as opposed to the Indic script systems.

The second section charts the development in the mid–twentieth century of Ol-Chiki and the contemporary movement to popularize this script. I suggest that as with the Roman script, the checked consonants played an important role in arguments for both the uniqueness and the suitability of writing the script for the Santali language. Yet as the script was still unknown to most of the population, a Roman script transliteration system was used to support the script's use and appeared in conjunction with the script in instruction primers, on signboards, and in print media. This Romanization, I argue, does not simply transliterate but also transduces the Ol-Chiki script, constructing an axis of

differentiation between Ol-Chiki, on the one end, and the missionary-derived Roman and the Indic scripts, on the other.

The last section outlines the use of Romanized orthography in Santali-language digital communication, examining both online social media platforms and the development of Ol-Chiki fonts and keyboards. Through an examination of Santali-language communication on a community-specific Facebook group, I suggest that the seemingly highly variable and nonstandard use of the Roman script to represent the glottalized consonants invokes the exposure and history of the multiple scripts used to write Santali across semiotic media. This is followed by a discussion on how this Romanized trans-script became integral to the development of fonts and keyboards used to type in Ol-Chiki. Consequently, what was originally seen by missionaries as a Romanized "transcript" to help non-native speakers properly pronounce Santali has now transformed into a baseline orthography to facilitate native speakers' digital communication in an altogether separate, independent orthography. This transformation, the article suggests, occurs through the diagrammatic construction of multiple axes of differentiation across a range of online and offline semiotic media present in the communicative environment.

Missionaries and the Artifactualization of the Checked Consonants

Since Santali speakers have lived among speakers of Indo-European languages for centuries, most Santals are bilingual or trilingual in various Indo-European languages. This long period of language contact has created conditions of convergence in phonology and syntax, leading to notable differences between Santali and other Austro-Asiatic languages spoken in Southeast Asia (Donegan and Stampe 2004). As many linguists have argued, much of the phonology is shared between Santali and Indo-European, with the exception of the presence of the word-final, postvocalic glottalized stops (see table 1 for a chart of Santal phonology). For instance, Ghosh (2008, 27), in his study of Santali, lists "native" and "borrowed" phonological subsystems, with aspiration appearing as a component of the latter and glottalization (or "checking") belonging to the former.

Linguists disagree on the phonemic status of these sounds. Neukom (2001) cautiously identifies them as phonemes, while Ghosh (2008) sees them as allophonic with voiceless stops. Osada, in his study of Mundari (1992, 28), a closely related Austro-Asiatic language with a similar pattern of word-final checked stops, argues that the stops are allophonic with their voiced counterparts, citing the optional alternations between glottal and voiced stops under certain conditions, an alternation also present in Santali. Despite the contested phonemic

Table 1. Santali Consonant Inventory (Adapted from Neukom 2001; Ghosh 2015)

	Bilabial	Alveolar	Retroflex	Palatal	Velar	Glottal
Stops (Voiceless)	p (p^{h})	t (t^{h})	ʈ (ʈ^{h})	c (c^{h})	k (k^{h})	
Stops (Voiced)	b (b^{h})	d(d^{h})	ɖ (ɖ^{h})	ɟ (ɟ^{h})	g (g^{h})	
[Stops, Glottalized]*	[p̓]	[t̓]		[c̓]	[k̓]	
Fricatives		s				h
Nasals	m	n		ɲ	ŋ	
Trill		r				
Flap			ɽ			
Lateral		l				
Glide	w			y		

* Phonemic status contested

status, in linguistic transcriptions of Santali the convention has been to represent these characters as checked voiceless stops, a convention that has been adopted by Ghosh, Neukom, and me. The contrast between the phonemic rendering of the checked stops and the phonetic rendering (Osada 1992) can be seen in table 2.

The convention of rendering these stops as separate graphemes within the linguistic literature aligns with an "artifactual ideology" (Blommaert 2008) first articulated by missionaries at the original codification of the Santali language and the emergence of a standardized Santali Roman script in the late nineteenth and early twentieth centuries. The American Baptist missionary Jeremiah Phillips produced the first known grammar of the Santali language, *An Introduction to the Santal Language*, in 1852. He chose to render the Santali examples in the Eastern Brahmi script, the script used for writing the Bengali language and the dominant script of the erstwhile colonial territory of Bengal. At this stage, Phillips made no distinction between the Indo-European phonology of Bengali and that of Santali, uncritically using the Eastern Brahmi graphemes to represent Santali sounds. However, he did note the presence of the checked consonants, writing in his phonetic summary that Santali had distinct "half-formed guttural sounds" not present in Bengali or Oriya (1852, 3).

The first clear articulation of these sounds as distinct and deserving of unique graphic representation came from the Norwegian missionaries who arrived in the Santal Parganas (in present-day Jharkhand state) in the late nineteenth century with the aim of establishing an independent and autonomous Santal mission. Missionaries such as L. O. Skrefsrud and, later, his prolific protégé Paul Olaf Bodding developed a Roman script for writing Santali, claiming it was linguistically superior to the Indic scripts used previously. The script included several

Table 2. Phonemic and Phonetic Rendering of Checked Consonant Features

Phonemic Rendering (Roman Script)	Phonetic Rendering (Osada 1992)
k̇	ʔ
ṗ	ʔb̤^{m}
ċ	ʔɟ^{m}
ṫ	ʔḍ^{m}

diacritics to isolate distinct Santali phonemes, particularly the checked consonant series, as seen as in table 2. The faithfulness of these diacritics to "correct" pronunciation was used to justify the script in relation to the other Indic scripts. "So far as the consonants are concerned," Bodding wrote in an early article on the Santali language, "any alphabet derived from the old Sanskrit or Devanagari alphabets is, with the exception of the checked consonants, much superior to our Roman alphabet" (n.d., 3).[2] Bodding claimed that the Roman alphabet, in which sounds like the checked consonants can be easily marked, could address the "uncertainty and lack of correct pronunciation, specifically with foreigners, both when reading or speaking the language" (1922, 5).

As Meek and I have argued elsewhere (Choksi and Meek 2016), the Roman script, originally devised to help "foreigners" with acquisition and fluency of the Santali language, soon came to "enfigure" Santali persons themselves.[3] Bodding goes on to say, "Santals have a mind much directed towards concrete and special subjects. To distinguish in writing between the different sounds is therefore something in accordance with their mental character. We have very little trouble in teaching them to write correctly, when they use our system. A better proof of its soundness is not needed" (1922, 5).

While in the first part of the essay, Bodding understands "correctness" as an index of the script for a "foreigner" interpretant who was meant to learn the script, he later resignifies this indexical relation as iconic of the native Santali "mental character." The interpretation of an indexical sign as iconic is what Gal, following earlier work by Irvine and Gal (2000), has called a process of "rhematization" (Gal 2016, 122).[4] Rhematization depends on the qualitative

2. Bodding's article was most likely written between 1910 and 1920, prior to his 1922 grammar.

3. By "enfigurement" we mean the "process of personifying some form in and through practice" (Choksi and Meek 2016, 229).

4. The process was originally called "iconization" in Irvine and Gal (2000) but then was reworked as "rhematization" to better reflect its Peircian origins. A *rheme* in Peircian terminology is the interpretation of a sign with respect to quality (firstness), and is part of Peirce's third trichotomy (Gal 2016, 122).

interpretation of a linguistic feature, and necessarily entails a contrast with another set of qualities, on an "axis of differentiation." According to Gal, this axis provides a "semiotic scaffolding" on which to evaluate and position iconic difference, "organized according to qualities picked out as shared by the expressive features that make up a register and also by the persona it indexes, in contrast to another pairing" (121).

In interpreting what was previously material meant to train non-native speakers into speaking Santali as inherent qualities of a cultural psyche, Bodding sets up an axis of differentiation where the qualities of "correct" and "incorrect" become aligned with a "native" and "non-native" persona, respectively. This contrastive pairing, based primarily on the interpretation of written material and the proper pronunciation of certain phonemes, contributes to the larger missionary project of delineating a bounded Santali ethnolinguistic group, distinct from neighboring tribes and caste Hindu communities (Carrin-Bouez 1986). Similar to what Irvine has shown for colonial linguistic interventions in Africa around the same period, the identification and valorization of putative phonemes (the phonemic status of which was often doubtful) and the bundling of linguistic features with person-types were critical for turning "linguistic practices into named 'languages' (supposedly) corresponding to ethnic groups" (Irvine 2008, 338).

As early as 1874, the Norwegian mission had started a printing press at its main station at Benagaria in what is now Dumka district, Jharkhand, publishing bibles, Christian hymnals, and transcriptions of Santali oral histories and folk songs, as well as a Roman script magazine. This allowed the Roman script to circulate broadly among both Christian and non-Christian Santals, and also helped connect the mission's far-flung settlements (Carrin and Tambs-Lyche 2008). Both Christian and non-Christian Santals continue to write in the Roman script, with some of the Santali language's longest-running publications, such as the Calcutta-based *Nawa Ipil* (New Star) and *Jugsirjol*, published in the script.

At the same time as the Roman script was being developed, Santals were also modifying the Indic scripts used for writing Santali. In 1890, Majhi Ramdas Tudu published *Kherwal Bongsho Dharam Puthi*, the earliest known nonmission publication in Santali, which outlined the traditional ritual practices of the Santal community (*Kherwal*) using a modified Eastern Brahmi script.[5] Tudu's work was in part a reaction to the increasing dominance of the Christian mission and converted Santals in the cultural sphere. An anti-Christian reform movement,

5. The spellings *Kherwal* and *Kherwar* are used interchangeably.

Table 3. Checked Features in Bengali Script-Santali (from *Tetre*, Published by M. Hansda Kaira, West Bengal, India)

Bengali	Roman Transliteration	Pronunciation
ঃ	ḥ (*visarga)*	ʔ
চ্	c	ʔɟ^{m}
ৎ	t	ʔḍ^{m}
প্	p	ʔḅ^{m}

which was nevertheless deeply influenced by the Christian mission (Anderson 2009), known as the Kherwar Movement, arose around the late nineteenth century and resulted in Santali language production and ethnolinguistic assertion in Indic scripts, primarily in Eastern Brahmi, but also Devanagari (used for writing Hindi) and Utkal (used for writing Oriya). The rise of the Kherwar movement established new semiotic arrangements, where non-native/native distinctions, iconized by script difference between Indic and Roman scripts, were reinterpreted to differentiate Santali persons and traditions from the cultural and religious project of the Santal mission.

While the development and modification of Santali writing in the Indic scripts occurred in opposition to the Christian mission's propagation of Roman, the artifactualization of the checked consonant stabilized as a feature of a new register of Santali language across graphic boundaries. Whereas the early representations of Santali in Indic scripts, such as the system used by Phillips, had no notation to distinguish the glottalized consonants, from the mid–twentieth century onward, Santali-language books and periodicals published in Indic scripts included modifications to indicate the checking of word-final stops. This marking system was not standardized across the entire Santali-speaking region, but many of the same principles were followed. For instance, the Santali-language print publication *Tetre*, published from Purulia district in West Bengal beginning in 1976, utilized what was by that time a largely accepted system of Bengali transcription of Santali characters (see table 3).

In this system, the phonological understanding of the checked consonants slightly differed from the Roman script devised by the missionaries. Like in Roman, the bilabial and palatals were represented by their equivalents in Bengali, although an Indic diacritic called a *virama*, which is used (in Indo-European languages) to mute the inherent vowel, was used to check the consonant (table 3, rows 2 and 4). However, the alveolar was represented with an entirely different

grapheme, the Eastern Brahmi character *khanda ta*, which is a special character used in Bengali orthography to articulate the alveolar in consonant conjuncts, although in Santali orthography, it is used strictly for the checked feature (table 3, row 3). The biggest difference between the Roman orthography and the Eastern Brahmi Santali orthography is the representation of the glottal stop /ʔ/, which in Roman is represented as a checked velar *k'*, but in the Eastern Brahmi is represented by the Indic character called a *visarga* (:) which is used for Sanskrit borrowings in Indo-European languages to impart breathiness to the preceding vowel (table 3, row 1).

The Eastern Brahmi–Santali orthography is much more ubiquitous in everyday life in the region of West Bengal, India, where I conducted my fieldwork. For instance, posters (such as the drama posters in fig. 1), magazines, newspapers, wedding invitations, and so on, all used this script, as it was the most easily understood by the vast majority of the Santali-speaking population in the region. While Santali in Eastern Brahmi is not formally taught or learned, since primary and secondary education occurred in the Bengali language, Santals could easily read out these characters, and wide exposure to the modifications, such as those of the checked consonants, resulted in easy recognition of the points in which the orthography departed from Bengali. Within local ideologies of differentiation in rural communities in southwest West Bengal, where a Santali-speaking plurality lived together with the more economically and socially dominant caste Hindu Bengali-speaking counterparts, these modifications also served to distinguish a Santal (*hoṛ* 'person') from a non-Santal (*diku*). For instance, early on in my fieldwork in a village in West Bengal, I found that my pronunciation of the checked consonants in the Eastern Brahmi renderings of Santali earned me praise from local Santali-speaking residents, who told me that unlike the average *diku*, I had more correct knowledge of the language. By the time of my fieldwork in 2011 it was already therefore apparent that the proper pronunciation of checked consonants was subject to ideological evaluation and served as a frame through which insider/outsider distinctions were negotiated.

Ol-Chiki Script and Its Romanization

In the 1930s, a Santal schoolteacher named Raghunath Murmu from the eastern Indian state of Orissa developed an independent script for Santali, which he called Ol-Chiki 'writing symbol'. The script was developed to provide a non-Roman, non-Indic alternative for writing Santali and also served as a way of unifying Santals across political and graphic boundaries. The development of

Figure 1. Santali language drama poster, Jhilimili, West Bengal (photograph taken by the author).

the script coincided with the call for the establishment of the independent, indigenous-majority state of Jharkhand, which was to be carved out of the border regions of Orissa, Bihar, and West Bengal in order to provide the various forest-dwelling indigenous groups a political voice in the soon-to-be-independent nation of India (Munda and Mullick 2003).

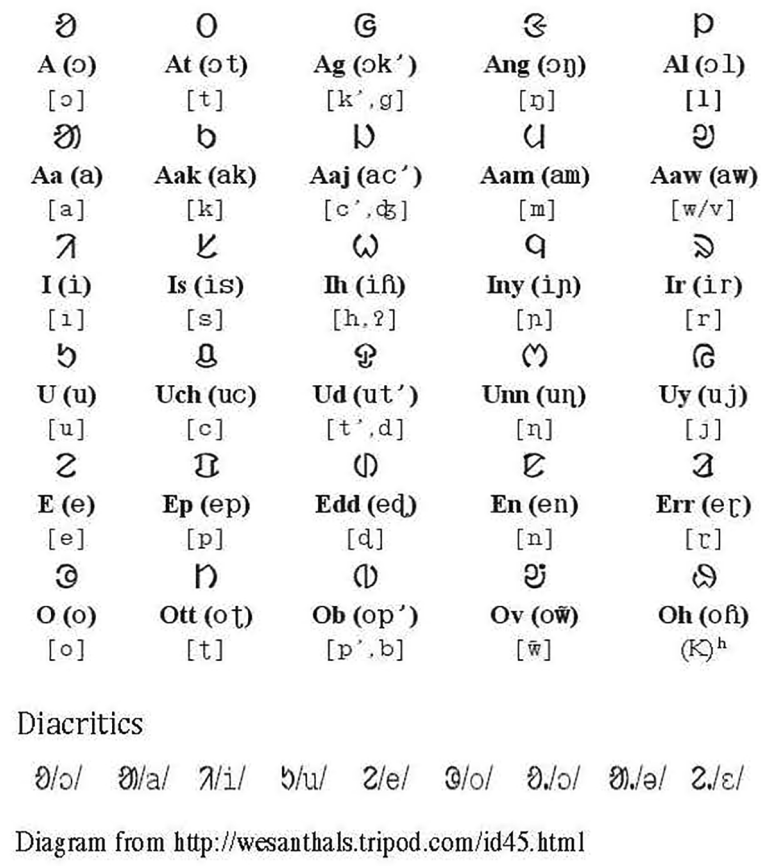

Figure 2. Ol-Chiki script, from http://wesanthals.tripod.com/id45.html

The script has 30 letters, with 25 consonants, 5 vowels, and several diacritics (see fig. 2). It is alphabetic, unlike Indic *abugidas* such as Eastern Brahmi. This meant that, instead of each letter carrying an inherent vowel, vowels and consonants are written separately, similar to the Roman script. Yet unlike the Roman script, each Ol-Chiki character, which is affixed to a particular phoneme, also iconizes a Santali lexeme. Hence, the script is pictographic as well as alphabetic (Mahapatra 1986). For instance, if one looks at the first letter in the first row in figure 3, the Santali letter *o*, which is phonetically pronounced /ɔ/, is named /lɔ/ 'fire' and thus looks like a flame. Similarly, *t*, phonetically indicating the alveolar /t/, is named /ot/ or 'Earth' and thus is round, shaped like the

Earth. This iconicity between sound, letter, and graphic shape extends throughout the whole system.

To justify the need for a new graphic system for writing Santali, Murmu (n.d.) wrote a prescriptive grammar called *Ronoṛ* 'lexicon', which would serve as a guide for future Ol-Chiki pedagogy. In the book, he explained how the Ol-Chiki script was better suited for the unique sounds of Santali than either Roman or the Indic scripts, and paid particular attention to the *taput' aṛang* 'checked consonants'. He wrote that the other scripts had misinterpreted the nature of these consonants and only Ol-Chiki represented them in the correct fashion, that is, as alternating allophonic variants between checked and voiced stops (listed in table 4). The term *taput' aṛang* would later be translated into English by Ol-Chiki advocates and in subsequent Unicode proposals as "semi-consonants" (Everson et al. 2002).[6] These consonants, when placed word finally, would be interpreted as "checked," but when appearing word medially preceding a vowel (represented in Ol-Chiki alphabetically), they would be realized by their voiced equivalent. For the Ol-Chiki rendering of the Santali word /daʔ/ 'water/rain', (ᱫᱸᱜ) is /uʔḍᵐ -a-oʔ/ (ᱫᱸᱜᱚ). The /uʔḍᵐ / (ᱫᱸ) appears word initially (before the vowel) and is therefore voiced, while the /oʔ/ (ᱜ) appearing word finally is glottalized. This contrasts with a word like /miʔḍᵐ/ 'one', (ᱩᱱᱫᱸ) where the /uʔḍᵐ/ (ᱫᱸ) is glottalized. However, if a word such as /daʔ/ appears in a sentence and is followed by a vowel, then it is pronounced as a voiced stop. This is accomplished through a deglottalization diacritic known as an *ohot* (ᱹ), which renders the word-final glottalized consonant as a voiced consonant, such as in the sentence *dag-ay* ᱫᱸᱜᱹᱭᱸᱜ 'it is raining', where /daʔ/ followed by a verbal indicative marker /a/ becomes pronounced as /dag/.

These innovations in the rendering of the checked consonants were used by Murmu and his followers to promote Ol-Chiki over what they called Ol-Urum 'dusty writing', which included the Indic and Roman scripts (Zide 1999). For instance, in an online tutorial of Ol-Chiki, the creators of the "WeSantals" group, a group founded to promote Santali language, culture, and the Ol-Chiki script, argue for the merits of Ol-Chiki in two sections that discuss the problems with Indic scripts and the problems with Roman script. In the first section, the initial reason given for the suitability of Ol-Chiki over Indic scripts is that "in Indic languages some phonetics [*sic*] like checked consonants /k', t', c', p'/ do not exist . . . there are no mechanisms to represent these unique Santali

6. This does not correspond to the linguistic use of the term *semiconsonant*, which is typically applied to sounds like glides, which could be considered either vowels or consonants.

Table 4. Checked Features in Ol-Chiki

Ol-Chiki Characters	-V / V-V Prounnunciation	V-# Pronunciation
ᱜ (ag)	g	ʔ
ᱡ (aj)	ɟ	ʔɟ^{m}
ᱢ (ud)	d	ʔḍ^{m}
ᱣ (ob)	b	ʔḅ^{m}

sounds." In the second section, they write, "Although the Roman script can nicely represent the checked consonants, it is not without deficiencies . . . it does not have any explicit mechanism to represent the glottal stop. Therefore to retain the beauty, specialty, peculiarity & sweetness of the Santali language, there is a need to use a script that can represent all sounds of the Santali language accurately, and is naturally appealing to all Santals."[7]

These statements in support of Ol-Chiki echo very closely the justification for the Roman script made by the missionary Bodding in the early twentieth century. On the one hand, the adequacy of the script is referenced through pronunciation of the "unique" phonetic features of Santali like the word final glottalized stops, not through linguistic arguments such as allophony. Although the multiple, position-based realizations of the "semi-consonant" graphemes hint at allophony, the governing ideology views the sounds as particular to Santali and casts Ol-Chiki as a more proper way of representing them than the Indic scripts. The indexical relation between sound and script is then rhematized, with the writers sweeping in a range of what Peirce has called "qualia" (Chumley and Harkness 2013; Gal 2016), including "beauty," "sweetness," or "peculiarity," and organizing them along an axis of differentiation bounded by Ol-Chiki on one end and Indic and Roman Ol-Urum scripts on the other.

Murmu was an innovator in that he devised several means to promote the script to a highly dispersed and differentiated audience. Not only was the Santali-speaking community spread in remote villages across several states; it was also marked by a wide distribution of educational attainment and familiarity with formal literacy. In addition, for those who had some level of formal schooling, it was in some combination of English and the dominant Indo-European vernaculars. In order to spread the script, Murmu yoked the learning of the script both to a religious movement called Sarna[8] that drew from the earlier Kherwar

7. See http://wesanthals.tripod.com/id45.html.

8. *Sarna* means "sacred grove" in Mundari and was said to be started by the Munda revolutionary Birsa Munda during the Munda insurrection in 1895. In 1963 it was revived as a religion by a Santal guru Hopna Besnao and coincided with the Jharkhand movement and was embraced by Raghunath Murmu and followers of the Ol-Chiki script (Carrin 2008).

movement (Orans 1965; Carrin and Tambs-Lyche 2008) and also wrote a play about two spirit lovers separated by war who used the script to communicate their "secret feelings" to one another (Lotz 2007). Similar to how Faudree describes for Mazatec language revival in Mexico (Faudree 2013), the movement to promote Ol-Chiki drew new literacy practices together with religious practices and performance genres such as song and dance to create political awareness around issues of language and autonomy, among both those with and those without exposure to formal literacy. Embedding the script in already existing affective circuits imbued the script with iconic signification beyond its association with named linguistic codes or phonemes (Choksi 2018).

In addition to writing plays and promoting the script through religion, Murmu also published a series of primers that were to be distributed across the Santali-speaking area by volunteers to help train young students in the script. These primers included an explanation of the letters of the script, the system of grapheme-phoneme correspondence for sounds such as the checked consonants, and the ideology behind the script. With the help of these primers, the script was taught in both urban and rural Santal settlements through a system of nonformal education. However, as the script gained popularity, demand for these primers increased and many who did not have access to these informal schools also started expressing interest in Ol-Chiki. Even during my fieldwork, after several years of Ol-Chiki education, Santali-language publishers told me that as opposed to books or periodicals in Ol-Chiki (which were few in comparison with the other scripts) the most popular publications in Ol-Chiki were primers. These primers often taught Ol-Chiki script through the graphic medium of Indic scripts but also transliterated the script into a nonstandardized variant of Roman. This Romanized transcription was not the Roman devised by missionaries, which required a specialized knowledge of diacritics, but a new form of Romanization meant to accompany Ol-Chiki characters for audiences still not proficient in the script.

In these primers, such as this one published in 2007 from a Santali-language publishing house in the eastern Indian metropolis of Calcutta that I picked up at a village fair for the equivalent of 20 US cents (fig. 3), one can see how Romanization functions alongside other scripts to facilitate Ol-Chiki pedagogy. On the first page of the primer, we see a list of Ol-Chiki letters, with their equivalents given in Eastern Brahmi, Devanagari, and Romanized transliteration. While for Santali-specific consonants, like the checked series, the Eastern Brahmi follows the standardized convention (using the *visarga* or the *virama*), the nonstandard Roman transliteration uses the voiced consonant without any diacritic.

Figure 3. First page of *Ol Adang*, Debdulal Murmu (Kolkata: Adim Publications, 2007)

For instance, if one looks at the third column in figure 3, one can see that the Ol-Chiki character ᱡ /oʔ/ is transliterated as "og," ᱣ /oʔʄ^{m} / is "oj," ᱫ /uʔd^{m} / is "ud," and ᱕ /oʔb^{m}/ is "ob." These renderings contrast with the Roman script used by the missionaries or the Indic script transliterations, which, through the use of diacritics, metalinguistically emphasize the checked nature of the consonants. As mentioned earlier, Murmu himself as well as subsequent Ol-Chiki advocates also mentioned the "peculiarity" of the *taput' aṛang* 'checked sounds' and Ol-Chiki's suitability in rendering these consonants, as opposed to Indic or Roman scripts. This suitability, they claim, lies in showing the alternation between voiced and glottalized stops with no diacritics, and therefore the Romanized

transliteration displays only the voiced consonant even word finally. Consequently, the distinction had to be learned and in fact formed a central component of Ol-Chiki pedagogy through which it became ideologically salient for speakers. In the weekend Ol-Chiki classes I attended in a suburb of Calcutta, started for children of Santals who had migrated to the city for work, teachers frequently corrected students in order that they would pronounce the word-final checked consonants properly when reading Ol-Chiki material.

Those who are skeptical of the Ol-Chiki project have criticized this graphic logic. For instance, an editor of a well-known Roman (missionary-derived) script magazine told me that one of the main problems with Ol-Chiki is that it is not clear on the status of the checked consonants. For instance, he said (as he wrote out the variants for me), if one is reading and writing *saohed* (literature) instead of *saohet* [*sic*], then how is one to master correct Santali pronunciation? Others have also attacked the logic invoking a certain view of "linguistic" validity. For instance, the linguist A. K. Pal (2006) writes in a Bengali-language article on the scripts of Santali that one of Ol-Chiki's major flaws is its confusion over the representation of the checked consonants. He draws from missionary sources to argue that checked consonants are phonemes in Santali and therefore must be represented as such, which is done in Eastern Brahmi and Roman, but not in Ol-Chiki. In the end he suggests that the modified Eastern Brahmi script, due to its accessibility and its representation of phonemes like the checked consonants, is the most suitable script for writing Santali.

The arguments for and against Ol-Chiki are facilitated by the continuous semiotic "transduction" (Silverstein 2003) between two semiotically distinct graphic systems, Ol-Chiki and Roman. This transduction emerges from the constant presence of Ol-Chiki alongside a Romanized transliteration within the linguistic landscape itself. For instance, during my fieldwork, most of the Ol-Chiki signs present in the marketplaces of the village where I worked also had Roman transliterations. Typical of these signs is an example in figure 4, which is advertising the Rusika Ramjham Orkestra, which is a local Santali-language touring band.

In addition to signs, almost all Santali-language media produced in Ol-Chiki that I came across also used this combination (Choksi 2017). This convention even extended to spontaneously produced writing. For instance, at the local high school, Santali students would decorate their doors of their dormitories with Ol-Chiki graffiti, and they would often use Ol-Chiki/Roman combinations. For instance, in figure 5, one can see the words *Sagun Daram* 'welcome' written

Figure 4. Poster from Jhilimili, West Bengal (photograph taken by author)

on the doors of the Class 11–12 hostel doors in Ol-Chiki and the nonstandard Roman combination.

While Roman appeared together in combination with Ol-Chiki, I never saw this particular Roman script alone with Eastern Brahmi–Santali. For instance, in the Santali-language drama posters in figure 1, or in Santali-language, Eastern Brahmi–script periodicals, Romanization did not appear. Thus, this form

Figure 5. Adjacent hostel doors with "Sagun Daram" (Welcome) written in Ol-Chiki (left) and Roman (right) scripts (photograph taken in Jhilimili, West Bengal, by author)

of Romanization was diagrammatically linked with the Ol-Chiki script, learned as part of Ol-Chiki pedagogy and circulating together with Ol-Chiki script on signs and in printed publications. The reasons for this are many. On the one hand, Ol-Chiki proficiency is still low and the Roman transliteration helps readers unfamiliar with the script. Yet this could be done in a variety of scripts, including Eastern Brahmi, which is easier to understand for readers. I argue elsewhere that the use of Roman and Ol-Chiki combinations is a matter of scale, where the dual-script, diagrammatic constellation is perceived by Santali speakers to transcend the regional connotations of Indic scripts such as Eastern Brahmi, instantiating the transborder territorial project of Ol-Chiki within the local linguistic landscape (Choksi 2015). This use follows a general perception of Roman script (and its association with the English language) in India as marking some translocal center rather than a regional periphery (LaDousa 2014; Proctor 2014). The Romanized transliteration draws on this more general ideology prevalent throughout South Asia to offer together with Ol-Chiki a Santal-specific scale that seeks to emancipate the Santali language from its position as a socially marginal and regionally delimited local variety in a caste-differentiated social milieu.

Romanization in Digital Communication

In the last decade, smart mobile phone technology has rapidly spread through India's countryside, including in the Santali-speaking areas, and with it, access to social media platforms such as Facebook and WhatsApp have become widely available. These social media platforms have become potent political forces throughout India and have contributed to shaping political campaigns, social movements and new forms of identity politics while also fueling rumors, hatred, and mob violence (Chaturvedi 2016; Neyazi et al. 2016). However, very little work in the South Asian context has been done on the semiotic affordances of "digital discourse" (Thurlow and Mroczek 2011) in relation to the highly multilingual and multigraphic milieu.

During my initial fieldwork in the Santali-speaking areas of southeast West Bengal, Internet technology was relatively rare. In fact when I settled in the village where I was to be based, I was the only one that had access to e-mail through a mobile network USB pen drive. However, when I returned to the same area in 2012, mobile towers even in remote areas had been upgraded, and cheap smart phones and data plans were made widely available. This facilitated the creation of several Santali-language social media sites, such as Facebook pages, and prompted the widespread use of WhatsApp, a Facebook-owned messaging service for everyday communication. However, as I have argued previously

(Choksi 2017), technology also facilitated a rapid rise in print media, particularly among the younger generation of Santals and particularly in Ol-Chiki script. Before computers and digital offset printing became available in local village markets, magazine editors and publishers either had to invest in a printing press, which was usually cost-prohibitive, or journey a long distance to a district town where a printing press was available. There was only one printing press in West Bengal that could print in Ol-Chiki, and therefore most Santali-language periodicals were printed in the Eastern Brahmi script. However digital offset printing and an increasing enthusiasm for the Ol-Chiki script by a younger generation of Santali writers and publishers due to changing political circumstances,[9] led to a proliferation of more recently published Santali-language magazines in Ol-Chiki script. In addition to print publications, which were available at bookstalls and at markets set up during the seasonal fairs (*pata*), many editors also create Facebook pages to promote their magazine. One such page is the "Sarjom Umul" (Shadow of the Sal tree), created by the editors of the Ol-Chiki quarterly magazine of the same name published in the Purulia district in West Bengal.

The explicit aim of the group is to "raise extreme faith among all Santals in Santali language, culture and religion and to bring real loyalty among Santals to literature and Santali script (Ol-Chiki),"[10] ideologically aligning the magazine and the group with a discourse of autonomy that accompanied the Ol-Chiki script. However, unlike the magazine, which mostly features short stories, poetry, and essays in Ol-Chiki script, the Facebook page serves as a community forum, displaying ongoing news and commentary of events or pertinent issues facing the regional Santali-speaking community. It is a publicly accessible group, in which all members are free to comment. Posts are usually in a combination of several languages, including English, Bengali, Hindi, and Santali and also in several scripts, including Roman, Eastern Brahmi, Devanagari, and Ol-Chiki. The messages on this site reflect the multilingual and multigraphic milieu in which most Santals reside.[11]

As in the built environment offline, Santali was the graphically most complex code, being represented in Roman, Eastern Brahmi, and Ol-Chiki on the

9. The changing circumstances arose as Jharkhand state was carved out of the southern districts of Bihar in 2001, excluding the demands of Adivasi communities living in adjacent areas of West Bengal and Orissa. Ol-Chiki script which had a history within the Jharkhand movement, became one way of expressing solidarity and territorial autonomy apart from the state-centric demand. For more, see Carrin (2008); Chattopadhyay (2014); and Choksi (2017).

10. See https://www.facebook.com/groups/665333906852163/about/.

11. This is found throughout South Asia (see Christina P. Davis, "Trilingual Blunders: Signboards, Social Media, and Transnational Sri Lankan Tamil Publics," in this issue)

site. At the time I collected the data in 2015, most of the Santali tokens were in Roman, followed by Eastern Brahmi. Ol-Chiki, which had not been included as an input font on Android-based mobile phones at the time, was restricted to GIFs uploaded to the site. The Romanized Santali presented an array of different combinations, particularly in relation to the checked consonants, which on first glance seem like variants but actually derive from phoneme-grapheme relations associated with the various Santali scripts.[12] Consequently, the choice of what Roman letter to use rhematizes an entire graphic repertoire, and also invokes a particular axis of differentiation present in the larger graphic ecology, both offline and online.

The following are a few examples of posts taken in 2015 from the Sarjom Umul group on Facebook. Example 1 is presented with an analysis of how the word-final glottal stops are represented:

@@Ato	se	Apanara**h**	District	rena**h**	nagam.
Village	Or	REFL-NML:INAN	District	GEN:INAN	history

'One's village or district's history'

In this post we see that the word-final glottal stops are represented with an "h," a feature present in several examples (table 5). As seen in table 4, row 1, the Eastern Brahmi convention for representing a word final glottal stop uses the *visarga* (:), which gives a breathiness to the preceding vowel. Thus, as one pronounces Santali in Eastern Brahmi, the graphic quality of "breathiness" is transduced phonetically as glottalization. Yet as one writes Santali in Roman, the baseline Bengali pronunciation, which all Santali speakers residing in West Bengal are aware of, ensures that the Romanized transduction of the *visarga* continues to retain the quality of breathiness. In this example of Romanization therefore, the quality of breathiness is transduced as a word-final "h," which then ends up being understood as an iconic index of a word-final glottal. The semiotic complexity of such an orthographic choice would be easily understood by Santali speakers residing in the region owing to shared formal education in the Bengali language, secondary knowledge of English (in which "h" is pronounced as breathy), and the prevalence of Eastern Brahmi-Santali in the linguistic landscape of Santali-speaking areas.

12. This follows a long line of studies on orthographic choice showing that even the presence of a few graphic variants can have contentious ideological consequences (Schieffelin and Doucet 1994; Jaffe 1996; Jaffe et al. 2012).

Table 5. List of Checked Features, Example 1

Script	Pronunciation
-a**h** (NML:INAN)	aʔ
-rena**h** (GEN:INAN)	rɛnaʔ

While it was common to see "h" used for the word-final glottal, one could also see the influence of the missionary-derived Roman. Example 2 is from the same thread discussing some famous Santali writers:

1956	sal	reak	30 Jun	Nathaniel Murmu	ar	Debi Soren	takinak	kurumutute
1956	Year	GEN:INAM	30 June	Nathaniel Murmu	and	Debi Soren	DUAL NOM:INAM	preparation-INST

'On June 30, 1956, by Nathaniel Murmu and Debi Soren's preparation . . .'

As shown in table 2, the missionary-derived Roman script uses "k" plus a diacritic to phonemically analyze the glottal stop as a checked velar, voiceless consonant. Hence, in digital tokens, the letter "k" has also become an acceptable variant to represent the word-final glottal, as in the tokens *reak* and *-ak* in example 2 (table 6). The poster did not bother to add diacritics, but the iconic relation with the missionary-derived Roman script is readily understood, given its contrast with "h," in example 1. In addition, the poster's knowledge of the missionary-derived Roman script is indexed by the content of the post as well, as he is referencing a Santali-language author from northern West Bengal, Nathaniel Murmu, who wrote novels in the missionary-derived Roman script.

The final example comes from a post by one of the administrators of the group, who is a younger person involved in the production of Ol-Chiki magazines. His rendering of the checked consonants followed the Ol-Chiki Romanization, which was perhaps the most common on the group page, as seen in example 3. This is a post about the annual commemoration of the 1855 Santal rebellion against the British colonial regime:

Hul	Maha	do	pata	lekate	bang	nel	kate**d**	furgal	larhai	rena**g**	mi**d**	unudu**g**	kami
Hul	Day	TOP	fair	like-ADV	NEG	see	CONV	freedom	fight	GEN:INAM	one	show	work

'Not having seen the Rebellion Day as an ordinary fair, it was an inspirational activity as part of the Freedom Struggle'

In this case the word-final glottal is represented by "g" as in *renag* (table 7, row 2), and the glottalized alveolar with "d" as in *kated* (table 7, row 1). These

Table 6. List of Checked Features, Example 2

Script	Pronunciation
-rea**k** (GEN:INAN)	rɛyaʔ
-a**k** (NML:INAN)	aʔ

Table 7. List of Checked Features, Example 3

Script	Pronunciation
-kate**d** (Converb particle)	katɛʔḍᵐ
-rena**g** (GEN:INAN)	rɛnaʔ
-mid ʻoneʼ	miʔḍᵐ
-unudug	unuɖuʔ

letters are the same as used to transliterate the names of the Ol-Chiki letters in the primer in figure 5 and also speak to the phonemic understanding underlying the Ol-Chiki representation of the checked consonants in which graphemes can have a voiced or glottalized realization depending on position. Hence the use of "g" or "d" in word-final position, pronounced as a glottal or as a glottalized alveolar stop, respectively, iconize the Ol-Chiki letters "g" ᱜ /oʔ/ and "d" ᱰ /uʔḍᵐ/, which also use one single grapheme to represent two phones. The transduction of the Ol-Chiki letters to the Roman "g" and "d" results from the widespread use of these characters in the Romanization of Ol-Chiki within the linguistic landscape and in Ol-Chiki pedagogy.[13] The use of these characters made by one of the administrators of the group in their post indexes the prevailing ideology within the group's stated mission of promoting the "Santali script (Ol-Chiki)."

The Facebook posts show how the Roman script used for digital discourse serves as a trans-script in the same way as Androutsopoulos discusses, where certain graphic choices invoke the knowledge of multiple scripts by those involved on digital platforms. In Androutsopoulos's research however, the digital space allows a relative freedom from the standard orthographic ideologies that

13. Just as 'h' is used for word-final glottal in many representations such as in example 1, 't' continues to be frequently used instead of 'd' for the word-final glottalized alveolar in digital discourse. For instance, several popular song titles transcribed on YouTube use "t" for common words such as *mit* /miʔḍᵐ/ 'one' since it may be more recognizable. See, e.g., "Mit Phuruk Handi tegin Bulen, Hiliya" (Sister-in-law, give me a little more rice beer!), by Soren Multimedia's virtual skeleton band at https://youtu.be/ilGNhjRebtc. Therefore, there continues to be an option to use 'd' or 't', but the increasing use of 'd' over 't' indicates the influence of Ol-Chiki trans-script on transcription practice.

are hegemonic in institutional spaces in European nation-states. This case reveals a more complex situation, offering both a degree of flexibility while also creating opportunities to promote emerging standard language ideologies that are embedded in the spread and socialization of scripts like Ol-Chiki. The digital socialization of Santali occurs through exposure to digital communication, as well as to print media.

This is illustrated through a small printed booklet titled *Mobile Romoj: Santali-Hindi SMS Beora* (Mobile phone fun: Santali-Hindi SMS couplets), which I picked up at a village fair in 2013, and published by the same Calcutta-based publishing house that produced the primer in figure 3. The book comprises several romantic couplets written in Hindi and Santali.[14] Though this is a print publication, the couplets are intended to be text-messaged (SMS) on mobile phones, and the content is specifically targeted toward the secondary school or college generation, who are comfortable with mobile phones and use them to communicate with potential romantic partners. As the publication is meant for digital communication, which occurs predominately in the Roman script, couplets are written in Roman along with an Eastern Brahmi transliteration.

Figure 6 presents a sample of one of the couplet sets (made up of six couplets divided into three lines) written in Santali, which translates as:

> My mind was searching for [your] mind,
> my soul was crying out for [your] soul
> I was crying out for a lover's love

The list of checked features is presented in table 8. As can be seen in rows 1 and 2, the word-final glottalized alveolar in the word /ləgiʔḍ^{m}/ is represented as a [d], and the word-final glottal stop in /raʔ/ is represented as a [g], which follows the Ol-Chiki convention of using voiced consonants in the word-final position to represent glottalized pronunciations. However, apart from the pronunciation itself, these choices are visually intertextual with the established digital convention, illustrated in the Facebook post in example 3, of writing a Romanized Santali in digital communication. These conventions allow speaker-authors to transduce the Ol-Chiki script, thereby creating a trans-script that ideologically aligns with the Ol-Chiki project, and differentiates itself from other scripts, including Roman. This trans-script is supported by the presence of Roman alongside Ol-Chiki in the nondigital linguistic landscape. Even though there

14. Though most in rural West Bengal have limited proficiency in spoken Hindi, many young people used Hindustani (Hindi-Urdu) to text romantic couplets (*shayari*) on their mobiles, in addition to Bengali or Santali.

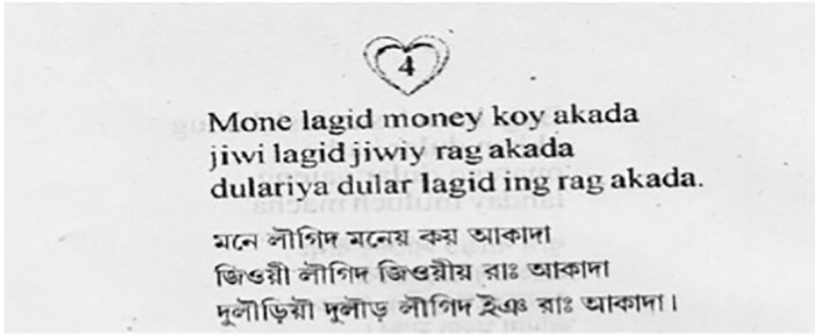

Figure 6. Page 2 of Disom Soren, *Mobile Romoj: Santali-Hindi SMS Beora* (Kolkata: Adim Publications, n.d.).

Table 8. List of Checked Features, Figure 6

Script	Pronunciation
lagid 'for'	ləgiʔḍᵐ
rag 'cry'	raʔ

is actually no Ol-Chiki present in the book, the chapbook serves as an example of graphic socialization in Ol-Chiki script.

Romanization and the Development of Santali-Language Fonts

The circulation of a Romanized Santali trans-script that furthers the Ol-Chiki project, I suggest, leads to a naturalization of Ol-Chiki within the digital space. However, the correspondences between Roman voiced consonants with the glottalized word-final consonants in Ol-Chiki not only are created through institutional means, such as Ol-Chiki pedagogy, or through the circulation and socialization of the orthography in digital spaces but also have become integral to the development of Ol-Chiki fonts for word processing and, more recently, as part of mobile phone keyboards, allowing for digital communication on media like WhatsApp or Facebook in Ol-Chiki script.

One of the reasons for the importance of Romanization is the linguistic trajectory of the development of information technology in India. Unlike in many countries, where hardware (such as keyboards) and software were developed in dominant national languages early on, India retained an English-language setup. Most keyboards on desktop or laptop computers do not have Indic characters alongside Roman ones, and software continues to be English language–based.

Keniston notes that despite the government and information technology companies' best efforts to provide local language digital resources, technology "widens the gap between who now both have power and English and the nineteen out of twenty Indians who don't. No one planned it this way, but the dominance of English as a computing language helps perpetuate existing iniquities in South Asia" (2014, 294).

"Localization," as Keniston has called it has for the most part failed in software creation in South Asia due to the association of technology with English among non-English-speaking populations. Consequently, even to use a Devanagari font, the script used to write Hindi, one of the world's most spoken languages and an official language of the Indian union, requires mediation through English-language interfaces and a Roman script–based keyboard, all of which "presupposes an advanced level of English" that most of the population does not possess (292). Yet I would suggest that in his analysis, Keniston confuses linguistic and graphic awareness. Even for a marginalized language like Santali, Roman characters have been a critical part of the development of Santali-language orthographies and have been learned by many literate Santali speakers through exposure and informal, noninstitutional means. In addition, the ability to operate computers and work with Ol-Chiki, while presupposing formal education, does not presume fluency with spoken or written English.

The development of the Ol-Chiki font, which, as Keniston argues for all Indian languages, is based on Roman script infrastructure, draws not only on English but also on the established conventions of Romanization that had been present at the outset of Ol-Chiki's development. Consequently, the ASCII fonts developed by R.C. Hansdah, a Santali-speaking computer science professor at the Indian Institute of Science, encoded the "semiconsonants" using the voiced Roman counterparts /g, d, j, b/ on the English language–based Roman script keyboard.[15] In their Unicode proposal, Hansdah and colleagues (2002, 10) describe the letters as "g, j, d, b" following the general Ol-Chiki convention, though they make reference to the missionary-derived Roman variants (k', c', t', p') in order to gloss the letters for a non-Santali-speaking audience.[16]

Previously as desktop or laptop computers were expensive to buy and computer literacy was quite low, the typing of Ol-Chiki was relegated to a few digital offset presses located in the market towns in the Santali-speaking areas.

15. These are fonts such as Ol Chiki classic, Ol Chiki royal, and Ol-Chiki *usara* 'handwriting', available at http://wesanthals.tripod.com/id19.html.

16. Ol-Chiki was accepted into Unicode in April 2008 (Baums 2016, 800).

Even then, however, the introduction of these fonts led to a substantial increase in the number of published material in Ol-Chiki, since it obviated the need to use printing presses. When computers were made more widely available in rural areas, Santals who were already acquainted with Ol-Chiki through the pedagogy could quickly transfer the skills to the Roman-based keyboard. For instance, my research assistant, who was a high school student who had studied Ol-Chiki at the local high school after Santali was first introduced as an optional subject in 2007, would, in his transcriptions of Santali spoken speech, use the voiced Roman consonants to transcribe the word-final glottals. This occurred even though he had very little exposure to computers or Ol-Chiki fonts before working with me. Typing Santali on a Roman keyboard then serves as a type of "embodied" sociolinguistic practice (Bucholtz and Hall 2016), further reinforcing the axis of differentiation in which Ol-Chiki, and its Romanized "trans-script," is considered as natural vis-à-vis spoken language, while other scripts are considered as somehow defective.

In 2017, Google announced that GBoard, the multilingual keyboard application for Android (the dominant platform used on mobile smart phones in India), added "22 Indic languages," including "Santali (Ol-Chiki and Latin)."[17] This has allowed for both Ol-Chiki and the missionary-derived Roman script (complete with diacritics) to be used as input fonts for social media platforms such as Facebook and text messaging applications such as WhatsApp. The ability to easily input Ol-Chiki, along with increasing proficiency in the script due to its inclusion as an optional subject in select schools in West Bengal, has led to what appears to be a slight decrease in the use of the Romanized trans-script on digital platforms. Transduction between scripts has therefore become integrated into the material infrastructure of the Romanized phonetic keyboard, and the expansion of this into digital applications such as GBoard ensures that the trans-script will continue to play a role in the production of Ol-Chiki, even if it is absent from the visual representation. Moreover, in providing Ol-Chiki and Latin (Roman) options, Google has responded to the ongoing conflicts between proponents of each script as to which should be used for writing Santali. The axis of differentiation, and the markers of that axis such as the word-final glottalized consonants, are artifactualized, on the layouts of the contrasting keyboards, and are now accessible to the entire population of smartphone-using Santali speakers.

17. See https://blog.google/products/search/gboard-android-gets-new-languages-and-tools/.

Conclusion

In this article, I trace the emergence and development of the Roman script among Santali speakers located in a highly diverse graphic milieu, in which several different scripts are simultaneously used to write the Santali language. While the missionaries developed a standard Roman script to represent unique Santali sounds, I show, through an examination of word-final glottalized consonants, how the Roman script offered different possibilities for phonetic realization that, though referencing the same set of sounds, could also reference different graphic repertoires. I argue this by focusing on the semiotic process of rhematization, in which indexical relationships between script and sound come to be seen as an iconic (or naturalized) resemblance between different scripts, and also between scripts and social personae. Hence, within the Roman script itself, the possible variants used to represent the checked consonants create an axis of differentiation that iconizes the several graphic repertoires present within the graphic ecology and their associated political projects and social affiliations.

While such trans-scripts have been discussed in relation to digital communication, which is seen as a privileged space for graphic innovation, I suggest that the development of Santali orthographies relied on processes of semiotic transduction that occurs both offline and online. I show this through a discussion of Ol-Chiki, an independently developed script for the Santali language, in which a type of Romanization, distinct from the missionary-derived Roman script, emerged alongside the Ol-Chiki script in print publications and signboards. This form of transduction was socialized in such a way that the trans-script was readily available once Santals started using Santali in digital domains, which reflected their commitment to the Ol-Chiki script without having to use the script itself. This was accomplished through deploying subtle contrasts in specific, ideologically weighted phonetic features, such as the word-final glottalized consonants. I then suggest another rhematized process of transformation whereby the trans-script becomes materially instantiated through the development of computer fonts to type Ol-Chiki. While the linguistic base of information-technology innovation in India and South Asia more broadly has been English, I argue that the long process of Romanization associated with Ol-Chiki made it readily available for transference to English-language keyboards, such that unique sounds could easily be physically transduced to Roman letters. The process once again created a situation whereby the specific sound-to-script correspondences between Ol-Chiki and the Romanized variety become naturalized through material instantiations on keyboards of phones or computers.

The process from "transcript" to "trans-script" discussed in this article mirrors a wider theoretical move in linguistic anthropology away from the study of script as a bounded entity or literacy as a specific, or specialized text-based practice to the study of what Debenport and Webster (2019) have recently called "graphic pluralism." Unlike theories of script in which "certain sign-type relationships have been naturalized" (391), the Romanized trans-script does not serve to either hegemonize indigenous language scripts or disrupt stable monolingual or monographic ideologies. Instead, Romanized repertoire indexically points to the continued presence of multiple scripts within the landscape, while also supporting the spread and legitimation of the newly created Ol-Chiki script. This complementary use of Roman with indigenous language scripts is prevalent in other indigenous language cases as well, such as in Bender's study of script use among the Eastern Cherokee in the United States (Bender 2002) or Daveluy and Ferguson's account (2009) of the graphic landscape in the Canadian Arctic. Moreover, Romanized trans-scripts can even invoke inscriptive practices beyond script, as suggested by Salomon and Nino-Murcia (2011) in their ethnography of literacy in an Andean village in Peru, where they demonstrate continuities between contemporary Romanized record-keeping and the inscriptional structure of pre-Colombian *khipu* knots.

The movement from "transcript" to "trans-script" evidenced in the Santali case, but also applicable to other cases in South Asia and beyond, is an example of what Bakhtin has called the decentralizing "centrifugal" (1981, 272) forces embedded in linguistic practice. The sociolinguistic realities present in places such as eastern India have ensured that scripts and languages remain in variable alignment, semiotically linked through heterographic repertoires such as the Romanized trans-script discussed here. However, as Bakhtin notes, centrifugal forces are always accompanied by centralizing "centripetal" (271) processes such as standardization and institutionalization, which may serve to eventually obscure the trans-scripts from the visible graphic landscape. As Santali in the Ol-Chiki script has now been introduced in several schools and colleges, and input tools in Ol-Chiki are now easily available for mobile phones, the Romanized trans-script already has become less visible in the graphic landscape. However, speaker-writers, both online and offline, continue to use multiple scripts on a daily basis both for Santali as well as other languages in their repertoire. As scripts like Ol-Chiki become increasingly subject to standardization, new forms of orthographic innovation may evolve that continue to draw on diverse graphic repertoires in order to articulate changing ideas of autonomy and political affiliation within the larger language community.

References

Andersen, Peter B. 2009. "Revival, Syncretism, and the Anti-colonial Discourse in the Kherwar Movement, 1871–1910." In *India and the Indianness of Christianity: Essays on Understanding—Historical, Theological, and Bibliographical—in Honor of Robert Eric Frykenberg*, edited by Richard Fox Young, 127–43. Cambridge: Eerdmans.

Androutsopoulos, Jannis. 2016. "Theorizing Media, Mediation, and Mediatization." In *Sociolinguistics: Theoretical Debates*, edited by Nikolas Coupland, 282–302. New York: Cambridge University Press.

Bakhtin, Mikhail. 1981. "Discourse in the Novel." In *The Dialogic Imagination: Four Essays*, translated by Caryl Emerson and Michael Holquist and edited by Michael Holquist, 259–422. Austin: University of Texas Press.

Baums, Stefan. 2016. "Recent Script Related Research." In *The Languages and Linguistics of South Asia: A Comprehensive Guide*, edited by Hans Henrich Hock and Elena Bashir, 798–802. Berlin: Walter de Gruyter Mouton.

Bender, Margaret Clelland. 2002. *Signs of Cherokee Culture: Sequoyah's Syllabary in Eastern Cherokee Life*. Chapel Hill: University of North Carolina Press.

Blommaert, Jan. 2008. "Artefactual Ideologies and the Textual Production of African Languages." *Language & Communication* 28 (4): 291–307.

Bodding, P.O. 1922. *Materials for a Santali Grammar*, vol. 1, *Mostly Phonetic*. Dumka: Santal Mission of the Northern Churches.

———. n.d. "On the Santal Language." Article in the Santalia archive, Nasjonalbiblioteket, Oslo.

Brandt, Carmen, and Pushkar Sohoni. 2018. "Script and Identity—the Politics of Writing in South Asia: An Introduction." *South Asian History and Culture* 9 (1): 1–15.

Bucholtz, Mary, and Kira Hall. 2016. "Embodied Sociolinguistics." In *Sociolinguistics: Theoretical Debates*, edited by Nikolas Coupland, 173–99. New York: Cambridge University Press.

Carrin, Marine. 2008. "Santal Religious Discourse and the Assertion of Adivasi Identity." In *Contemporary Society: Tribal Studies; Professor Satya Narayana Ratha Felicitation Volumes*, edited by Georg Pfeffer and Deepak Kumar Behera, 24–37. New Delhi: Concept.

Carrin, Marine, and Harald Tambs-Lyche. 2008. *An Encounter of Peripheries: Santals, Missionaries, and Their Changing Worlds: 1867–1900*. New Delhi: Manohar.

Carrin-Bouez, Marine. 1986. "De la langue au discours : Une dialectique du repli et de la modernisation dans une minorité tribale de l'Inde." *Langage et Société* 35 (1): 67–91.

Chattopadhyay, Pradip. 2014. *Redefining Tribal Identity: The Changing Identity of the Santhals of South-West Bengal*. New Delhi: Primus Books.

Chaturvedi, Swati. 2016. *I Am a Troll: Inside the Secret World of the BJP's Digital Army*. New Delhi: Juggernaut Books.

Choksi, Nishaant. 2015. "Surface Politics: Scaling Multiscriptality in an Indian Village Market." *Journal of Linguistic Anthropology* 25 (1): 1–24.

———. 2017. "From Language to Script: Graphic Practice and the Politics of Authority in Santali-language Print Media, Eastern India." *Modern Asian Studies* 51 (5): 1519–60.

———. 2018. "Script as Constellation among Munda Speakers: The Case of Santali." *South Asian History and Culture* 9 (1): 92–115.

Choksi, Nishaant, and Barbra Meek. 2016. "Theorizing Salience: Orthographic Practice and the Enfigurement of Minority Languages." In *Awareness and Control in Sociolinguistic Research*, edited by Anna Babel, 228–52. New York: Cambridge University Press.

Chumley, Lily Hope, and Nicholas Harkness. 2013. "Introduction: QUALIA." *Anthropological Theory* 13 (1–2): 3–11.

Daveluy, Michelle, and Jenanne Ferguson. 2009. "Scripted Urbanity in the Canadian North." *Journal of Linguistic Anthropology* 19 (1): 78–100.

Debenport, Erin, and Anthony K. Webster. 2019. "From Literacy/Literacies to Graphic Pluralism and Inscriptive Practices." *Annual Review of Anthropology* 48:309–404.

Donegan, Patricia, and David Stampe. 2004. "Rhythm and the Synthetic Drift of Munda." *The Yearbook of South Asian Languages and Linguistics* 7:3–36.

Everson, Michael, R. C. Hansdah, and N. C. Murmu. 2002. "Revised Proposal for Encoding the Ol Chiki Script in the UCS." In *Universal Multiple Octet Coded Character Set*. International Organization for Standardization. http://std.dkuug.dk/jtc1/sc2/wg2/docs/n2505.pdf.

Faudree, Paja. 2013. *Singing for the Dead: The Politics of Indigenous Revival in Mexico*. Durham, NC: Duke University Press.

Gal, Susan. 2016. "Sociolinguistic Differentiation." In *Sociolinguistics: Theoretical Debates*, edited by Nikolas Coupland, 113–38. New York: Cambridge University Press.

García, Ofelia, and Li Wei, eds. 2014. *Translanguaging: Language, Bilingualism and Education*. London: Palgrave Macmillan.

Ghosh, Arun. 2008. "Santali." In *The Munda Languages*, edited by Gregory D. S. Anderson, 11–98. London: Routledge.

Irvine, Judith T. 2008. "Subjected Words: African Linguistics and the Colonial Encounter." *Language & Communication* 28 (4): 323–43.

Irvine, Judith T., and Susan Gal. 2000. "Language Ideology and Linguistic Differentiation." In *Regimes of Language: Ideologies, Polities, and Identities*, edited by Paul Kroskrity, 35–83. Santa Fe, NM: School of American Research Press.

Jaffe, Alexandra. 1996. "The Second Annual Corsican Spelling Contest: Orthography and Ideology." *American Ethnologist* 23 (4): 816–35.

Jaffe, Alexandra, Jannis Androutsopoulos, Mark Sebba, and Sally Johnson, eds. 2012. *Orthography as Social Action: Scripts, Spelling, Identity and Power*. Berlin: Walter de Gruyter.

Keniston, Kenneth. 2014. "Language, Power, and Software." In *Culture, Technology, Communication: Towards an Intercultural Global Village*, edited by Charles Ess, 283–324. Albany, NY: SUNY Press.

LaDousa, Chaise. 2002. "Advertising in the Periphery: Languages and Schools in a North Indian City." *Language in Society* 31 (2): 213–42.

———. 2014. *Hindi Is Our Ground, English Is Our Sky: Education, Language and Social Class in Contemporary India*. Oxford: Berghahn.

Li, Wei, and Zhu Hua. 2019. "Tranßcripting: Playful Subversion with Chinese Characters." *International Journal of Multilingualism* 16 (2): 145–61.

Lotz, Barbara. 2007. "Casting a Glorious Past: Loss and Recovery of the Ol Script." In *Time in India: Concepts and Practices*, edited by Angelika Malinar, 235–63. New Delhi: Manohar.

Mahapatra, Sitakant. 1986. *Modernization and Ritual: Identity and Change in Santal Society*. New Delhi: Oxford University Press.

Munda, Ramadayal, and S. Bosu Mullick, eds. 2003. *The Jharkhand Movement: Indigenous Peoples' Struggle for Autonomy in India*. Copenhagen: International Work Group for Indigenous Affairs.

Murmu, Raghunath. n.d. *Ronod'* [Grammar]. Jhargram: Marsal Bamber.

Neukom, Lukas. 2001. *Santali*. Munich: Lincom Europa.

Neyazi, Taberez Ahmed, Anup Kumar, and Holli A. Semetko. 2016. "Campaigns, Digital Media, and Mobilization in India." *International Journal of Press/Politics* 21 (3): 398–416.

Orans, Martin. 1965. *The Santal: A Tribe in Search of a Great Tradition*. Detroit: Wayne State University Press.

Osada, Toshiki. 1992. *A Reference Grammar of Mundari*. Tokyo: ILCAA.

Pal, Animeshkanti. 2006. "Shaontali Bhashar Jonno Lipi Nirbachan" [Selection of script for Santali language]. *Lokshruti*, December, 1–20.

Phillips, Jeremiah. 1852. *An Introduction to the Sántál Language; Consisting of a Grammar, Reading Lessons, and a Vocabulary*. Calcutta: Calcutta Schoolbook Society.

Proctor, Lavanya Murali. 2014. "English and Globalization in India: The Fractal Nature of Discourse." *Journal of Linguistic Anthropology* 24 (3): 294–314.

Salomon, Frank, and Mercedes Niño-Murcia. 2011. *The Lettered Mountain: A Peruvian Village's Way with Writing*. Durham, NC: Duke University Press.

Schieffelin, Bambi B., and Rachelle Charlier Doucet. 1994. "The 'Real' Haitian Creole: Ideology, Metalinguistics, and Orthographic Choice." *American Ethnologist* 21 (1): 176–200.

Silverstein, Michael. 2003. "Translation, Transduction, Transformation: Skating 'Glossando' on Thin Semiotic Ice." In *Translating Cultures: Perspectives on Translation and Anthropology*, edited by Paula G. Rubel and Abraham Rosman, 75–107. Oxford: Berg.

Singh, Udaya Narayana. 2001. "Multiscriptality in South Asia and Language Development." *International Journal of the Sociology of Language*, no. 150.

Thurlow, Crispin, and Kristine Mroczek, eds. 2011. *Digital Discourse: Language in the New Media*. New York: Oxford University Press.

Zide, Norman. 1999. "Three Munda Scripts." *Linguistics of the Tibeto-Burman Area* 22 (2): 199–232.