Hannah Carlan*

“In the mouth of an aborigine”: language ideologies and logics of racialization in the Linguistic Survey of India

https://doi.org/10.1515/ijsl-2018-0016

Abstract: The Linguistic Survey of India (LSI), edited and compiled by George Abraham Grierson, was the first systematic effort by the British colonial government to document the spoken languages and dialects of India. While Grierson advocated an approach to philology that dismissed the affinity of language to race, the LSI mobilizes a complex, intertextual set of racializing discourses that form the ideological ground upon which representations of language were constructed and naturalized. I analyze a sub-set of the LSI’s volumes in order to demonstrate how Grierson’s linguistic descriptions and categorizations racialize minority languages and their speakers as corrupt, impure, and uncivilized. I highlight how semiotic processes in the text construct speakers as possessing essential “ethnic” characteristics that are seen as indexical of naturalized linguistic differences. I argue that metapragmatic statements within descriptions of languages and dialects are made possible by ethnological discourses that ultimately reinforce an indexical relationship between language and race. This analysis of the survey sheds light on the centrality of language in colonial constructions of social difference in India, as well as the continued importance of language as a tool for legitimating claims for political recognition in postcolonial India.

Keywords: language surveys, language ideologies, racialization, British colonialism, India

1 Introduction

The Linguistic Survey of India (1903–1928), hereafter the LSI, was the first attempt to capture the spoken languages and dialects of British India as part of the wider goal of the colonial government to learn, through bureaucratic documentation, about the social makeup of its subjects. The enormous compendium consists of upwards of 10,000 pages of sketches of the grammatical, phonological, and

*Corresponding author: Hannah Carlan, University of California, Los Angeles, CA, USA, 
E-mail: hannahcarlan@ucla.edu
orthographic features of a majority of the vernacular languages of India, edited and compiled over three decades by Irish ethnologist and philologist, George Abraham Grierson. Grierson espoused novel ideas about the political implications of a linguistic survey, particularly regarding the need to move away from methods that directly linked ethnology with philology, stating in the introductory volume that “questions of racial origin” cannot be adequately addressed using philological data (Grierson 1927: 29). In a parallel movement to Franz Boas in the same historical moment, Grierson aimed to establish a more scientific method of linguistic analysis based on grammar, as distinct from the study of racial and cultural difference. Boas begins his Handbook of American Indian Languages (1911) much in the same way as Grierson begins the LSI, with an exposition on the independence of language from “culture” and “physical type” (Boas 1911: 8–9), noting examples of groups whose languages have changed while their other features have remained stable, and vice versa. Where Boas offers the case of “North American negroes [...] of African descent” who are “essentially European” by “culture and language” (1911: 8), Grierson recounts the example of the Brahui speakers of Baluchistan, who speak a Dravidian language, despite having “no physical characteristics entitling ethnologists to class them as members of the Dravidian race of India” (1927: 93). Grierson thus, like Boas, attempted to break from established norms of colonial ethnology, dominated by evolutionary theories of racial progress, as well as philology, long concerned with tracing the genealogical origins of Indian populations vis-à-vis Europeans.¹ Rather than focus on the ways in which Grierson inevitably reinforced a well-established pattern of cultural evolutionism in British colonial knowledge about India, I instead want to draw attention to the ways in which Grierson’s attempt to elevate the scientific study of modern Indian languages was inseparable from existing representations of Indian society.

Despite Grierson’s renunciation of earlier forms of overtly racialist ethnological and philological forms of knowledge, the text is steeped in semiotic processes of erasure and racialization that make it difficult to disentangle linguistic data from ethnological theories about racialized human difference. This occurs through an intertextual incorporation of ethnological data from

¹ While there is no historical evidence that Grierson read Boas’ work or vice versa, the intellectual orientations of Grierson in the introductory volume of the LSI are strikingly similar to Boas’ early writings in the Handbook. That said, while Grierson’s introductory volume to the LSI shares many of Boas’ concerns with moving beyond ranking languages in an evolutionary hierarchy, as well as distinguishing between race, language, and culture as independent phenomena, the rest of my analysis will show how Grierson ultimately relied heavily on the assumptions that Boas worked to dismantle in his Handbook.
various sources, which function as evidence for classifying and distinguishing racialized groups’ forms of speech. Thus, the text contains a complex and contradictory amalgamation of quite progressive stances on the inability to infer “racial origins” from linguistic knowledge alongside a reliance on established evolutionary paradigms that were based on fundamental assumptions about (1) the affinity of language with race and ethnicity and (2) the ability to categorize, rank, and enumerate populations (and their languages) as distinct entities in an evolutionary hierarchy (Gal and Irvine 1995). Below, I examine how ideologies about language in the LSI are intimately tied to racializing discourses that construct speakers of minority, especially “tribal”, languages as corrupt, impure, and uncivilized. Rather than drawing conclusions about Grierson as an individual whose good intentions were constrained by the scholarly milieu of his time, I instead analyze the complex logics of the text itself, and its lasting implications in struggles for political recognition in postcolonial India.

### 1.1 Language and the production of racial difference in colonial knowledge

The LSI, much like the colonial censuses and gazetteers of its time, was a technique of knowledge production whereby scholar-administrators, missionaries, and civil servants recorded, and thereby created, knowledge about colonized peoples that could then be used to facilitate institutional control over them (Cohn 1981: 231; Foucault 1994). The goals and methods of the LSI were inextricable from the larger project of the newly created decennial census. It was not only designed as a supplement to population data but was itself a language census that provided estimates for numbers of speakers, as well as maps of the distribution of languages in space. Documentation of India’s linguistic geography was included beginning with the third decennial census of 1901, in the form of a report on the initial results of the LSI by Grierson (Grierson 1903a). Each subsequent census drew heavily from the LSI, with further linguistic data

---

2 I use the term “tribal” throughout the article as a reflection of Grierson’s own terminology, rather than an acceptance of the category and its application to specific groups. Grierson’s selective usage of “tribal” to describe certain ethnic groups is part of the larger process whereby a hierarchy was reified not only between Europeans and Indians, but between Indians themselves, on the basis of the characteristics of their languages. The LSI thus supplemented the ongoing discursive production of indigeneity as a colonial category of identity by asserting its connection to language.
incorporated into the censuses of 1961 and 1971 (Nigam 1972), as well as several state-specific censuses carried out after 1984. There has thus been an intimate connection between linguistic research and colonial and post-colonial governance in India, often seen as a positive and “fruitful” partnership (Bright 1968; Nigam 1972).

My analysis of the LSI highlights three processes of racialization through which languages are interpreted and classified: the mobilization of colonial tropes of India’s “martial” races (see Section 4.1), the positioning of grammars in a hierarchy of development as more or less capable of “abstract” thought (see Section 4.2), and the application of colonial ethnographic representations of “mixed” and “impure” racial groups to understandings of language (see Section 4.3). I argue that the metapragmatic and language ideological statements contained within the linguistic descriptions are made possible by the ethnological data embedded in the text. In the words of Frantz Fanon, these racialized speakers are “overdetermined from the outside” (2008 [1952]: 95), such that ethnological data produced by colonial ethnographers becomes the ideological ground upon which representations of their languages are built. In the context of Algeria, Fanon shows that the “white gaze” renders the African “a new type of man, a new species” (Fanon 2008 [1952]: 95), much in the same way that tribal forms of speech in the LSI are rendered as not-fully-languages, but as the corrupted, impure jargons of primitive others (Rosa 2016). Thus, even while Grierson professed an understanding of language as distinct from both culture and race in a parallel move to Boas in the North American academy, the text, by relying heavily on established evolutionary paradigms, inadvertently helped to reinforce an essentializing relationship between language and race in colonial India.

The “cultural process whereby linguistic units come to be linked to social units” (Gal and Irvine 1995: 970), also termed linguistic differentiation, is a central way in which colonial powers exacted a technique of divide-and-rule (Irvine and Gal 2000). Linguistic differentiation relies on semiotic processes through which ways of speaking became indexically linked to distinct forms of personhood, such that metalinguistic and metapragmatic statements about language could be seen as objective forms of knowledge about populations (Kroskrity 2000; Irvine and Gal 2000; Agha 2005). British colonial knowledge worked not only to racialize colonized groups, but also to nativize them along lines of supposedly primordial “ethnic” difference - to mark populations as not only other to Europeans but also internally incommensurable with one another (Pierre 2012). Such ideologies about language are intimately tied to the discursive and semiotic process of racialization - a set of sociohistorically specific interpretive practices that continuously make and re-make race and its
boundaries (Omi and Winant 2014; Pierre 2012; Chun and Lo 2016; Rosa 2016; Alim et al. 2016; Shankar 2015). Understanding racialization requires attention to the hermeneutic strategies deployed in colonial texts, which, as Irvine and Gal have argued, were based on the Darwinian premise that “languages were natural objects, consequences of spiritual and even biological differences between populations” (Gal and Irvine 1995: 968). This objectification of language as natural object in colonial discourse resulted in essentialized portraits of “weak” and “strong” languages, which were further seen as iconic of speakers’ intellectual capacities and features of their personhood (Errington 2001: 34; Erington 2007; Said 1979). Racialization constitutes the makings of what Silverstein has described as “indexical orders”, a process whereby metapragmatic discourses rooted in language ideologies are naturalized and become the semiotic ground upon which successive orders of indexical value are built (Silverstein 2003; Inoue 2004).

Even as Grierson rejected attempts to uncover ancient ethnological relationships through the reconstruction of linguistic genealogies as in the tradition of William Jones (Trautmann 2006: 15), his work was nonetheless steeped in the logics of modernity that explicitly placed Indian society at an earlier stage of development (Chakrabarty 2000). The LSI, with its foundation in the methods and logics of comparative philology, applied the teleological notion of progress to a theory of stages in the “development” of Indian languages from crude to civilized (Bauman and Briggs 2003). These processes effectively worked to racialize Indians as not only not-European, but also positioned within an internal hierarchy of social and ethnological difference. Racialization is not a singular or inevitable project, but necessarily entails competing ideas of where the boundaries between populations lie (Omi and Winant 2014). As such, contradiction and complexity is inherent to all racial projects, and in the case of British colonial knowledge about India and the LSI in particular, Grierson’s competing desires to elevate the study of Indian languages even as he dismissed many of them as primitive are evidence of the heterogeneity of the colonial project.

2 Crafting the survey

Trained in mathematics and Indian languages at Trinity College, Dublin, Grierson first gained the idea for the survey from his mentor Robert Atkinson, before he went on to complete nearly thirty years in the Indian Civil Service. He was stationed in Bankipore, Bihar, where he worked until 1885, rising ultimately
to the posts of Commissioner and Collector. Grierson conducted ethnological and linguistic research during this period, which formed the basis for two of his works, *Seven grammars of the dialects and subdialects of the Bihari language* (1883–1887) and *Bihar peasant life* (1885), before beginning the work of proposing a colony-wide linguistic survey. Aided by the support of the renowned Orientalist Max Müller and the Sanskrit philologist Sir Monier Monier-Williams, Grierson proposed the idea for a systematic survey of India’s languages and dialects at the 1886 International Congress of Orientalists in Vienna. The proposal was swiftly rejected by the central government, which cited its exorbitant cost and the shortage of competent officers who could be spared from their normal duties.³

For several years after Vienna, Grierson wrote hundreds of letters to the provincial governments and Home Department seeking support for the survey. The prospect of an all-India linguistic survey slowly gained favor with the Government of India once it was established that the results could usefully coincide with the publication of the following decade’s census (Pandit 1975: 73). After five years of negotiation, the survey was approved in 1891 on the basis that it would take two years to complete with an allotted budget of Rs. 2,000 per year, with Grierson to be the sole officer dedicated to working full time on the project. Grierson would go on to employ Ram Gharib Chaube, chief clerk and translator for the LSI, who was not credited for his role and was largely forgotten when Grierson retired from the civil service and returned to England in 1899, where he would go on to edit and compile the survey until its completion in 1927 (Amin 2011).

Once plans were finally in place, it was decided that the provinces of Madras, Burma, and the princely states of Hyderabad and Mysore would be excluded from the survey. In letters between Grierson and the various provincial governments seeking support for the survey in the late 1880s, administrators from the Madras presidency and Burma decided that they could not spare the financial or administrative resources.⁴ The final area covered by the survey thus extended from Kashmir in the north, the North West Frontier

---

³ The survey was initially proposed to cost 363,000 rupees and would have required a team of 14 assistant directors, with provincial civil servants trained to carry out data collection in a uniform way, so that the survey did not devolve into “a mass of paperasseries”, or red tape (Singh 1969: 142).

⁴ See Singh (1969: 156–167) for reproductions of letters between Grierson and the Home Department in which the decisions to exclude the Madras Presidency and Burma from the survey are discussed. The Nawab of Hyderabad suggested in 1891 that the linguistic survey co-occur with the Archaeological Survey of India, but the Indian government did not want to spare their experts. Once the plans for the LSI were being finalized in 1889, Hyderabad said that it could not contribute funds for the project (Singh 1969: 150–152).
Provinces in the West, to Assam in the East and the Bombay Presidency in the South (cf. Figure 1). Ultimately, however, Grierson did include a volume dedicated to the Munda and Dravidian languages (Volume 4), written by Norwegian Indologist Sten Konow, using secondary materials and specimens returned from the Central Provinces, Odisha, and Bihar.

With the survey approved, and the collection of data left to untrained local administrators, a simple method was devised for capturing the necessary information for basic descriptions of each form of speech. District-level officers were asked to complete a simple questionnaire for each of the languages and dialects in their territories that would then be sent to Grierson to be analyzed.5

The questionnaire that volunteer data collectors were asked to complete contained four columns. Column I was to contain the names of every local dialect spoken in the district “by which it is locally known” (Singh 1969: 203). Column II required the “popular classification” of the dialect based on what “intelligent people of the locality think about their own language” (Singh 1969: 204). Column III was to contain figures for the number of speakers, while Column IV was left for remarks containing any special information about the dialect or if any local literature was available.

---

5 The questionnaire that volunteer data collectors were asked to complete contained four columns. Column I was to contain the names of every local dialect spoken in the district “by which it is locally known” (Singh 1969: 203). Column II required the “popular classification” of the dialect based on what “intelligent people of the locality think about their own language” (Singh 1969: 204). Column III was to contain figures for the number of speakers, while Column IV was left for remarks containing any special information about the dialect or if any local literature was available.
“specimens” were then to be collected for each dialect surveyed, including (1) a translation of a passage from the Biblical Parable of the Prodigal Son6 (2) “a piece of folklore or some other passage in narrative prose or verse” (Grierson 1927: 17) and (3) a standard list of words and sentences to be translated. In its final product, the LSI presents 179 languages and 544 dialects, with short descriptions of the grammatical and phonological rules of each, followed by reproductions of the three specimens that were returned by data collectors.

2.1 Scope of investigation

Given the enormous complexity of the finished product, an article of this length can only begin to explore how processes of racialization arise within the LSI. I have included here analyses of portions of Volumes 3, 4, 5, and 6, which contain information on the languages classified as members of the Tibeto-Burman, Munda and Dravidian, Bihari and Oriya, and Eastern Hindi families or groups, respectively, with additional analysis from Volume 1, the final published volume that constitutes an introduction to the survey as a whole. These portions were chosen because they contain information about the minority languages that are most susceptible to the processes of erasure and racialization that I describe below. When dealing with a text of this magnitude, there are of course numerous other examples to be analyzed, but I have chosen these as representative of the larger systematicity with which speakers of tribal languages in particular are racialized through ideologies about language, ethnicity, and personhood.

A primary arena of analysis that is absent in this article is the ideological role that literature and literacy plays in the survey (Carlan forthcoming). Grierson adhered to an Enlightenment-era understanding of the fundamental interrelatedness of language, literature, and “nationality”. His effort to shed light on the vernacular languages of India was simultaneously an attempt to elevate the study of vernacular literatures in India, which had mostly hitherto been denounced as worthless. This was evidenced in “Anglicist vs. Orientalist” debates on the relative value of English vis-à-vis vernacular languages as a

---

6 The usage of the Parable was contentious, as the government wanted to avoid suspicion that the Survey was connected to missionary work. In a letter from Grierson to J. P. Hewett, Secretary of the Government of India, on 13 April 1896, Grierson explains that the Parable was selected because “a large number of translations were already available which could be used for samples to the translators”. For those native translators who did not know English, they could thus rely on one of the approximately 40 different existing translations of the passage. There were also changes to the Parable in which the mention of a “swine” was to be replaced with “goat” so as not to be “offensive to native prejudices” (Singh 1969: 264).
medium of education in India. While the former entirely denounced the study of Indian literatures as worthless vis-à-vis European literature (exemplified by Thomas Babington Macaulay’s infamous *Minute on Indian Education* [1972]), the latter was rooted in the revolutionary Sanskrit philology of William Jones that promoted a hearty study of India’s classical languages and literatures as a window onto the history of Indian and European human relatedness. Grierson, contrary to both positions, advocated the need to study India’s vernacular literatures, which he had dedicated much of his time in the civil service to researching. The central place of literature within the survey goes hand-in-hand with the broader processes of racialization that occur on the basis of ethnological texts; while Grierson is careful to acknowledge any body of vernacular literature, no matter its size, his definition of literature is still nonetheless highly Eurocentric and does not account for oral forms of literature. The presence of written vernacular literatures is highly influential for Grierson’s decisions about whether to classify a form of speech as a language or a dialect, and is steeped in the Herderian tradition of viewing literature as evidence of distinct nationalities (Bauman and Briggs 2003: 163–196). Indeed, Grierson’s assertion that it is the presence of a written vernacular literature that constitutes a separate “nationality” of a group has been the foundation upon which ethno-linguistic movements in postcolonial India have coalesced. The Maithili Movement, for example, has drawn inspiration from Grierson’s classification of Maithili as “the most important distinct language of Bihari, and the only one with a literary tradition” (Brass 2005: 62).

In addition, this short article is also unable to address the legacy of the *LSI* both for its lasting utility for later studies of linguistic variation in India (Gumperz 1957, 1958; Ferguson and Gumperz 1960; Gumperz 1971a; Shapiro and Schiffman 1981; LaDousa 2004), as well as its role in subsequent language policy and planning (Mawdsley 2002), and the rise of linguistic nationalism and counter-movements like the Dravidian movement (Lelyveld 1993; Trautmann 2006). Finally, there are pertinent questions I am unable to address about how the *LSI* affected the subsequent development of the discipline of linguistics in the late colonial and postcolonial period in India, all of which constitute fruitful areas for further analysis.

---

7 While it is not within the scope of this article to discuss Bauman and Briggs’ interpretation of Herder (but see Piller 2016 for one critical view), I do believe that their discussion of Herder’s views on the importance of vernacular literature for the construction of nationality are highly applicable to the *LSI*. 
3 The politics of counting in the LSI

An all-India language survey of the magnitude of the LSI reflected a larger move under British colonial administration at the turn of the century to collect systematic data on the population of India. The LSI was to serve the purpose of giving civil servants the “command of language” required in order to rule with authority (Cohn 1996) and mirrored the growing concern with population management under nineteenth and twentieth century European modes of colonial governmentality (Foucault 1991).

Much as the objectification of caste in the colonial census set the stage for communitarian politics during the nationalist movement and after Partition (Appadurai 1993; Cohn 1987), Grierson’s work enabled the formation of a new criterion upon which Indian groups could be grouped and categorized: the form of their speech, captured through a brief grammatical skeleton and a few translated passages. The LSI formed the basis of the “linguistic areas” that guided subsequent scholarly investigations of the distributions of regional languages (Gumperz 1971b), which have furthermore been the basis for political contestations over statehood in the post-independence era. With the move to divide independent India along lines of linguistically “homogenous” states, language has become intimately tied to the imagination of regional identity and has spawned ongoing separatist movements of the twentieth and twenty-first centuries on the basis of linguistic difference (Mawdsley 2002), most recently in the case of the separation of the majority Telugu-speaking region of Andhra Pradesh into the new state of Telengana in 2014. The construction of community along linguistic lines laid out in the LSI has thus been a crucial site of mobilization in postcolonial Indian politics (Sarangi 2009: 207). Indeed, “Grierson’s maps and boundaries have been periodically used by political pressure groups to ‘claim’ an area for a language group”, demonstrating the political significance of Grierson’s having identified, named, and separated previously fluid and unnamed languages, despite his ambivalent stance toward the ability to clearly demarcate languages, dialects, and their speakers (Pandit 1975: 78).

Dipesh Chakrabarty has argued that the inextricability of measurement and government in colonial projects has had a lasting impact on political imagination in independent India, particularly for the measurement and distinction of ethnic groups (Chakrabarty 1994). This process involved the transformation of “fuzzy” communities into “enumerated” ethnic identities, that elevated ethnicity as a category to be compared and ranked in scales of relative “backwardness” (Chakrabarty 1994: 150). The salience of ethnicity as an instrument of colonial
rationality and control was similarly applied to language, although the distinction between ethnicity and language is extremely muddied in the survey. The quantification of languages and dialects and the counting of their speakers are intimately tied to ideologies about ethnic and racial difference. In the survey, each language or dialect is presented with estimates of the numbers of speakers and where they are located, using the census of 1891 as a basis and modifying those figures to reflect current numbers collected from the survey.

In deciding how to quantify the number of speakers for each language and dialect, Grierson is often forced to rely on secondary ethnological data to rationalize how to group forms of speech. This is especially true where speakers have made claims to speak an independent language or dialect, and the author questions the legitimacy of such claims. Decisions about how to quantify the number of speakers of a dialect are inevitably tied up with ideologies about speakers’ ethno-racial characteristics. This process is clear in the portions of Volume 4 that describe the Munda and Dravidian languages spoken in the tribal region known as “Chota Nagpur”, which today makes up parts of the states of Jharkhand, Odisha, West Bengal, Bihar, and Chhattisgarh. In the entry for the “dialect” of Asuri, the author presents ethnological and linguistic data written in an article from the Asiatic Society of Bengal by Reverend F. Hahn (Hahn 1900), a Prussian missionary stationed in Chyebasa (now Chaibasa, Jharkhand) at the Gossner Evangelical Lutheran Mission. After describing the Asur tribe as broken into “totemestic sections” that are similar to “other Aborigines in Chota Nagpur”, Grierson (as editor) expresses uncertainty about the figures that were reported on the numbers of speakers of different “sub-dialects” in the survey’s specimens, especially the figures for Mundari and Manjhi. Manjhi is described as “extremely difficult to classify” due to its being a “very corrupt and mixed form of a dying language” (Grierson 1906: 145). As a result, when quantifying the speakers of Asuri, figures are reported under the label “so-called Manjhi” and “so-called Mundari” (Grierson 1906: 136), while the entirety of the Asuri dialects are described as “rapidly dying out, and their total disappearance can only be a question of time” (Grierson 1906: 136).

The same is true for another tribal group from Chota Nagpur, Birhar, the entry for which begins with a description by ethnologist Herbert Risley that describes them as “eke[ing] out a miserable living” in the jungle (Grierson 1906: 102). After this short description, the number of speakers is estimated at 1,234 based on information collected for the survey in three districts of Chota Nagpur. However, the author also mentions that there are “some few Birhars” in other districts, yet “no estimates are available, and their number is unimportant” (Grierson 1906: 102). The descriptions of forms of speech as corrupt mixtures contribute to decisions about how to (or not to) quantify them in the larger
statistical project of the survey. Subsequent censuses have followed Grierson’s logics of enumeration by ceasing to include languages with less than 10,000 speakers beginning with the census of 1971, dramatically decreasing the number of languages represented (Devy 2014: 6). Representations of these racialized speakers, described in various ways as primitive, are therefore bound up with ideologies about their forms of speech as corrupt, mixed, and impure. These data offer a complementary case of the ideology of “languagelessness” (Rosa 2016) that has structured educational and census policies targeting minority groups in the U.S., especially Spanish speakers, who have historically been deemed racially and linguistically mixed, and therefore unable to speak any language properly (Leeman 2004). The division between the “Aryan” languages of north India and the Munda and Dravidian families of the south was crucial to Grierson’s crafting of the survey - indeed, he argued that the “aboriginal tribes” of the south had succumbed to “Aryan civilization and influence”, with their languages being “the first thing to yield” (Grierson 1903a: 33; cf. Ramaswamy 1997: 41).

### 3.1 Linguistic mapping

Even as the LSI relies on language ideological statements for deciding whether and how to classify and quantity groups of speakers, Grierson simultaneously emphasized the need to avoid fixed and definitive representations of the linguistic landscape of British India. Boundaries between languages and dialects were seen as unstable, and shaded into one another gradually, such that a clear line could not be established. Javed Majeed has argued that Grierson, in his hesitation to fix boundaries between languages, embraced “a different kind of epistemological style, one that is more open-ended, provisional, and conscious about the limits of knowledge” (Majeed 2015: 224). As such, Majeed argues that colonial knowledge more broadly was not a “coherent and hegemonic instrument of rule”, but instead was “a more fractured, dialogically produced, potentially open-ended, and socially unstable constellation of ideas and practices” (Sengupta and Ali 2011: 6).

Majeed’s argument that the LSI flouted and destabilized the norms of colonial scholarship is representative of much of what Grierson explicitly stated were the aims of the survey. Grierson’s ambivalence toward cartographic sensibilities was not novel, as colonial scholarship on the languages of India had acknowledged the fluidity of boundaries between groups of speakers for at least half a century before the LSI. In 1853, Sir Thomas Erskine Perry, retired chief justice of the Supreme Court of Bombay, published
an article entitled “On the geographical distribution of the principal languages of India, and the feasibility of introducing English as a *Lingua Franca*”, later described in the census of 1961 as “the first attempt to give a correct estimate of the linguistic situation in India” (Singh 1969: 58). The article exhibits the same skepticism toward mapping languages that is present in the introductory volume of the *LSI* (1927), as Perry “observes that the limits of two neighboring languages often occur in wild unexplored and unpeopled tracts of the country so as ‘to prevent the tracing of a precise boundary line’” (Singh 1969: 58). While both Perry and Grierson did their best to explain the limits of establishing borders between languages, they nonetheless did create linguistic maps, even as they were seen as abstractions from the reality of sociolinguistic life. As the *LSI* was itself created as an instrument for government, Grierson had to make judgments about classification and enumeration that often relied on language ideological assessments of forms of speech as more or less pure, unmixed, and civilized, which were furthermore based on statements about speakers’ ethnic and racial characteristics as evidence of their identities as distinct populations.

### 4 Linguistic difference as ethnoracial difference in the *LSI*

Even as the *LSI* challenged many of the assumptions of philological and ethno-logical work of its time, the text continues to rely on deeply entrenched logics of racialization that characterized colonial knowledge about India. The move away from racial understandings of language and culture is part of the broader intellectual context of the ethnological work in the early nineteenth century, deeply influenced by the work of Franz Boas, who broke with the evolutionary paradigm in his effort to demonstrate that language, race, and culture were not mutually constitutive, but instead are independent, historically and environmentally conditioned phenomena (Stocking 1968). Boas, in his effort to decouple language from race and culture, sought to engage the scientific study of these categories on their own terms, as independent processes. In doing so, however, he worked to reify race as an essential feature of biology, separate from culture, to be studied by physical scientists (Visweswaran 1998). For Grierson, then, the insistence on avoiding discussions of racial origin in the introductory volume is subsumed by the centrality of race to the construction of the West as distinct from the colonized world in social scientific knowledge (Trouillot 1991).
In the section of Volume 1 entitled “General results of the survey,” Grierson lays out his clearest explanation of his approach to philology and ethnology: “Nowhere are there presented stronger warnings against basing ethnological theories on linguistic facts than in India. There are many instances of tribes which have in historic times abandoned one language and taken to another” (Grierson 1927: 28). Grierson draws inspiration here from Max Müller, who early in his career advocated for the notion of the racial unity of the “Aryan” people in the effort to argue for the kinship of Europeans and Indians. Towards the end of the nineteenth century, however, “race science” had grown in influence, and Müller made a famous proclamation in 1891 that “the unholy alliance between philology and physiology had hitherto done nothing but mischief” (Müller 1892: 185). Indeed, it was the move to de-link language from race that may have actually ignited the rise of racialist thought in the nineteenth century (Trautmann 1997: xxiv).

While Grierson aligns himself explicitly with Max Müller’s critique of the “unholy alliance” of philology and ethnology in the introductory volume, he was also in close contact with Herbert Hope Risley, who argued that the caste system evolved from systems of endogamy meant to preserve racial purity and championed anthropometric techniques as a method for classifying the races of India (Risley 1908). The imprint of Risley on the LSI is felt most acutely in Volume 4, which contains descriptions of dialects collected from tribal regions of India where both Grierson and Risley collected ethnological data. This volume was initially written by Dr Sten Konow of Christiana, Norway, was reviewed by Mr V. Venkayya, Government Epigraphist from Madras, and ultimately finalized and edited by Grierson (Grierson 1906: 1). Many of the volumes entail this kind of distributed authorship (Majeed 2011), making it difficult to distinguish between the positions of Grierson and those of his collaborators. Yet, as he states in the introductory notes to the volumes that he did not initially write himself, Grierson holds himself “responsible for all the statements contained in” the volumes as its principal editor (Grierson 1906: xv).

Grierson and Risley, as scholarly collaborators in similar ranks of colonial administration, were well acquainted with one another’s work. Whereas there are myriad citations of writings as supplementary sources throughout the survey, Risley is the only one who is quoted at length, but never actually cited. There is constant mention throughout Volume 4 of his work, usually beginning with “According to Mr Risley”, followed by long passages of quotations containing descriptions of the ethnic and racial makeup of a given ethnic or caste group. Often, the labels attributed to different dialects are interchanged with the names of castes, which in Risley’s formulation are equivalent to racial differences. Other sources are cited after each entry under the heading “Authorities”, but Risley’s voice is so authoritative that it need not even be
attributed. His words, however, reverberate throughout the text, making it difficult to disentangle the position of Risley from either Grierson as editor or the authors of the volumes written by other specialists.

Grierson presents sections of Risley’s writings on anthropometry and his attendant scheme of racial classification of the physical types of India throughout Volume 4. Many of the dialects listed are accompanied by ethnographic sketches of the features of the tribe as detailed in Risley’s writings. A description of the physical characteristics of the “Dravida race” is the first information presented in the introduction to this volume. The text quotes Risley’s anthropometric measurements and descriptions of the “Dravidian type”: “the lips are thick; the face wide and fleshy; the features coarse and irregular […] the figure is squat, and the limbs sturdy. The colour of the skin varies from dark brown to a shade closely approaching black” (Grierson 1906: 1). There are slippages in this early volume between the categories of language and race, which on the one hand must be distinguished, while on the other hand they are presented as intimately related:

According to Müller, man can only have developed a real language after having split up into races, and the various languages in actual use must therefore be derived from different racial bases. Nay, it seems even necessary to assume that the individual race had often split up into further sub-divisions before developing a language of its own. All the languages of one race are not, therefore, necessarily derived from the same original. (Grierson 1906: 2)

Thus, language is presented as an essential form of human difference that was a necessary consequence of racial difference. The problem for Grierson was that because language is a poor indication of racial origin, it is necessary to draw on other tools (namely, grammar) to determine the relationships of speakers to one another. Yet, Grierson constantly falls back on information about “ethnic differences” to distinguish linguistic groups from one another that are otherwise grammatically identical.

This occurs throughout Volume 4, where the authors and editor (Konow, Venkayya, and Grierson) describe and classify languages on the basis of corresponding ethnological characteristics. The entry on “Ho or Larka Kol” contains a description of a “tribe” that is said to speak a language called “Ho-kaji”, that is described as “exactly like Mundari”, the language of another tribe in the region (Grierson 1906: 118). The entry states:

In Singbhum they are usually known as the Larka Kols, i.e. the fighting Kols. Mr. Bradley-Brit rightly remarks that they have fully justified this name. ‘As far back as their annals go, they are found fighting, and always crowned with victory, driving back invaders or carrying war and devastation into the enemy’s lands.’ (Grierson 1906: 116)
Although the Ho are presented as an independent tribe with an independent language, their language is further described “almost identical with Mundari. The only difference of importance is the different treatment of the cerebral r. It is retained in Mundari, but dropped in Ho” (Grierson 1906: 118). This single difference in phonetic pronunciation is the only linguistic evidence that the authors offer for distinguishing this dialect from its neighbors, whereas the remainder of the entry draws on the differences in ethnicity between Mundari and Ho speakers.

The entry continues by describing the area in which the tribe is said to live, with accompanying figures estimating the number of speakers across those territories. Thus, the grammatical criteria upon which dialects are to be classified as distinct in the LSI are here subsumed under an ethno-historical consideration of what makes this group essentially different: their primordial identity as a fighting tribe. The decision to classify Ho as an independent language is made solely on the basis of the brief ethnological descriptions derived from secondary sources, rather than on the basis of grammatical description alone (the “scientific” method Grierson advocates).

The discourse of the “fighting tribe” is reminiscent of earlier colonial portrayals of the “martial races” of India, and as such relies as much on notions of racial difference as it does on construction of masculinity (Sinha 1995). The colonial imagination of certain “martial races” in India have been constructed through a juxtaposition of these groups to other “effeminate” Indians, both of which were derided as inferior to the colonizer even as they were situated hierarchically vis-à-vis one another in an internal scale of Indian development. There is thus a constellation of features that work to co-construct the discourses of the martial and the effeminate - not only through race, ethnicity, and gender, but also, in the LSI, through language.

“Linguistic practices that index social groups or activities appear to be iconic representations of them - as if a linguistic feature somehow depicted or displayed a social group’s inherent nature or essence (Irvine 1989). This process entails the attribution of necessity to a connection (between linguistic features and social groups) that may be only historical, contingent, or conventional” (Gal and Irvine [1995: 973], original formatting).

In the example of Ho, the description of the language is made possible through both the erasure of internal heterogeneity, such that the tribe is seen as contiguous and their language unchanged across the vast territories in which they live, as well as the racialization of the speakers as a unified tribe with a primordial identity. There is thus no basis for distinguishing their language given its “identical” features to Mundari, save for these descriptions of their essential ethno-racial difference as a “fighting tribe”.

Brought to you by | University of California - Los Angeles - UCLA Library
Authenticated
Download Date | 7/9/18 2:47 PM
4.1 Masculinity and racialization: the dialects of Bihari

The prominence of masculinity as a co-construct alongside race is present both in the earliest and the latest volumes of the LSI. In Volume 5, Specimens of the Bihari and Oriya languages, published in 1903, and in the introductory volume published in 1927, we see tropes of effeminacy that echo the writings of earlier orientalist writers, particularly Alexander Dow and Robert Orme, through which British colonizers came to see themselves as both physically and intellectually superior to their Indian subjects (Said 1979; Fanon 2008 [1952]). The languages and dialects of Bihar were most familiar to Grierson from his time researching the area as a civil servant, so he had a level of expertise that informed his decisions about how to represent the languages and dialects from the region. Having written on the subject of Bihar’s languages since the 1880s, Grierson had long been embroiled in the tricky decision of how to classify, group, and distinguish its languages and dialects, and was well-acquainted with the cultures of the region.

In his early writings on the languages and cultures of Bihar, he was the first to designate “Bihari” as a language distinct from Hindi, Bengali, and Oriya, and ultimately relegated Bhojpuri, Maithili, and Magahi to the status of Bihari dialects (Hastings 1996). He reached this conclusion after originally designating Maithili as an independent language due to its mutual unintelligibility with Hindi and Bengali, writing in 1881 that the region where it is spoken, Mithila, “is a country with its own traditions, its own poets, and its own pride in everything belonging to itself” (Grierson 1881: 2). By the time he published the volume on Bihari specimens in the LSI in 1903, however, he had reduced Maithili to the status of a dialect, and began to distinguish it from neighboring dialects on the basis of the “ethnic peculiarities” (Grierson 1903b: 3–4) of its speakers. These peculiarities are what justify his decision to classify them not only as separate dialects, but as indexical of differences in nationality, conceived primarily as features of personhood - strong vs. weak, educated vs. boorish, masculine vs. effeminate.

Grierson distinguishes Bhojpuri from Magahi on the one hand and Maithili on the other on the basis of what he sees as their essential ethnic differences. Describing the peasantry of Magadha as “oppressed for centuries” by “contending Musalman armies” that have left the people “uneducated and enterprising,” he then goes on to assess the “national character” of Magadha:

There is an expressive word in Eastern Hindostan which illustrates the national character. It is ‘bhades,’ and it has two meanings. One is ‘uncouth, boorish,’ and the other is ‘an inhabitant of Magadha.’ Which meaning is the original, and which is the derivative, I do not know: but a whole history is contained in these two syllables.

(Grierson 1903b: 5)
Mithila, on the other hand, is home to a Brahman sect that, despite “conquest after conquest”, has held fast to its “ancient peculiarities” (especially their “uncivilised pride”) that has resulted from their “Brahminical domination” (Grierson 1903b: 4). Its current population is said to “multiply and impoverish the earth” while refusing to take up other forms of life than agriculture (Grierson 1903b: 4). Contrastingly, the speakers of Bhojpuri are presented as superior to their closely related neighbors:

Suffice it to say here that Maithili and Magahi are the dialects of nationalities which have carried conservatism to the excess of uncouthness, while Bhojpuri is the practical language of an energetic race, which is ever ready to accommodate itself to circumstances, and which has made its influence felt all over India. The Bengali and the Bhojpuri are two of the great civilizers of Hindustan, the former with his pen, and the latter with his cudgel. (Grierson 1903b: 4)

While the speakers of Maithili and Magahi are characterized as inherently oppressed, multiplying uncontrollably, and clinging to their ancient ways, Bhojpuri is presented as enterprising and masculine:

They form the fighting nation of Hindostan. An alert and active nationality, with few scruples, and considerable abilities, dearly loving a fight for fighting’s sake, they have spread all over Aryan India, each man ready to carve his fortune out of any opportunity which may present itself to him. (Grierson 1903b: 5)

The speakers of Bhojpuri are described as “an enterprising tribe found in numbers all over the land”, including outside India, whereas “the Maithili” are a “timid, staying-home people” (Grierson 1903b: 16). The speakers of Bhojpuri are thus seen as possessing these essential ethno-national characteristics that further map indexically onto the features of their language, seen as iconic of simplicity and terseness: “Such are the people who speak Bhojpuri, and it can be understood that their language is a handy article made for current use, and not too much encumbered by grammatical subtleties” (Grierson 1903b: 5). This notion re-appears in the description of Bhojpuri verb conjugation, where Grierson writes that Bhojpuri has “cast aside all that maze of verbal forms which appalls [sic] the student when he first attempts to read Maithili or Magahi” (Grierson 1903b: 42). Similarly, the Magahi language is “condemned by speakers of other Indian languages as being rude and uncouth like the people who use it”, which Grierson attributes to the fact that there is indiscriminate usage of term “re” at the end of questions, usually used for inferiors, but applied by Magahi speakers even to superiors. The result has been a popular song, which Grierson attaches to the beginning of the section: “Magah is a land of gold. The country is good, but the language is vile.
I lived there and have got into the habit of saying ‘re.’ Why, ‘re,’ do you beat me for doing so?” (Grierson 1903b: 30).

The ethnic characteristics of Bhojpuri speakers thus become the semiotic ground upon which ideologies about language arise through Grierson’s metapragmatic statements about Bhojpuri grammar. These racializing discourses of essential ethnic difference thus work in concert with ideas about masculinity in constructing a linguistic and “national” hierarchy between the speakers of Bhojpuri, Maithili, and Magahi.

4.2 Grammar and “abstract thought”: The tibeto-Burman family

Volume 3, The Tibeto-Burman family, written by Sten Konow and approved by Grierson, contains some of the clearest examples whereby early linguistic science was able to racialize speakers by positioning their grammars as evidence of primitiveness. Max Müller’s earlier writings on the stages of language development and the relationship of grammar to abstract thought are cited as evidence for classifying the Tibeto-Burman languages as members of a pre-modern form of speech. We are told that “the need for abstract nouns grew with the progress of civilization”, and that because the Tibeto-Burman languages are only able to express concrete nouns, and do not possess a “real verb”, that they are unfit for expression of abstract thought (Grierson 1909: 46).

Grierson/Konow draw on Müller’s description of this as a failure to distinguish between form and substance, where form is the general category needed to think in the abstract. Whereas Boas attributes this to a cultural difference rather than an “inability to form them” (Boas 1911: 66) - “Primitive man [...] is not in the habit of discussing abstract ideas” (Boas 1911: 64) - the lack of “form” words in the Tibeto-Burman languages is seen as a direct result of being in an earlier stage of development. “Most Tibeto-Burman languages further evince a difficulty in forming words for abstract ideas. This is again a consequence of the uncivilized state of the tribes speaking them” (Grierson 1909: 5). Thus, they are unable to conceive of “a father in the abstract, who is not the father of any particular individual” because this is “an idea which requires a certain amount of reflection” that is precluded by the structure of the language (Grierson 1927: 46). The fact that these languages are isolating (monosyllabic), and devoid of particles, is further seen as a source of their insufficiency for higher thought. We are told that the absence of particles doesn’t necessarily “preclude higher thinking”, as the Chinese have demonstrated, yet the Tibeto-Burman languages “have nevertheless remained in the stage of individual conceptions and are
unable to give expression to abstract ideas” (Grierson 1909: 5). These features of Tibeto-Burman grammar are seen as the natural consequence of their speakers being “tribes in a primitive stage of civilization” (Grierson 1909: 5), thus reinforcing the collapse of distinctions between philological and ethnological data in the text.

Grierson echoes such statements on the relationship of language to thought in various volumes of the LSI, including in the last-published introductory volume (1927) in which he urges a move away from using philological data to speculate on racial origin. Here, however, he continues to rely on an evolutionary approach that viewed linguistic structure as both reflective and constitutive of biological characteristics of populations. Grierson describes Pashto as “a strong, virile language, [...] capable of expressing any idea with neatness and accuracy” (1927: 106). Pashto phonology, however, is described as “harsh-sounding”, and forms the basis for an indexical relationship between linguistic features and forms of personhood: “The rugged character of its sounds suits the nature of its speakers and of the mountains that form their home, but they are most inharmonious to the fastidious ears of other oriental lands” (Grierson 1927: 106). Thus, Grierson echoed much of the evolutionary assumptions about linguistic structure as an index of cultural evolution in his final writings for the LSI.

### 4.3 Intertextuality and racialization

There is an enormous number of secondary sources, largely colonial and missionary ethnological texts, that are integrated into the Grierson’s analysis and influence the way in which he describes the features of the languages and dialects represented. Intertextuality has been described as a process that “imbues texts with order, unit, and boundedness”, such that the dialogic, heteroglot nature of the speech genre becomes foundational to its authority (Briggs and Bauman 1992: 149; Bakhtin 1981). The re-entextualization of ethnological documents within the body of the text often seamlessly blurs the distinctions between the classificatory goals of the survey and the discursive production of colonial power over native, especially tribal, subjects. The inclusion of these sources under the heading “Authorities” within each entry contributes further to the production of ethnological data as “fact” in the LSI, which relies on orientalist discourses and tropes that support racist ways of conceptualizing Indian society.

An example of how this intertextuality works to racialize linguistic groups is evident in Volume 6, on the dialects of Chhattisgarhi, which includes several languages spoken by tribal groups in central India. In the section of Volume 6 on the specimens of the Baiga people of Chhattisgarh, Grierson reproduces a
portion of text from Sir Benjamin Robertson’s Report on the census of 1891 of the Central Provinces. Robertson then quotes Captain James Forsyth (b. 1838 - d. 1871), a civil servant who worked as a settlement officer, and whose travelogues were published posthumously in *The Highlands of Central India* (1872). Forsyth writes:

> The real Baiga of the hill ranges is still almost in a state of nature. They are very black, with an upright, slim, though exceedingly wiry frame, and show less of the negretto type of feature than any other of these wild races. [...] the Baiga is the very model of a hill aborigine.
> (Grierson 1904: 233)

This passage represents an isolated tribe living as one with forest animals, naïve in their dealings with traders and merchants, with no legal authority and a patriarchal system of self-government. We are then told that the Baiga dialects of Chhattisgarhi are a “corrupt form”, “composite jargon”, (Grierson 1904: 235) or a “broken dialect”, (Grierson 1904: 238) and unworthy of full description beyond inclusion of the collected specimens. The same is true of the entry for Kalanga and Bhulia, two other dialects of Chattisghari, which are described “mere corrupt jargons of an uneducated people. It is unnecessary to analyze their corrupt grammatical forms” (Grierson 1904: 251).

Indeed, in the introductory section to the entry for Chhattisgarhi, its six “broken dialects” are described as “spoken by aborigines who have forsaken their ancestral tongues, and try to speak the language of their Aryan-speaking neighbours,” the correctness varying “a great deal upon the personal equation of each speaker” (Grierson 1904: 25). Thus, the Chhattisgarhi dialects are the unsuccessful attempts of tribal groups to acquire the more sophisticated Aryan languages, which become broken admixtures “in the mouth of an aborigine” (Grierson 1904: 25). The survey’s linguistic descriptions are thus intimately tied to forms of knowledge produced by colonial ethnographers, in which racialized “tribal” groups were not only denied coevalness (Fabian 1983) but also “overdetermined from the outside” (Fanon 2008 [1952]: 95), such that their ethno-racial characteristics shaped the representation of their languages, and vice versa.

### 5 Conclusion

This article has explored how ideologies about language intersect with processes of racialization in the *Linguistic Survey of India*, demonstrating how the text
reinforces essentializing relationships between language, ethnicity, and personhood despite the editor’s attempt to move beyond methods of using comparative philology to determine “racial origin”. While the LSI was limited in what it could accomplish due to its embeddedness within the project of knowledge production in service of colonial administration, it had a lasting impact on the politics of identity formation in the late colonial and postcolonial contexts, as well as in sociolinguistic and anthropological scholarship on India. Grierson set a precedent, beginning with his contribution of a volume on language to the census of 1901, that would bind linguistic research to census operations, and consequently to structures of governmentality, well into the post-colonial era. The legacy of this connection is particularly salient given the decision by the Indian government in 1955 to reorganize the states along lines of imagined, yet contestable, linguistic homogeneity first constructed by Grierson, sparking further proliferation of movements by minority linguistic groups for political recognition in various regions.

The importance of language as a locus of ethnic and regional identity has also been reaffirmed by the new linguistic survey, entitled the People’s linguistic Survey of India (PLSI), currently being published in fifty volumes beginning in 2011. The PLSI, edited by G. N. Devy, is meant to be “not a repeat, a substitute, a replacement or a sequel to Grierson’s Linguistic Survey of India,” (Devy 2014: 24), but seeks to preserve endangered languages as well as the knowledge and practices of minority groups. The new survey intends to be a community-driven project through which linguistic documentation can work in tandem with efforts for promoting “sustainable development”, bringing the concerns of marginalized communities into closer dialogue with the state (Devy 2014: 25). Where the LSI was to be a tool for aiding the efficiency of British administration in the hopes of bringing “the official class into easier and closer communion with the people” (Singh 1969: 129), the PLSI similarly aims to “create[] closer links between the government and linguistic communities” in order to make India’s diverse linguistic and cultural communities represented in the project of economic development (Devy 2014: 24–25). There thus continue to be affinities between knowledge production about language and forms of governance in the postcolonial context, with the PLSI proposing to act as a tool for resisting the erasure of indigenous and marginalized “voices”, even as it aims to bring such groups closer to the state efforts at development that often threaten them development projects that threaten them. This tension is reflective of the broader ways in which development discourse has become central to the imagination of Indian identity in the post-colonial context (Gupta 1998).
The production of knowledge about language in the LSI was not only shaped by, but also predicated on, the racializing logics of colonial ethnology that preceded it. Grierson’s attempt to classify and enumerate language-as-grammar was inevitably mired in modernist assumptions of evolutionary progress that allowed him to represent many of the already racialized forms of speech (and speakers) he encountered as less worthy of analysis than their more advanced counterparts. The legacy of Grierson’s project is palpable within ongoing knowledge production about language in India and related movements for political recognition on linguistic and cultural grounds. As such, this investigation demonstrates the need for greater historical awareness of the motives and effects of ongoing conceptual slippages between linguistic description and statements about human difference in postcolonial India and beyond.

Acknowledgements: I am grateful to E. Annamalai, whose immense intellectual generosity made this article possible. Thanks also to Vinay Lal, Nivedita Nath, Nana Osei-Opare, Sebastiaan Broere, Naveen Kanalu, Mohsin Malik Ali, and Matt Reeck for their early input on this article. I also want to thank Erin Debenport, Alessandro Duranti, Candy Goodwin, Paul Kroskrity, Norma Mendoza-Denton, Constantine Nakassis, Elinor Ochs, Jonathan Rosa, Shalini Shankar, Michael Silverstein, and the numerous members of the UCLA Discourse Laboratory for their comments at various stages of writing. Thanks also to the anonymous reviewers and the editors of this special issue for their helpful feedback. All errors and shortcomings are my own.

References


