

The Contributions of J. H. Hutton and M. V. Sreedhar to the Study and Documentation of Sümi Language: A Community Linguist's Reading

Inato Y. Shikhu

PhD Scholar (Folklore) &
Member, Sümi Literature Board, Nagaland.

Abstract:

The written history of the Sümi language, known today within the community as Sütsa, has been shaped by several overlapping phases of documentation, interpretation and institutional development. Much of the earliest written record was produced by colonial administrators, Christian missionaries and external scholars who attempted to represent a previously largely unwritten language through the orthographic and analytical conventions available to them. Among the most influential figures in this history are John Henry Hutton and M. V. Sreedhar. Hutton's *Rudimentary Grammar of the Sema Naga Language, with Vocabulary* (1916) represents an important early attempt to describe Sümi grammar systematically, while his *The Sema Nagas* (1921) preserved extensive ethnographic, lexical and cultural information. Several decades later, M. V. Sreedhar's *Sema Phonetic Reader* (1976) and *A Sema Grammar* (1980), published by the Central Institute of Indian Languages, brought a more systematic phonetic, phonological and grammatical approach to the study of the language.

This article offers a critical reading of these contributions from the perspective of a Sümi community linguist rather than that of a formally trained theoretical linguist. The analysis focuses particularly on orthography, vowel representation, tonal notation, phonological representation, lexical accuracy and morphological analysis. It argues that the works of Hutton and Sreedhar remain indispensable historical and linguistic resources, but that their value should not prevent contemporary Sümi scholars from critically examining their limitations. Examples include Hutton's inconsistent representation of tone through macrons and the letter *f*, his representation of certain consonantal sequences such as *pf*, his treatment of verbal nominalisation through *Ke-*, and his representation of Sümi vowels. Sreedhar's work, although substantially more systematic, likewise contains lexical and orthographic forms that require correction when compared with contemporary standard Sütsa. His treatment of tones and his proposed *c/ch* distinction also provide useful material for reconsidering the possibilities and limitations of orthographic reform.

The article further examines the relationship between external documentation and community knowledge. It demonstrates that both Hutton and Sreedhar depended substantially upon Sümi speakers and earlier Sümi publications, showing that linguistic documentation has never been exclusively an external scholarly enterprise. The discussion therefore proposes a three-stage history: colonial and missionary documentation, systematic linguistic documentation, and contemporary community-based institutional standardisation. The final phase is particularly significant because the Sümi community is no longer merely the object of linguistic description. Through the Sümi Literature Board, indigenous scholars, community linguists and emerging institutional mechanisms, the community is increasingly participating in decisions concerning spelling, orthography, education, research ethics and language development.

The article argues that community linguistics should not be understood as a rejection of professional linguistics. Rather, it should be understood as a complementary mode of knowledge in which native competence, lived language use, historical memory and institutional responsibility interact with formal linguistic expertise. The central question is therefore not whether Hutton and Sreedhar were “right” or “wrong”, but what happens when a speech community itself rereads the documentation of its language and assumes responsibility for correcting, preserving and developing it. Such a perspective contributes to wider discussions of indigenous language documentation, linguistic authority, language revitalisation and community control.

Keywords: Sema; Sütsa; Community linguistics; Orthographic standardisation; Indigenous language revitalisation

Introduction

The Sümi tribe is one of the major indigenous communities of Nagaland, inhabiting Zünheboto District, Niuland District, and Dimapur District, with substantial populations in Kiphire District, as well as villages in Wokha and Mokokchung districts, and six recognized villages in the Tinsukia District of Assam. The current Sümi population is estimated at approximately 300,000.

The history of an indigenous language in written form is rarely a simple history of the language itself. It is also a history of who first wrote it, through which alphabet, according to whose linguistic assumptions, with whose assistance, and for what purposes. The written history of Sümi language illustrates this particularly well. The language is spoken by the Sümi people of Nagaland and adjoining areas, yet for much of the history of its written documentation it was represented in publications under the name “Sema”. The term occurs throughout colonial records and in the writings of scholars such as John Henry Hutton and M. V. Sreedhar. The historical terminology needs to be retained when discussing their works because it reflects the titles and linguistic categories of their period; nevertheless, the language is referred to as Sütsa in the present article when discussing the contemporary language.

The distinction between historical terminology and present community terminology is not merely cosmetic. “Sema” was the name used in official and external contexts for many decades, whereas within

the Sümi community the people referred to themselves as *Sümi*. The Government of Nagaland officially adopted “Sümi” as the name of the tribe following a notification dated 30 September 1995. The semantic composition of the ethnonym *Sümi* is also significant in Sümi oral tradition: *Sü* means ‘tree’, while *mi* means ‘people’. According to traditional accounts, *Süpu*, the eldest son of Khepiu, was born under a large tree, and this association is understood to have given rise to the expressions ‘son of a tree’ and, subsequently, ‘people of the tree’ (*Sü* ‘tree’ + *mi* ‘people’). Subsequently, the Sümi Literature Board adopted the term *Sütsa* to designate the language, combining *Sümi* with *tsa/atsa*, meaning ‘language’. Thus, *Sütsa* may be understood as ‘Sümi language’ or ‘the language of the Sümi’. The movement from Sema to Sümi and then to Sütsa therefore represents not simply a change in vocabulary but a wider movement from externally assigned naming towards indigenous linguistic and cultural self-identification.

The first substantial published references to Sümi language occurred in the broader colonial documentation of the Naga languages. Grierson’s *Linguistic Survey of India* classified the language within the Tibeto-Burman family. The early twentieth century subsequently witnessed the emergence of Sümi literacy through the combined activities of colonial administration and Christian mission education. In 1906, Rev. H. B. Dickson and Ivulho Shohe produced *SIMI'TSA KAKU*, described in the present study as an early printed Sümi book, while the same year saw the publication of the first Sümi hymnal, *Alhou Kishe-Le*. These publications are important because they preceded Hutton’s 1916 grammatical work and demonstrate that the history of written Sümi cannot properly begin with an external scholar alone. The early writing system emerged through interactions among missionaries, Sümi speakers and the practical requirements of literacy.

Hutton’s *Rudimentary Grammar of the Sema Naga Language, with Vocabulary* (1916) nevertheless occupies a special position in this history. It was the first substantial work devoted specifically to a systematic description of Sümi grammar. The book included discussion of the alphabet and sounds, numerals, parts of speech, syntax, calendar terminology and vocabulary. Hutton subsequently incorporated substantial lexical and ethnographic material into *The Sema Nagas* (1921). His importance is therefore difficult to overstate. Even where his representations are inaccurate from a contemporary Sümi perspective, they preserve an early linguistic record that would otherwise be unavailable.

M. V. Sreedhar represents a substantially different phase. His *Sema Phonetic Reader* (1976) and *A Sema Grammar* (1980), published by the Central Institute of Indian Languages, belong to a period when descriptive linguistics had become more systematic and when Indian language planning increasingly recognised the importance of documenting indigenous languages. Sreedhar’s work examined speech organs, speech sounds, phonemes, positional variants, phonetic drills and orthography, while *A Sema Grammar* extended the analysis to phonotactics, grammatical classes, morphological construction, causatives, phrases, clauses and other grammatical structures. The *Phonetic Reader* explicitly sought to describe the sounds of Sema, teach their pronunciation and suggest a scientific script. The introductory

discussion of the human speech organs was also deliberately written so that readers without formal linguistic training could understand the description of individual sounds (Sreedhar, 1976, p. 6).

The contributions of both scholars are therefore foundational. Yet foundational does not mean infallible. A written representation becomes authoritative precisely because later writers use it as a reference. An error made in an early grammar can consequently become an error reproduced in subsequent grammars, dictionaries, school books, religious texts and scholarly works. The problem becomes more serious in a language whose standard written form is still undergoing institutional consolidation.

This article consequently asks a broader question than whether particular forms in Hutton or Sreedhar are correct according to contemporary Sūtsa. It asks: What happens when an indigenous speech community rereads the linguistic documentation of its own language from within the community rather than simply receiving that documentation as external scholarship?

The question is important because contemporary approaches to indigenous-language documentation increasingly recognise that communities should not remain passive providers of linguistic data. Documentation becomes more useful when community members participate in the planning, interpretation, production and application of linguistic materials. Research on language revitalisation has similarly stressed community participation, capacity-building and the development of locally controlled archives and resources.

The argument advanced here is not that professional linguists are unnecessary. Nor is it that native speakers automatically possess all the theoretical tools required for linguistic analysis. Rather, the argument is complementary: formal linguistic expertise and community linguistic knowledge answer different but interconnected questions. A trained linguist may provide an analytical framework for phonology or morphology; a native speaker may recognise that a supposedly recorded lexical item is not part of the language as it is actually spoken. One may describe a sound scientifically; the other may know when that sound has been represented in a way that does not correspond to community usage. The strongest future scholarship on Sūtsa should therefore bring these forms of knowledge together.

Methodology and Positionality: Reading Sūtsa as a Community Linguist

This study employs a critical historiographical, comparative and community-based linguistic methodology. It is not presented as an experimental phonetic study, nor does it attempt to reproduce the methods of a trained laboratory phonologist. Instead, it combines historical textual analysis, comparative lexical analysis, orthographic comparison and native-speaker interpretation.

The principal texts examined are Hutton's *Rudimentary Grammar of the Sema Naga Language, with Vocabulary* (1916), Sreedhar's *Sema Phonetic Reader* (1976) and Sreedhar's *A Sema Grammar* (1980). Hutton's *The Sema Nagas* (1921), Hutton and Bor's "The Use of Tones in Sema Naga" (1923), Dickson

and Shohe's *SIMI'TSA KAKU* (1906), and Teo's phonological study are used where they help establish historical connections or contextualise the linguistic observations.

The methodology consists of four related procedures. First, the historical documents are read comparatively in order to identify how the written representation of Sütsa changed over time. Secondly, selected lexical items are mapped against contemporary Sütsa spelling. Thirdly, particular grammatical claims are examined through morphological parsing from the perspective of a native speaker. Fourthly, orthographic representations are interpreted within the broader history of Sümi literacy, publishing, religious literature, education and institutional language planning.

The study is therefore best understood as community-based insider research. The author is not a trained linguist but has developed sustained engagement with Sütsa through several forms of community language work. This includes serving as Editor and Publisher of *Sümi Zümülhü*, a daily Sümi newspaper, from 2011 to 2015; serving as a member of the Sümi Literature Board since 2021; participating as Member Secretary of the BA Sümi Honours Curriculum Development Cell since 2022; and serving as Controller of Examinations of the Academic Council of the Sümi Literature Board.

These experiences do not transform the author into a theoretical linguist. They do, however, provide a different kind of linguistic access. A language is not encountered only as an abstract phonological system. It is encountered in newspapers, schoolbooks, hymns, Bible translations, public notices, examination papers, literary works, conversations and institutional debates. A community linguist works precisely at this intersection between language as a system and language as a lived social practice.

This positionality is important because the present study deals with linguistic matters that can legitimately be questioned by professional linguists. The absence of phonetic transcription is one such limitation. The article does not provide IPA forms for the disputed words because the author does not claim the specialist training required to undertake instrumental phonetic analysis. This is a limitation, but it should not be treated as a reason to dismiss the native-speaker evidence altogether. Rather, it identifies an area where community knowledge and professional linguistic analysis should meet. A future collaborative study could test the observations presented here acoustically and articulatorily.

The article therefore adopts a position of critical humility. It neither claims that every contemporary Sütsa spelling must be linguistically optimal nor assumes that every historical form is wrong simply because it differs from present spelling. The concern is more specific: where a historical representation demonstrably conflicts with contemporary native usage, it should be identified; where an older form preserves a genuine historical or dialectal feature, it should be preserved as historical evidence; and where the evidence is uncertain, the uncertainty should be acknowledged.

Biographical Profiles of Pioneer Scholars

J. H. Hutton (1885–1968)

John Henry Hutton was a British colonial administrator and anthropologist who served in the Indian Civil Service (ICS) as Deputy Commissioner of the Naga Hills district. Though not formally trained as an academic linguist, Hutton was a field ethnographer. His long residency in the Naga Hills enabled him to achieve conversational fluency in several local vernaculars, including *Sütsa*. His publications *Rudimentary Grammar of the Sema Naga Language with Vocabulary* (1916) and his monograph *The Sema Nagas* (1921) served as the earliest systematic documentation of Sümi social structure, oral traditions, vocabulary, and basic syntax (Hutton, 1916, 1921). His linguistic work laid the initial empirical foundation upon which subsequent language documentation in the region was built.

M. V. Sreedhar (1928–2005)

Dr. M. V. Sreedhar was a professional linguist associated with the Central Institute of Indian Languages (CIIL) in Mysore, India. Operating within the framework of mid-twentieth-century structural linguistics, Sreedhar conducted field investigations across Nagaland between 1970 and 1974 (Sreedhar, 1980, p. v). His primary contributions to Sümi documentation include the *Sema Phonetic Reader* (1976) and *A Sema Grammar* (1980), published under the CIIL Grammar Series. Unlike his colonial predecessors, Sreedhar applied formal descriptive methodologies, analyzing speech organs, phonemic inventories, positional variants, phonotactics, and syntactic structures (Sreedhar, 1980, pp. 19–186). His work sought to establish a scientific script, standardized orthography, and formal pedagogical grammars for indigenous languages.

Historical Phases of Sümi Language Documentation

The history of Sümi language documentation can usefully be understood through three overlapping phases: colonial and missionary documentation; systematic linguistic documentation; and community-based institutional standardisation.

Colonial and missionary documentation

The first phase was shaped by the practical circumstances of early literacy. Missionaries needed to translate religious materials and teach literacy, while colonial officials required linguistic information for administration. Early writing conventions were consequently influenced heavily by English orthographic conventions.

The 1906 *SIMI'TSA KAKU* and *Alhou Kishe-Le* are particularly significant. They demonstrate that Sümi was already being transformed into a written language before Hutton's 1916 grammar. The problem was that the available writing conventions did not always correspond neatly with Sümi phonology. The vowel system was particularly difficult. Contemporary *Sütsa* recognises the vowels *a*, *e*, *i*, *o*, *u*, and *ü*, and the sixth vowel, *ü*, does not correspond straightforwardly to the English vowel categories through which early writers approached the language.

An early example is the word for “come”. In the 1906 hymnal it was written *eghilo*, whereas contemporary Sütša writes *ighilo*. The difference is not merely a typographical matter. It demonstrates the difficulty of representing a Sümi vowel through an English-based orthographic system. Similar difficulties affected consonantal representation. According to the International Phonetic Alphabet (IPA), the close front vowel [i] is orthographically represented by the letter *i* in Sütša, whereas in English it may be represented by the letter *e* in certain contexts. The discrepancy in orthographic representation may have resulted from an interpretation of the vowel sound based on established English spelling conventions, rather than from a deliberate alteration of the original Sütša form.

Forms involving *gh*, *q*, and *x* were difficult to represent consistently. The sound represented today by *xü*, for example, was previously represented through an underlined *khü* and only later came to be represented with *x*.

These difficulties should not, however, be treated simply as evidence of incompetence. Early writers were attempting to create a written system for a language that had not previously been widely represented in print. They were working without modern recording technology, without established Sümi linguistic terminology and, in many cases, through the mediation of a limited number of speakers. The appropriate historical judgement is therefore double: their representations contain errors, but their work made literacy possible.

This is particularly important when discussing Hutton. He should not be judged according to the standards of twenty-first-century descriptive phonetics alone. His 1916 grammar was pioneering precisely because it attempted to describe a language whose grammatical structure had not previously been systematically presented in a dedicated grammar. His work included word classes, syntax, numerals, vocabulary and other aspects of the language. Moreover, Hutton recognised that tone was important. The introduction to *The Sema Nagas* notes the tonal nature of the language and gives examples in which meaning changes with tone. The work also records dialectal differences, elisions and variations in pronunciation.

One particularly valuable aspect of Hutton’s work is his recognition of variation. He noted that different Sümi areas had differences in pronunciation and vocabulary and that some communities shortened words more than others. He also observed the frequency of vowel and syllable elision. This is important for contemporary standardisation because it demonstrates that the historical Sümi speech community was not completely uniform. Standardisation therefore involves choices among forms that have existed within a larger dialect continuum rather than simply correcting a language that was previously uniform.

Hutton’s Grammar and the Problem of Representation

Hutton’s *Rudimentary Grammar* is groundbreaking, but its weaknesses become visible when read against contemporary Sütša. One of the clearest examples concerns tone.

Hutton's representation of tone

Hutton attempted to distinguish tones through macrons and the letter *f*. In one case he used a macron, as in *Achuī*, to distinguish a high-toned form from another *Achui*. Elsewhere, however, he used the letter *f* to indicate a tonal distinction. His sentence “*Akhoh pfulunannif mono?*” is an important example (Hutton, 1916, p. 28). In contemporary Sütša this is written “*Akho pulunani mono?*” In this context, the historical *f* should not be treated simply as an ordinary consonant corresponding to a contemporary Sütša consonant; rather, it forms part of Hutton's older method of representing tonal distinction. Notice also that the word *Akhoh* ('load') written with a final *h* by Hutton has dropped the silent *h* in modern spelling.

The problem becomes more striking because Hutton was not entirely consistent in his representation. In another sentence he wrote “*Sakhalu-no Abor'limi ipfu ghe*” (Hutton, 1916, p. 12), while a specimen of a song or story omits the *f* in a closely related expression. The contemporary form given by the community is “*Sükhahu no Aborlimi yipughi.*” These inconsistencies matter because orthography becomes instructional. Learners encountering different representations of tonal distinctions may reasonably conclude that the difference is grammatical or lexical when it may instead arise from orthographic practice.

The use of *f* as a tonal marker continued in Sümi writing for some time. The contrast between *Apu* ‘father’ and *Apuf* ‘son’, and between *Akini* ‘ear’ and *Akinif* ‘wealth’, illustrates the practical consequences of this convention. Contemporary standard Sütša no longer uses final *f* as a tonal marker.

It is therefore important to distinguish Hutton's tonal use of the letter *f* from his separate representation of consonantal sequences such as *pf*. They should not be treated as one phenomenon.

The case of *pf*

The case of *pf* requires a different explanation. Unlike the *f* discussed above, *pf* is not being treated here as merely another device for marking tone.

In the common Sümi language, the consonantal sequence or sound represented by *pf* is not generally found. There are, however, some Aphuyemi villages, Sümi ancestral area, from which dialectal forms containing a sound resembling *pf* are reported. Thus, it would be inaccurate to state categorically that no Sümi variety has such a sound. The more precise statement is that *pf* is not a regular feature of the common or standard Sütša speech used across the wider Sümi community, although it occurs in certain Aphuyemi dialects.

This distinction is important because it prevents contemporary standardisation from erasing legitimate dialectal variation. A sound may be absent from the common standard language while still existing in a particular local variety. Dialectal forms should therefore be documented even when they are not incorporated into the standard orthography.

The concern arises from Hutton's representation of forms such as *pfulunannif* and *ipfu* in "*Akhoh pfulunannif mono?*" and "*Sakhalu-no Abor'limi ipfu ghe*". From the perspective of common contemporary Sütsa, the presence of *pf* is unusual. The question, therefore, is why Hutton represented the forms in this way.

One possible explanation is linguistic influence from the Angami environment, rather than a tonal convention. This should, however, be stated cautiously. The available evidence does not by itself establish that Hutton consciously borrowed *pf* from Angami or that he directly modelled these forms on Angami phonology. Nevertheless, the possibility is worthy of consideration because Hutton worked extensively among the Naga peoples and subsequently authored the important monograph *The Angami Nagas* (1921), published in the same general period as his work on the Sema. His familiarity with Angami language and society therefore provides relevant historical context when examining his representation of consonantal sounds in neighbouring Naga languages.

The point is not to conclude that Hutton's *pf* must have originated from Angami. Rather, the observation invites further comparative research into the phonological systems available to colonial-era observers and the ways in which sounds in one Naga language may have been interpreted through categories familiar from another.

This is particularly relevant because linguistic documentation does not occur in isolation. Researchers working in multilingual environments inevitably encounter neighbouring languages. When a sound is unfamiliar, the analyst may interpret it through a sound category that already exists in another language with which they are familiar. Such a possibility deserves investigation in the case of Hutton's representation of *pf*.

The present argument can therefore be stated precisely:

Hutton's use of *pf* should be treated separately from his use of *f* as a tonal marker. While *f* belongs to his inconsistent system of tonal representation, *pf* represents a different phonological and orthographic question. The sequence is absent from common contemporary Sütsa but occurs in the dialects of some Aphuyemi villages. Its appearance in Hutton's writing may reflect his interpretation of a particular pronunciation or a neighbouring linguistic influence, including the possibility of Angami influence, although the available evidence does not permit a definitive conclusion.

This formulation also leaves room for future research. A comparative study of Hutton's *Rudimentary Grammar*, *The Sema Nagas* and *The Angami Nagas*, together with contemporary recordings from Aphuyemi villages, could determine whether the historical *pf* corresponds to a genuine dialectal feature, an external phonetic interpretation, or a combination of the two.

Hutton's representation of vowels

Hutton's representation of vowels is equally important. In his discussion of interrogatives he gives the sentence "*Enakha mishi kyo, ketami mishi ghi kya?*" The contemporary form identified by the community is "*Inakha mishi keo, kütami mishi kea?*" The difference between *Enakha* and *Inakha*, and between *ketami* and *kütami*, demonstrates the difficulty early documenters faced in distinguishing Sümi vowel qualities. The Hutton form is historically valuable because it shows what the language sounded like to an external observer, but it should not automatically be treated as the correct contemporary spelling. The original passage occurs in Hutton's discussion of alternative questions (Hutton, 1916, p. 23).

Hutton's Treatment of Nominalisation

A theoretical error in Hutton's work occurs in his treatment of how nouns are derived from verb roots. Hutton asserted that nouns are formed from verbs strictly by prefixing the element *Ke-* to the root. According to him, a noun may be formed from a verb by prefixing *Ke-* to the root and, in some cases, adding *-mi*. His examples include *ti* 'die' becoming *Ketimi* 'dead man or dead men' (Hutton, 1916, p. 9).

There are several ways in which these errors can be understood. The word for 'die' in Sütsa is *thi*, not *ti* as rendered by Hutton. The use of a compound noun in his explanation further complicates his analysis, particularly when considered alongside the spelling errors in his transcription. The verb *thi* 'die' becomes the noun *akithi* when nominalised. Similarly, the compound noun meaning 'dead man' or 'dead men' should be *kithimi*, not *ketimi* as rendered by Hutton. In Sütsa, *mi* functions as a nominal suffix referring to a person, particularly in compounds denoting a person or an individual associated with a particular place, group, or category. Thus, Hutton's rendering of *kithimi* as *ketimi* appears to involve not only a spelling error but also a misrepresentation of the underlying morphological structure.

When analysed from an insider native speaker's perspective, this explanation is incomplete and structurally inaccurate for Tibeto-Burman morphology. Nominalising mechanisms (normalisers) in Sütsa function through specific prefixing and infixing structures involving *Aki-*, *Aku-*, *Akü-* (prefixes) and *-ki-*, *-ku-*, *-kü-* (infixes). The element *Ke-* rarely acts as a standalone nominalising prefix, and cannot adequately account for the full range of forms. When prefixation occurs, it almost universally begins with *A-* (*Aki-*, *Aku-*, *Akü-*). When nominalization occurs internally, it operates via infixation (*-ki-*, *-ku-*, *-kü-*). There is no infix *-ke-* for normalisers in Sütsa grammar. English prescriptive models fail to capture these internal infixing dynamics:

Verb Root (Sütsa)	English Meaning	Hutton's Asserted Rule	Actual Correct Nominalised Form	Morphological Process
Po	Run	*Kepo	Akupo (Running / Run)	Prefixing (<i>Aku-</i>)
Kive	Play	*Kekive	Akikive (Play / Game)	Prefixing (<i>Aki-</i>)
Kükha	Wrestle	*Kekükha	Akükükha (Wrestling)	Prefixing (<i>Akü-</i>)
Kimiye	Love	*Kekimiye	Kikimiye (Love -	Infixing/Prefixing

			abstract noun)	(Kiki-)
Tolu	Win	*Ketolu	Tokulu (Victory / Winning)	Infixing (-ku-)
Hebide	Disobey	*Kehebide	Hekibide (Disobedience)	Infixing (-ki-)

The significance of this comparison is not that Hutton was incapable of identifying nominalisation. He correctly recognised a relationship between verbal roots and nominal forms. His limitation was analytical: the descriptive category he used did not capture the internal morphological processes operating in Sùtsa. This illustrates an important methodological principle. A grammatical category may be visible to an external observer while its internal structure remains difficult to analyse without detailed knowledge of the language.

Hutton, Dickson and Shohe: Documentation as an Interconnected Process

A further contribution of this study is the identification of a relationship between Hutton's 1916 work and the earlier *SIMI'TSA KAKU* of Dickson and Shohe. The point is not to accuse Hutton of plagiarism. Such a judgement would require a much more extensive investigation into his working papers and research practices. The more modest and defensible observation is that linguistic documentation is intertextual.

Dickson and Shohe's 1906 primer described the consonant *G* through an English comparison involving "gun" and "singing". Hutton's 1916 grammar gives a strikingly similar explanation of *G*, and his descriptions of *J*, *R* and *Y* likewise correspond closely to the earlier primer. The 1906 formulations occur on page 5 of *SIMI'TSA KAKU*, while Hutton's corresponding description appears on page 1 of the 1916 grammar.

The importance of this observation is methodological rather than accusatory. Later linguistic works often become the sources from which subsequent scholars construct their own descriptions. Hutton's work became an important reference for Sreedhar; Sreedhar consulted Hutton and Hutton and Bor's work on tone; later researchers, including Amos B. Teo, drew on both Hutton and Sreedhar. In addition, Sreedhar relied heavily on Avito Kibami's 1961 *Sumi Tsaithi*. Thus, the history of Sùmi documentation is better understood as a chain of textual and human transmission than as a series of isolated scholarly achievements.

This has a major implication for community linguists. If an inaccurate form enters a published grammar, later scholars may reproduce it because the grammar appears authoritative. A mistake can consequently travel across generations. The responsibility of community informants is therefore not insignificant. They are not merely sources from whom information is extracted. They are participants in the production of linguistic knowledge.

Sreedhar and the Emergence of Systematic Linguistic Description

M. V. Sreedhar's contribution represents a major advance in the history of Sümi documentation. His *Sema Phonetic Reader* was published in 1976, followed by *A Sema Grammar* in 1980. Unlike the earlier colonial grammatical tradition, Sreedhar worked within a more explicit linguistic framework. His *Phonetic Reader* began with the human speech mechanism and proceeded to the phonemic inventory, descriptions of individual sounds, phonetic drills and orthography. The stated objectives were to describe Sema sounds accurately, teach their pronunciation and suggest a scientific script.

This achievement deserves strong recognition. Sreedhar's work made Sütsa available to readers in a linguistic framework that was much more systematic than anything that had preceded it. His description of speech organs and sound production also made the material accessible to readers without formal linguistic training. In *A Sema Grammar*, he extended the analysis to phonemes, phonotactics, grammatical classes, morphological construction, causatives, invariables, phrases and clauses. The grammar therefore remains an important historical linguistic resource.

The significance of *A Sema Grammar* is also demonstrated by the fact that contemporary researchers continue to engage with many of the structural questions it addressed. The grammar recorded the language at a particular historical moment and therefore provides an important point of comparison for present-day Sütsa.

Yet the very systematic nature of Sreedhar's work means that its discrepancies require careful attention. A systematic grammar can have greater long-term influence than a scattered ethnographic record because teachers, researchers and learners may treat it as an authoritative description. Consequently, its errors deserve to be identified precisely rather than ignored.

Lexical discrepancies in the *Sema Phonetic Reader*

The *Sema Phonetic Reader* contains several forms that differ from contemporary standard Sütsa. The following comparison is based on the lexical forms identified in the present study:

Sreedhar's Transcription (1976)	Page Citation (Sreedhar, 1976)	English Gloss	Modern Standard Sütsa Orthography
Sephe	p. 25	Breath	Shephe
Yeqinhey	p. 25	Quiet	Iqünhei
Awnhey	p. 25	Elbow	Aunhe
Khago	p. 29	Gate	Khagho
Alaw	p. 30	Inside	Alou
Ipo	p. 30	My father	Ipu
Buzu	p. 31	Feel	Buju
Zoga	p. 31	Playing	Zügha
So	p. 32	Hard	Sho

Aso	p. 32	Bread	Asho
Ayyepu	p. 33	Star	Ayepu

These differences should not all be classified as the same kind of error. Some involve consonantal representation, some vowels, some final segments and others a combination of these. The distinction matters because it prevents the criticism from becoming a blanket rejection of Sreedhar's work.

Sreedhar's reliance on native informants is particularly relevant here. His 1980 grammar records extensive fieldwork between 1970 and 1974 and identifies numerous Sümi speakers who assisted with the collection and checking of data. These included Avito Kibami, Vihoto Sheyipu, Lukhashe Sema, I. Lozhevi Sema and Inai Sema, among others. The grammar also states that additional data were collected later to fill gaps.

This information is important because it complicates the simple distinction between "external scholar" and "community knowledge". Sreedhar did not work without Sümi participation. On the contrary, his grammar was substantially dependent upon native speakers. The problem is that the published text may not always reveal which forms were supplied by which speaker, how competing forms were evaluated, or whether a particular form was checked against multiple speakers.

This is precisely where contemporary community linguistics can make a contribution. A future linguistic documentation project should record not merely the form of a word but its speaker, locality, age, dialect, meaning, grammatical function and degree of community acceptance. Such metadata would make it possible to distinguish an individual pronunciation from a widely accepted lexical form.

Tone and Diacritics in Sreedhar

Tone constitutes another area in which Sreedhar made an important contribution while also leaving questions for later researchers.

Sütsa has three contrastive tonal levels, generally described in contemporary terms as high, mid and low. Hutton had already recognised three tones, and Hutton and Bor (1923) specifically addressed tone. Sreedhar also recognised three tones, using low, high and level terminology. Later phonological work has continued to treat tone as central to the structure of Sütsa.

Sreedhar attempted to represent tone orthographically through diacritics acute (´) for high tone, grave (`) for low tone, and left mid-tones unmarked (Sreedhar, 1976, pp. 50–52). While many assignments were accurate, several entries misclassified mid-tones as high tones or altered word roots:

Salá ('share') (Sreedhar, 1976, p. 51): Correctly applied acute accent for high tone, compatible with present usage.

Sò ('drinking water') (Sreedhar, 1976, p. 51): Applied grave accent correctly for low pitch, but misspelled the consonant root. The correct spelling is **Shò**.

Ká ('going astray') (Sreedhar, 1976, p. 51): Incorrectly marked with an acute high tone accent; the native sound is a mid tone. The correct spelling is **kawuve** (or *Kawuve*).

Pú ('spying') (Sreedhar, 1976, p. 51): Incorrectly marked with an acute accent; native speech uses a mid tone where employing a diacritic marker is unnecessary.

Sreedhar also correctly applied tone markers in several instances (Sreedhar, 1976, pp. 51–52), though spelling adjustments are required for modern standards: **Aqò** ('pit'), **Aqhò** ('brain'), **Abò** ('stem'), **Awnhèy** ('elbow' *Aunhè*), and **Thonhèy** ('dip' *Thonhèi*).

The important point is that Sreedhar's tonal system should not be dismissed because some individual tone assignments are questionable. On the contrary, his attempt to establish a systematic tonal orthography was a significant advance over the inconsistent methods found in earlier writing. The contemporary system can therefore be viewed partly as an inheritance from this history while recognising that its implementation requires correction and refinement.

The *c/ch/chh* Question

An orthographic observation made by Sreedhar involved simplifying aspirated palatal affricates (Sreedhar, 1976, pp. 53–55). In traditional Sümi writing, the unaspirated affricate sound is written as *ch* (e.g., *chu* = 'eat'), while the aspirated counterpart is written with a double-*h* as *chh* (e.g., *chhu* = 'dig'). Similarly, *chi* ('cultivate', unaspirated) contrasts with *chhi* ('full', aspirated). This convention results in words containing multiple *h* characters, such as *Chhochhou* (crier/herald) or *Chhochhothi* (a local fruit).

Sreedhar proposed a more economical system in which *c* would represent the unaspirated sound and *ch* would represent the aspirated sound. Under such a system, forms could be simplified:

Present form	English word meaning given by Sreedhar	Sreedhar-type form
<i>Amiché</i>	Any object that becomes smaller	<i>Amice</i>
<i>Miche</i>	Taste	<i>Mice</i>
<i>Chunani</i>	Will eat	<i>Cunani</i>

Under this proposed system, written forms would streamline as follows:

Ace = a sunny place

Acè / *Acecu* = an area of land for cultivation

Acè = existed

Acequ = porcupine

Acine = rites and rituals

Acuyí = frog

If Sreedhar's *Che* ('hit') were written in present *Sütsa* orthography, it would be *chhe*. While the Sümi Literature Board has retained the double-*h* (*chh*) convention as a unique, linguistically valid marker of aspiration within its orthography, Sreedhar's exercise demonstrated structural possibilities for script optimisation.

The value of this observation lies in its demonstration that historical scholarship can still generate useful ideas for contemporary orthographic discussion. The purpose here is not to recommend replacing the present *ch/chh* system. The existing system has historical continuity through church literature, hymnals, books and education. Orthographic change must therefore consider not only phonemic efficiency but also literacy, familiarity, cultural continuity and the cost of transition.

This is a central principle of language standardisation: a theoretically economical spelling system is not automatically the best social orthography. A writing system belongs to a community as much as it belongs to a linguistic analysis. The question must therefore be whether a proposed change improves clarity, learnability and consistency sufficiently to justify the disruption it would create.

Vowel Omission in A Sema Grammar

A particularly important pattern appears in Sreedhar's 1980 grammar: the omission of vowels, especially *i*, *u* and *ü*, from several lexical forms (Sreedhar, 1980, pp. 20–45). The following examples are identified in the present study:

Sreedhar's Entry (1980)	Page Citation (Sreedhar, 1980)	English Gloss	Correct Standard Sütsa Form	Omitted Vowel Element
Kla	p. 22	Wedding	Küla	Omitted central vowel <i>ü</i>
Mbomsi	p. 24	Fool	Mubomzü	Omitted vowels <i>u</i> and <i>ü</i>
Mza	p. 25	Graze	Müza	Omitted central vowel <i>ü</i>
Mna	p. 25	Stink	Müna	Omitted central vowel <i>ü</i>
Axamnu	p. 28	Flower	Axamunu	Omitted internal vowel <i>u</i>
Mekda	p. 30	Plaster	Meküda	Omitted central vowel <i>ü</i>
Atna	p. 31	Couple	Atüna	Omitted central vowel <i>ü</i>
Homxa	p. 33	Circle	Homüxa	Omitted central vowel <i>ü</i>
Akmho	p. 35	Corpse	Akumo	Omitted internal

				vowel <i>u</i>
Amphóbò	p. 38	Hearth	Amiphobo	Omitted initial vowel <i>i</i>

The pattern is especially interesting because contemporary phonological research provides evidence that forms such as *kūla* may contain a minor syllable whose vowel can be phonetically reduced or omitted in rapid speech. Teo's later analysis, for example, discusses *kūla* and related forms and shows how such words can be realised in different phonetic ways, including a reduced form resembling *kla* (Teo, 2012, p. 368; Teo, 2014).

This observation introduces an important distinction. A vowel that is reduced or not fully realised in speech is not necessarily a vowel that should be removed from the orthographic representation. Contemporary standardisation must distinguish phonetic reduction from orthographic deletion. The presence of *Kla* in Sreedhar may therefore preserve something about actual pronunciation or the spelling practices of his period, but it does not necessarily establish that *Kla* should be the standard written form.

The same issue is visible in the Sümi Bible of the period. It is probable that Sreedhar recorded informants during a period when rapid spoken contractions were being experimented with in written form, a trend also visible in the complete Sümi Bible translation commissioned in 1978 and published in 1985:

Ipekulu 36:26

"...choqhe kepu ghenguno tsandiyino ghokepua lhobidi mla." (Baibel, 2013, O.T. p. 101)

Leshele 118:1 & 5

"...pa kikimiye ye lhokta kutoke." (Baibel, 2013, O.T. p. 671)

"Ikijiksa lono ino AMPEU kuke;" (Baibel, 2013, O.T. p. 671)

Rommi 3:17

"Panonguno alakishi laghi ye kiukeno mtha." (Baibel, 2013, N.T. p. 208)

Rommi 12:21

"Alhoksa pelo nongu theqhiluvepe kevalo ikemu akivi pesu alhoksa theqhiluvelo." (Baibel, 2013, N.T. p. 220)

In modern standard Sütsa, the underlined words above are correctly spelled with their full vowels as **mülà**, **lhoküta**, **ikijiküsa**, **mütha**, and **alhoküsa**. According to Sümi elders, the artificial omission of these vowels during that period was driven by a desire to minimise the physical page count and printing costs of the Sümi Bible.

This distinction is crucial. A community linguist should not turn an oral explanation into an unquestionable historical fact without documentary evidence. At the same time, community testimony

should not be dismissed simply because it does not originate in an academic publication. It should be recorded, identified as testimony and, where possible, tested against publishing records.

Lexical Forms That Are Not Contemporary Sümi Words

The most serious lexical concern in Sreedhar's 1980 grammar concerns forms identified by the present native-speaker analysis as not belonging to contemporary Sütsa:

Khuwasny ('rain', p. 42): Not a Sümi word.

Wenkyé ('what', p. 44): Not a Sümi word.

Aqmbó ('graveyard', p. 45): An erroneous transcription (likely attempting *akumo* = corpse/graveyard).

Awwe ('forest', p. 48): An erroneous transcription (likely attempting *ave* = forest).

These errors underscore the vulnerability of external language documentation when disconnected from native peer review. Words such as *Aqmbó* 'graveyard' and *Awwe* 'forest', appear to be possible attempts to represent other Sümi words, but the precise intention cannot be established from the form alone. A researcher may accurately record what an informant says at a particular moment and yet misunderstand the lexical identity of the word. This is particularly likely when the researcher is working across languages and when two sounds or forms are phonetically similar. A published error can subsequently become a "fact" simply because it has appeared in a book.

Intertextual Borrowing in Colonial and Structural Linguistics

Linguistic historiography often treats published monographs as isolated achievements. However, textual analysis reveals unacknowledged intertextual borrowing among pioneer documenters. When J. H. Hutton formulated his consonant sound guide in *Rudimentary Grammar* (1916), he referred to a published *SIMI 'TSA KAKU* (1906) that was written by Rev. H. B. Dickson and Ivulho Shohe a decade earlier.

Similarly, M. V. Sreedhar relied heavily on Hutton's works (1916, 1921), Hutton and Bor's tonal study (1923), and Avito Kibami's *Sumi Tsaithi (English-Sema Dictionary)* (1961) as primary sources (Sreedhar, 1980, p. 193).

Decades later, when Amos B. Teo conducted his research on Sümi tonal phonology and phonetics phonology (*Sumi tone: A phonological and phonetic description of a Tibeto-Burman language of Nagaland*), he likewise built upon the foundational work of Hutton and Sreedhar, engaging with their earlier descriptions and analyses of the language. His primary field informants were senior native language administrators, including H. S. Rotokha (Secretary of the Sümi Literature Board) and Hokishe Yephthomi (Chairman of the SLB), alongside native lexicons such as *Anglo-Sümi Tsashe* by Scato Swu and Hokishe Yephthomi (Teo, 2014, p. 3). Because of this community-guided oversight, Teo's work avoided the spelling distortions and artificial entries that affected earlier colonial and structural studies.

This pattern demonstrates an essential lesson: early literature forms an interconnected web. When an error is introduced in an early text, such as Hutton's tonal *f* or Sreedhar's dropped central vowels, it propagates through academic literature as established fact unless actively intercepted and corrected by community scholars.

Subsequent Developments in Sütša Orthographic Conventions

Since Amos B. Teo's fieldwork in Sümi-inhabited areas, certain aspects of Sütša orthographic practice have undergone further modification. At the time of Teo's fieldwork, tonal distinctions in Sütša were represented through consonantal conventions rather than diacritics. In particular, the letter *h* was used to indicate low tone, while doubling the following consonant (*hh*, *kk*, *nn*, *ll*, etc.) was used to indicate high tone. These conventions were particularly useful in distinguishing otherwise identical forms whose meanings differed according to tone.

For example, an unmarked form such as *Akümla* could be represented in different ways depending on its tonal pattern:

Unmarked form	Meaning	Tonal pattern	Representation in the earlier convention	Tone indicator
<i>Akümla</i>	work	Low on final syllable	<i>Akümlah</i>	<i>h</i>
<i>Akümla</i>	easy	High on final syllable	<i>Akümla</i>	consonant doubling (<i>ll</i>)
<i>Akümla</i>	melt	Low on first syllable	<i>Ahkümla</i>	<i>h</i>

This system illustrates an important stage in the development of Sütša orthography. Rather than representing tone through dedicated diacritical marks, tonal contrasts were encoded through additional consonantal symbols or consonant doubling. Such conventions were practical in contexts where diacritics were not readily available, but they could also make the written form appear structurally different from the underlying phonological representation of the language.

A similar issue can be observed in the representation of aspiration. At the time of Teo's work, the sequence *ch* was used in certain words to represent aspiration. For example, *Chhitoi* 'full' was written as *Tchitoi*. This convention appears to have been influenced by orthographic practices familiar from English and other European writing traditions. From the perspective of subsequent Sütša standardisation, however, the use of *tch* was considered unnecessary and potentially inconsistent with the emerging internal principles of Sütša orthography.

It is therefore important to distinguish Teo's description of the orthography as it existed during his fieldwork from the orthographic conventions subsequently adopted by the Sümi community. Teo's work

should be understood within the documentary context available to him at the time; it should not necessarily be treated as a representation of the orthographic standard currently adopted by the Sümi community.

A significant development took place in 2019, when a Sümi General Meeting resolved to introduce a more systematic use of diacritics for the representation of tone. Under the revised convention, the grave accent (à) is used to indicate low tone, the acute accent (á) to indicate high tone, while the mid tone remains unmarked. This represents a movement away from the earlier consonantal method of marking tone and toward a system in which tonal information is represented directly on the vowel.

The same standardisation process also addressed the representation of aspiration. The earlier convention was discontinued, and was adopted for representing the aspirated sound occurring in words beginning with *c*. Thus, the later standardisation sought to bring the orthography closer to the phonological structure of Sütsa while reducing conventions that were perceived as externally influenced or unnecessarily complicated.

The difference between the earlier and subsequently standardised systems can be summarised as follows:

Feature	Earlier convention reflected in Teo's documentation	2021 standardized convention
Low tone	<i>h</i> used as a tonal indicator	Grave accent (e.g., à)
High tone	Consonant doubling (<i>hh</i> , <i>kk</i> , <i>nn</i> , <i>ll</i> , etc.)	Acute accent (e.g., á)
Mid tone	Generally dependent on the absence of the above markers	Left unmarked
Aspiration after/with <i>c</i>	<i>tch</i>	<i>chh</i>
General principle	Tone encoded through consonantal additions/doubling	Tone represented directly through diacritics
Orthographic orientation	Reflects conventions in use during earlier documentation	Community-standardised system

Orthography is not a static representation of a language; it is a historically developing system shaped by linguistic analysis, practical usage, and community decision-making. Teo documented Sütsa according to the conventions and linguistic practices available during his fieldwork. The subsequent 2019 resolution represents a later stage in the community's attempt to standardise those conventions.

This distinction is particularly important when comparing the works of Hutton, Sreedhar, Teo, and contemporary Sütsa scholarship. Each researcher worked within a particular documentary and orthographic environment. Consequently, differences between their representations should not automatically be interpreted as differences in the language itself. Some differences may instead result

from changes in orthographic convention, differences in analytical methodology, or the gradual refinement of community-based linguistic knowledge.

The documentation of these successive stages is therefore important for future scholarship. When subsequent researchers examine Sütsa several decades from now, they will be able to distinguish the orthographic conventions used by Hutton and Sreedhar from those documented by Teo and from the standardised conventions adopted by the Sümi community in 2019. A future study may then compare these systems systematically, examining what has changed, what has remained stable, and whether particular earlier representations reflected genuine linguistic features or merely the orthographic conventions of their respective periods.

In this sense, the present standardisation itself becomes part of the linguistic history of Sütsa. Recording the transition from the earlier consonant-based tonal notation to the current diacritic-based system is therefore not merely a matter of spelling reform; it is also an act of linguistic documentation. It creates a historical record against which future scholars can evaluate the development of Sütsa orthography and assess the extent to which earlier descriptions corresponded to the language as understood and standardized by its speech community.

This also reinforces the central argument of the present study: Sütsa linguistic documentation should be read as a cumulative and evolving tradition rather than as a series of isolated works. The orthographic practices of one period become the reference point for the next, while later community intervention can confirm, modify, or correct earlier representations. In this continuing process, community scholars and institutions play an increasingly important role in ensuring that the written representation of Sütsa reflects not only earlier documentation but also the linguistic knowledge and preferences of its contemporary speech community.

From External Documentation to Community Linguistics

The history discussed above should not end with a list of errors. Its larger significance lies in the changing relationship between linguistic scholars and the Sümi community.

Both Hutton and Sreedhar depended upon community members. Hutton worked for an extended period in the Naga Hills and acquired considerable familiarity with Sütsa. Sreedhar's grammar explicitly identifies numerous Sümi speakers who assisted him. Later researchers such as Amos B. Teo likewise worked with Sümi community members and drew upon earlier Sümi publications. The history therefore demonstrates that community knowledge has always been present, even when the published authorship belonged to outsiders.

The difference today is that the community increasingly has institutional mechanisms through which it can participate in the production, evaluation and standardisation of linguistic knowledge.

This is the foundation of the concept of the community linguist used in this article. The term does not mean a person who possesses the same training as a professional linguist. It refers to a community member who has developed substantial linguistic competence through sustained engagement with the language and who participates in its documentation, teaching, publication, standardisation or institutional development.

Community linguists can perform functions that external linguists cannot easily perform. They can recognise whether an unfamiliar word is genuinely part of the language. They can identify whether a pronunciation belongs to a particular village or age group. They can recognise whether a grammatical form is natural in ordinary speech. They can understand the cultural meaning behind a word. They can determine whether a proposed spelling is likely to be accepted by teachers, writers, pastors, journalists and ordinary speakers. Most importantly, they remain in the language community after the external researcher has completed the field visit.

This does not make community linguists infallible. Native speakers can disagree, dialects can differ, memories can be imperfect and community institutions can themselves make mistakes. The proper conclusion is therefore not that community authority should replace academic authority, but that the two should be brought into sustained dialogue.

Community Linguistics and Language Identity

The relationship between language and identity is particularly important in the Sümi case. The change from “Sema” to “Sümi” and the emergence of “Sütsa” are not simply terminological developments. They are expressions of linguistic self-definition.

A community’s writing system is part of this process. Orthography determines how children first encounter their language in school. It determines how names are written, how newspapers are printed, how literature is published and how the language appears in digital communication. A spelling system can therefore become a visible representation of identity.

This explains why debates over Sütsa spelling can become emotionally charged. Some Sümi speakers believe that the forms established by early writers are inherently authoritative and that changing them amounts to distorting the language. This position deserves respect because the historical spelling system is connected to church literature, older books, hymns and the literacy experiences of several generations.

Nevertheless, historical continuity cannot by itself establish linguistic correctness. The early writing system was developed under circumstances very different from those of contemporary Sütsa. Much of the early documentation was produced when the language was being reduced to writing for the first time and when English orthography provided the main model available to external writers. Since then, Sütsa has expanded into education, administration, journalism, literature, religious publication, higher

education and digital communication. A written language that enters new domains inevitably requires institutional standardisation.

Standardisation should therefore not be understood as an attempt to erase history. It is better understood as a continuation of the historical process of making Sütša usable in new contexts. The pioneering writers established a foundation; the present generation has the responsibility to examine that foundation critically and refine it where necessary.

Standardisation as Continuation Rather Than Rejection

The debate surrounding the Sümi Literature Board's work on spelling can therefore be reframed. The issue should not be "old spelling versus new spelling". The more useful question is: What kind of writing system will enable Sütša to function effectively for present and future generations while respecting its historical record?

The present work suggests that three principles should guide the answer;

First, historical forms should be documented rather than erased. Hutton's spellings, Sreedhar's spellings and early missionary forms are valuable historical evidence. They should remain available to researchers because they reveal the history of the language and its written representation.

Secondly, historical forms should not automatically be imposed on contemporary writing. Where a spelling is demonstrably inconsistent with contemporary pronunciation or grammar, standardisation has a legitimate reason to revise it.

Thirdly, reforms should be systematic rather than personal. Orthographic decisions should be based on explicit principles that can be explained to teachers, writers, students and the wider community. A standard writing system cannot depend on the preferences of whichever committee happens to be in office at a particular moment.

This is where the institutional development of the Sümi Literature Board becomes significant. The preparation of the BA Sümi Honours curriculum under Nagaland University represents a transition in which Sütša is moving further into formal higher education. A language used at university level requires consistent terminology, spelling, grammar and pedagogical materials. The development of such resources makes systematic standardisation increasingly necessary.

A Comparative Reading of Hutton and Sreedhar

The contributions and limitations of Hutton and Sreedhar can be summarised comparatively:

Dimension	J. H. Hutton	M. V. Sreedhar	Community Linguistics
Historical phase	Colonial/ethnographic	Systematic linguistic	Contemporary institutional
Main works	<i>Rudimentary Grammar</i>	<i>Sema Phonetic Reader</i>	SLB, native scholars and

	(1916); <i>The Sema Nagas</i> (1921)	(1976); <i>A Sema Grammar</i> (1980)	contemporary research
Primary orientation	Ethnography and basic grammar	Phonetics, phonology and structural grammar	Standardisation, education and language vitality
Main strength	Early documentation and preservation	Systematic linguistic description	Native authority and institutional application
Treatment of tone	Macrons and <i>f</i> , inconsistently applied	More systematic tonal diacritics	Acute high, grave low, mid unmarked
Representation of <i>pf</i>	Requires separate phonological consideration	—	Dialectal and comparative investigation
Vowel representation	Strong English influence	More systematic but still contains discrepancies	Contemporary standardisation
Morphology	Important early observations but incomplete analysis	More systematic structural analysis	Native functional and grammatical interpretation
Community role	Informants and cultural guides	Informants and linguistic collaborators	Co-researchers, reviewers and institutional decision-makers
Major limitation	English-based representation and inconsistent analysis	Lexical, orthographic and vowel discrepancies	Requires continued technical linguistic training
Historical significance	Foundation	Systematisation	Consolidation and future development

The comparison demonstrates that the three phases should not be treated as mutually exclusive. Contemporary Sümi language development depends on all three. Hutton provides historical depth. Sreedhar provides systematic linguistic description. Community linguists provide continuity between documentation and living usage. The relationship is therefore cumulative rather than oppositional.

The Wider Significance of the Community-Linguist Perspective

The originality of the present approach lies less in claiming that Hutton or Sreedhar made errors than in asking what those errors mean for the future of an indigenous language.

Identifying a spelling error in an old grammar is not, by itself, a major theoretical contribution. What becomes significant is the recognition that published linguistic knowledge has consequences for later language development. Once an external description becomes authoritative, the community may begin to reproduce it even when it does not correspond to native usage. The publication itself becomes part of the language ecology.

This creates a form of documentary inheritance. Each generation inherits not only the language spoken by its elders but also the books produced by previous generations of scholars and writers. The inherited books may contain accurate descriptions, approximations, historical variants and errors. The community must therefore become capable of reading its linguistic inheritance critically.

This is where the community linguist has a distinctive role. A community linguist is not simply a native speaker who disagrees with a linguist. The more constructive role is to become sufficiently familiar with linguistic concepts to engage with professional scholarship and sufficiently grounded in the community to test that scholarship against actual language use.

The future of Sütsa requires precisely this combination. Community linguists must develop basic knowledge of phonetics, phonology, morphology, syntax, sociolinguistics and language planning. Professional linguists, in turn, should recognise that technical expertise does not eliminate the need for native linguistic authority. The strongest scholarship will emerge when both sides accept the limits of their own knowledge.

This is particularly important in the context of Indigenous language revitalisation. Contemporary scholarship increasingly argues that language documentation and revitalisation should not be separated. Documentation becomes more useful when it produces resources that communities can actually use for teaching and revitalisation. The Sümi case provides an opportunity to apply this principle to a language that is not merely being archived but is actively being standardised and incorporated into education.

The community-linguist model also provides a way of moving beyond the binary between “traditional knowledge” and “academic knowledge”. The two should not be treated as rival systems. Native speakers possess knowledge that is acquired through lifelong participation in the language, while academic linguistics provides analytical tools for describing that knowledge. A productive collaboration translates one kind of knowledge into the other without reducing either.

Institutional Safeguards: The Sümi Cultural Research Ethics Committee (SCREC)

To prevent future misrepresentation and ensure that research ethically serves the speech community, the Sümi Literature Board has established the Sümi Cultural Research Ethics Committee (SCREC). Comprising university professors, college educators, legal experts, and senior community scholars, SCREC operates in coordination with tribal administrative networks and village councils across Sümi-inhabited territories. Under current SCREC guidelines:

1. External researchers must submit research proposals, data collection tools, and elicitation lists for institutional review prior to field investigation.
2. Elicited data and field transcriptions must undergo vetting by designated community linguists to verify orthographic and semantic accuracy.
3. Published works must acknowledge native informants and deposit copies in community repositories for local educational use.

This institutional mechanism transforms the indigenous community from passive research subjects into active, sovereign partners in linguistic documentation.

Conclusion

The history of Sümi language documentation demonstrates that the development of a written language is a cumulative process involving many actors. Missionaries and Sümi writers created the earliest printed materials. Colonial administrators documented vocabulary, grammar and ethnography. Later professional linguists brought systematic phonetic, phonological and grammatical description. Contemporary Sümi scholars and institutions are now taking increasing responsibility for standardisation, education, research ethics and language development.

Within this history, the contributions of J. H. Hutton and M. V. Sreedhar remain indispensable. Hutton's *Rudimentary Grammar of the Sema Naga Language, with Vocabulary* (1916) was a pioneering attempt to describe Sümi grammar at a time when systematic linguistic documentation was extremely limited. His work preserved lexical, grammatical and cultural information from an earlier period. His recognition of tone, dialectal variation, elision and other features demonstrates that his linguistic observations were often more sophisticated than a superficial reading might suggest. His *The Sema Nagas* (1921) further preserved an extensive ethnographic record that remains valuable for historical research.

At the same time, Hutton's work contains limitations that must be acknowledged. His reliance on English orthographic conventions resulted in problematic representations of Sümi vowels and consonants. His representation of tone was inconsistent, particularly in his use of macrons and the letter *f*. The *pf* issue, however, needs to be considered separately. Unlike the tonal use of *f*, the occurrence of *pf* concerns consonantal representation. While *pf* is not a regular feature of common contemporary Sütsa, it occurs in the dialects of some Aphuyemi villages. Its presence in Hutton's writing therefore cannot simply be dismissed as an error without further investigation. At the same time, the possibility of influence from the Angami linguistic environment deserves consideration, particularly in view of Hutton's extensive work on the Angami and his authorship of *The Angami Nagas*. The available evidence, however, does not permit a definitive conclusion about the source of his *pf* representation.

Hutton's representation of forms such as *Enakha* and *ketami* likewise demonstrates the difficulty early documenters faced in hearing and representing Sümi vowel distinctions. His grammatical treatment of nominalisation through *Ke-* does not adequately represent the more complex normalising processes evident in contemporary Sütsa.

Sreedhar's work represents a major advance. His *Sema Phonetic Reader* (1976) brought a systematic treatment of speech organs, phonemes, sound production, phonetic drills and orthography. His *A Sema Grammar* (1980) extended the analysis to phonotactics, morphology, syntax and other grammatical structures. His attempt to establish a scientific orthography and to represent tone systematically was a significant contribution.

Yet Sreedhar's work also requires critical reading. The lexical discrepancies identified in the *Sema Phonetic Reader*, the questionable tonal assignments, the proposed *c/ch/chh* distinction and the vowel omissions in *A Sema Grammar* demonstrate that even systematic linguistic description can contain inaccuracies. The forms *Khuwasny* and *Wenkyé*, identified from the perspective of contemporary native Sütša as not being Sümi words, are particularly important warnings about the consequences of insufficient community verification. The point is not to diminish Sreedhar's achievement. It is precisely because his works are important that they must be read critically.

The central argument of this article is therefore one of appreciation through criticism. To criticise the pioneering works is not to dishonour their authors. On the contrary, serious scholarship honours important works by examining them carefully enough to understand both their contributions and their limitations.

The larger contribution of the article is the proposal to understand the history of Sümi documentation as a movement through three phases: colonial and missionary documentation, systematic linguistic documentation, and community-based institutional standardisation. The third phase does not reject the first two. It builds upon them while changing the relationship between the community and linguistic scholarship.

The community is no longer merely the source from which linguistic data are extracted. It is increasingly becoming an active participant in deciding how that data are interpreted, written, taught, archived and transmitted. The development of the Sümi Literature Board, the involvement of Sümi linguists, the preparation of university-level Sütša curriculum and the establishment of the Sümi Cultural Research Ethics Committee all point towards this institutional transformation.

This development is particularly important because language standardisation is sometimes misunderstood as an attempt to replace an established tradition. For some Sümi speakers, any change to the spelling system inherited from earlier writers can appear to be distortion. Such concerns should not be dismissed. Historical spellings are part of the community's literary inheritance, particularly through church literature, hymns, Bible translations and older publications. However, historical inheritance and contemporary standardisation are not mutually exclusive.

A written language must change as the domains in which it operates change. Sütša is now being used not only in informal communication and religious literature but also in journalism, publishing, education, examinations, digital communication and higher education. This expansion requires consistency. Standardisation should therefore be understood as a continuation of the work begun by the pioneers rather than a rejection of them.

The responsibility of the contemporary generation is consequently twofold: to preserve the historical record and to improve the contemporary system. Historical forms should be retained as evidence. Contemporary standards should be developed according to explicit linguistic and community principles.

The role of community linguists is central to this task. They are not substitutes for professional linguists, just as professional linguists are not substitutes for native speakers. Their responsibilities are complementary. Professional linguists can provide theoretical and technical expertise; community linguists can provide native linguistic competence, cultural understanding, historical memory and knowledge of actual community usage. The future of Sütsa will be strongest when these forms of expertise work together.

There is also an ethical dimension. External researchers working on Sütsa should recognise that a published linguistic description does not end its life when the research project ends. It may be used by teachers, students, writers and future researchers for decades. An inaccurate form can therefore acquire authority simply through publication. Community review is consequently not an obstacle to research; it is a means of improving research quality.

The Sümi experience also speaks to the wider field of Indigenous language scholarship. The movement towards community-controlled documentation and revitalisation reflects a broader reconsideration of who has the authority to describe an Indigenous language and for whom linguistic documentation is ultimately produced.

The question with which this article began can therefore be answered more clearly: What happens when an indigenous community reads the linguistic documentation of its language from within? It discovers both the value and the limitations of the inherited record. It recognises the labour of earlier scholars while identifying representations that require correction. It learns that linguistic knowledge has always been collaborative, even when the published authority was external. And, most importantly, it begins to assume responsibility for the future of its own language.

For Sütsa, this is not merely an academic exercise. It is part of the larger process of maintaining linguistic identity in a rapidly changing world. The challenge is not simply to produce more books about Sütsa. The challenge is to ensure that the books produced about Sütsa are accurate, that the books produced in Sütsa are consistent, and that the language remains a living medium through which future generations can know who they are and where they come from.

The work of Hutton and Sreedhar therefore belongs neither to a museum nor to an untouchable canon. It belongs to the history of Sütsa and should continue to be read, taught, questioned and used. Their work laid foundations. The task of the present generation is to build upon those foundations with greater linguistic precision, stronger community participation and clearer institutional responsibility. In that continuing process, the community linguist occupies a crucial position, not between the community and linguistics, but within the community and in dialogue with linguistics.

REFERENCES:

1. American Baptist Mission, Impur. (1925). *Ongko kaku atughuu* [Beginner's arithmetic in Sema Naga].

2. Bible Society of India. (2013). *Baibel [Sümi Bible]*.
3. Dickson, H. B., & Shohe, I. (1906). *Simi'tsa Kaku [The first reader in Sema Naga]*. Director of Public Information, Assam.
4. Grierson, G. A. (1903–1928). *Linguistic survey of India* (Vol. 3, Tibeto-Burman family). Office of the Superintendent of Government Printing, India.
5. Hinton, L. (2010). Language revitalization in North America and the new direction of linguistics. *Transforming Anthropology*, 18(1), 35–41. <https://doi.org/10.1111/j.1548-7466.2010.01068.x>
6. Hutton, J. H. (1916). *Rudimentary grammar of the Sema Naga language, with vocabulary*. Assam Secretariat Printing Office.
7. Hutton, J. H. (1921a). *The Angami Nagas*. Macmillan and Co.
8. Hutton, J. H. (1921b). *The Sema Nagas*. Macmillan and Co.
9. Hutton, J. H., & Bor, N. L. (1923). The use of tones in Sema Naga. *Journal of the Royal Asiatic Society*, 55(1), 103–109.
10. Kibami, A. (1961). *Sumi tsaithi [English-Sema dictionary]*.
11. Kinimi, N. (1927). *Ateghiu akikihui Simi kaku [Arithmetic in Sema Naga]*. Subdivisional Officer, Mokokchung.
12. Leonard, W. Y. (2017). Producing language reclamation by decolonising “language”. *Language Documentation and Description*, 14, 151–169. <https://doi.org/10.25894/ldd146>
13. Sreedhar, M. V. (1976). *Sema phonetic reader* (CIIL Phonetic Reader Series, No. 15). Central Institute of Indian Languages.
14. Sreedhar, M. V. (1980). *A Sema grammar* (CIIL Grammar Series, No. 7). Central Institute of Indian Languages.
15. Swu, S., & Yephthomi, H. (2004). *Anglo-Sümi Tsashe (Anglo-Sumi dictionary)* (2nd ed.).
16. Teo, A. B. (2012). Sumi (Sema). *Journal of the International Phonetic Association*, 42(3), 365–373. <https://doi.org/10.1017/S0025100312000254>
17. Teo, A. B. (2014). *A phonological and phonetic description of Sumi, a Tibeto-Burman language of Nagaland*. Asia-Pacific Linguistics.