elpublishing

Language Documentation and Description

ISSN 1740-6234

This article appears in: *Language Documentation and Description, vol 18*. Editors: Candide Simard, Sarah M. Dopierala & E. Marie Thaut

Introducing the Sylheti language and its speakers, and the SOAS Sylheti project

CANDIDE SIMARD, SARAH M. DOPIERALA & E. MARIE THAUT

Cite this article: Simard, Candide, Sarah M. Dopierala & E. Marie Thaut. 2020. Introducing the Sylheti language and its speakers, and the SOAS Sylheti project. In Candide Simard, Sarah M. Dopierala & E. Marie Thaut (eds.) Language Documentation and Description 18, 1-22. London: EL Publishing.

Link to this article: http://www.elpublishing.org/PID/196

This electronic version first published: August 2020

Introducing the Sylheti language and its speakers, and the SOAS Sylheti project

Candide Simard
The University of the South Pacific

Sarah M. Dopierala
Goethe University Frankfurt

E. Marie Thaut
SOAS, University of London

Abstract

Sylheti is a minoritised, politically unrecognised, and understudied Eastern Indo-Aryan language with approximately 11 million speakers worldwide, with high speaker concentrations in the Surma and Barak river basins in north-eastern Bangladesh and south Assam, India, and in several diasporic communities around the world (especially UK, USA, and Middle East). This paper briefly describes the Sylheti language from a variety of historical, socio-cultural, political, and linguistic angles, with a focus on the context of the Sylheti spoken among diaspora speakers living in London, which is the home base of the SOAS Sylheti Project (SSP), also introduced here, that helped facilitate some of the research presented in the papers comprising this volume.

Simard, Candide, Sarah M. Dopierala & E. Marie Thaut. 2020. Introducing the Sylheti language and its speakers, and the SOAS Sylheti project. *Language Documentation and Description* 18, 1-22.

1. Introduction

The Sylheti language counts around 11 million speakers, mostly living in the Surma river basin that is today largely within the Sylhet Division in north-eastern Bangladesh, and in the Barak river basin in the neighbouring Indian states of Assam, Tripura, and Manipur. Historically, undivided Greater Sylhet (geographically composed of both Surma and Barak river basins) was part of the region of Assam in British colonial India, but after the creation of independent India and Pakistan following the Partition of 1947, Barak Valley remained in India, and Surma Valley became part of East Pakistan. In 1971, East Pakistan, with a dominant population of mostly ethnic Bengalis, far removed geographically and culturally from officially Urdu-speaking West Pakistan, seceded to form Bangladesh. Today's political borders do not correspond to historical, cultural, and linguistic community borders.

Official speaker data is lacking because Sylheti is not politically recognised as a language in Bangladesh, or in India. Reports by Sylheti speakers themselves indicate that it is spoken by a significant portion of the population in a large area that ranges from Jiribam in the Indian state of Manipur in the east, through the three Barak Valley districts Cachar, Hailakandi, and Karimganj in the state of Assam, India, and in the west in the four Surma Valley districts Sylhet, Moulvibazar, (eastern) Sunamgonj, and (north-eastern) Hobigonj in Bangladesh, and to the south in (northern) North Tripura, (northern) Unakoti, and (northern) Dhalai, in the Indian state of Tripura. A significant migrant community of Sylheti speakers is also reported in the area of Hojai, Assam, India, dating back to before Partition. Sylheti is not spoken uniformly in this linguistically diverse area. Figure 1 shows the partitioned political areas where the Sylheti language is reported to be most spoken.

Figure 1. Sylheti language map

Today, Sylheti is also spoken in several diasporic communities in South Asia and worldwide. It is, for example, the language spoken by most of those in the UK who identify as having a Bangladeshi origin (see Section 1.2).

1.1. The Sylheti language

Sylheti is an Eastern Indo-Aryan language, classified as a member of the Bengali-Assamese continuum (Lewis et al. 2014). It is not officially recognised in Bangladesh, where it is simply referred to as a dialect of Bengali[1] by the government (Faquire 2012); it has, equally, no legal status in India. Sylheti is often dismissed as 'slang' or as a corrupted version of Bengali, even by some of its own speakers, for whom it is not a language in its own right. In the linguistic landscape of Bangladesh, standard Bengali is the only official national language, which continues to have an important political dimension. It has played a significant role in the events that led to the founding of Bangladesh as an independent nation in 1971. In 1952, in what was then Pakistan, students demanding equal status for Bengali as a national language rioted in the streets of Dhaka and several were killed. They are now known as 'the language martyrs', remembered every year on the day of *Ekushay* on 21st February (see Schendel 2013, inter alia). Today, Bangladesh does not recognise any of the more than forty minority languages spoken within its borders, and language communities continue to struggle for recognition. In India, there were also deadly protests in 1961 in Silchar in Assam's Cachar district[2] against the imposition of Assamese as the state's only official language. Standard Bengali, one of the 22 recognised scheduled languages of India, was ultimately established as the district's official language, but Sylheti did not get any recognition.

Sylheti-speaking areas of Bangladesh and India are characterised by diglossia, where standard Bengali is the language of education and literacy and Sylheti is the vernacular variety used in everyday interactions. Although Bengali and Sylheti are closely related (as are both to Assamese), the academic consensus on mutual intelligibility between Sylheti and Bengali

[1] We use 'Bengali' to refer to the language spoken in Bangladesh and India by over 200,265 million speakers (6th or 7th largest in the world, see https://www.ethnologue.com/guides/ethnologue200), and to denote the ethnicity of its speakers. We are aware that the endonym (native name) 'Bangla' is now sometimes preferred for the language and 'Bengali' for the ethnicity.

[2] See https://bit.ly/2XKTVMA (accessed 2020-07-19)

ranges from 'unintelligible' to 'hardly intelligible' (Chalmers 1996). The claim of mutual intelligibility by some speakers of both Sylheti and Bengali may be more an effect of the speakers' exposure to both languages; speakers of Sylheti who have never learned Bengali often report that they do not understand it to any functional degree.

Sylheti, politically unrecognised and minoritised in modern history, has lower prestige than Bengali, and many of its speakers only openly admit to speaking Bengali, or Sylheti-Bengali. There is reported language shift in the Sylheti-speaking regions of Bangladesh and India, as well as in the diaspora with Bengali replacing Sylheti, as some parents do not speak Sylheti to their children, reducing the number of future Sylheti speakers. In Bangladesh and Assam Sylheti endonyms have practically been replaced by place names in Bengali. Many younger speakers do not recognise [silɔt] 'Sylhet', and only know the Bengali term [silɛt], with some claiming, ironically, that [silɔt] must be a foreign creation. There are, however, factors that can contribute to raising the status of a language variety, and two such factors are relevant for Sylheti. Firstly, it is used as a marker of identity. There is a strong regional identity evident in Greater Sylhet, in Bangladesh, and in Assam, India (which is often concomitant with a national Bangladeshi or Indian identity). This shared identity is similarly evident in the diaspora (Bhattacharjee 2012). The second factor is the presence of a written standard. In the South Asian context, the significance of a recognised script to enhance the status of a language is particularly understandable, given the noted linkages between the existence and use of a script and a literary tradition, classical language, recorded history, cultural authenticity, and power (Brandt 2014). Sylheti has its own script, Sylheti Nagri (see Figure 2). Although the scholarship on Sylheti Nagri is nowadays limited, Constable *et al* (2002: 4) attest that it is not directly derived from the Eastern Nagri[3] script, it is rather 'a member of the group which includes all the major scripts in use today in northern India (other than Urdu which uses a modified Arabic script), such as Gurmukhi (Punjabi), Gujarati, Devanagari, Bengali and Oriya'.

[3] The Eastern Nagri script was first created to write Sanskrit and later adopted by regional languages like Bengali and Assamese. The Bengali Unicode block of characters is created from the Eastern Nagri script and contains character variants, like for the 'r', that is different in Bengali and Assamese.

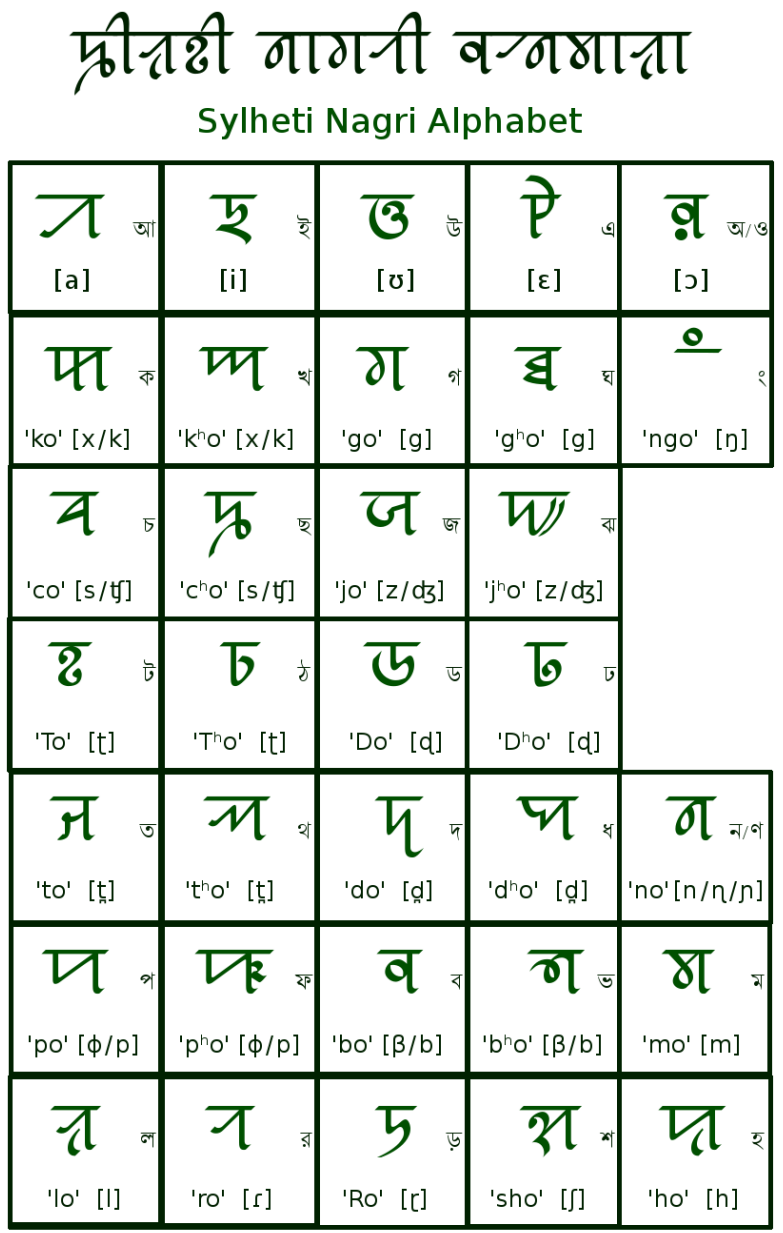

Figure 2. Sylheti Nagri alphabet

Sylheti Nagri has a distinct literary tradition; culturally notable are the *puthi-poṛa* (where *puthi* means 'book, manuscript' and *poṛa* 'to read, recite'). They are in effect musical performances (Kane 2017) where, traditionally, the performer, surrounded by their family or by listeners in the village, recited *puthi* for amusement and education in genres ranging from poetic fairy tales to romantic and religious stories. According to Kane (2017), from the 16th century onwards, the Sylheti *puthi* played an important role in communicating Islamic ideals to the people in the Sylhet region, as they were written in the spoken vernacular language. In the 1970s, the use of Sylheti Nagri was discouraged by the new Bangladesh government, and it fell out of use when its remaining printing presses were destroyed after 1971. Its decline can be attributed, firstly, to nationalist ideas originating in the colonial era of only politically recognising languages that have a widely-used written form, and, secondly, to the resulting widespread literacy in standard Bengali throughout Bangladesh, and in Barak Valley in Assam, India.

The cultural importance of the Sylheti Nagri script is nonetheless recognised by scholars, for instance, in the British Library Endangered Archives Programme project *Archiving Sylheti Nagri Texts*[4] led by Professor Anuradha Chanda of Jadavpur University (India) that has preserved 103 handwritten and printed texts in the Sylheti Nagri script from South Assam. Moreover, Sylheti Nagri is undergoing a revival today in Greater Sylhet and the diaspora, bringing with it great symbolic significance. For example, it was featured in a 2017 exhibition and workshop held in Tower Hamlets (London, UK) titled *Bangla is not my mother tongue*[5] by the visual artist and spatial designer Osmani Saif.

Eisenlohr (2004) argues that the presence of a language in new technologies can lead to it being better appreciated, by establishing a positive association with modernity and relevance to current lifestyles. Internet connectivity is still limited for many around the world, but mobile devices are allowing more and more people, especially those in minority/minoritised communities, to access information and take part in knowledge creation (see Sylheti Wikipedia incubator, Thaut 2019, 2020). Sylheti speakers of a certain age, those unfamiliar with the existence of the Sylheti Nagri script and its literary history, often express the opinion that Sylheti cannot be considered a language because it is not written. Younger speakers are not hampered by such considerations and have started communicating in Sylheti digitally, through text messaging and on social media platforms, using the scripts of the

[4] https://doi.org/10.15130/EAP071 (accessed 2020-06-19)

[5] http://www.saifosmani.com/bangla-is-not-my-mother-tongue/4593497983 (accessed 2020-06-19)

languages in which they are literate. Unicode characters for the Sylheti Nagri script have been created,[6] and this has helped to increase the online presence of Sylheti. Unicode characters are a first step, and the creation of fonts is the next, so that the Unicode characters can be read ('rendered') in text processing software and internet browsers, for instance. A final development is the creation of dedicated keyboards to make fonts easily accessible to users; keyboards for the Sylheti Nagri script are available for Android devices, and the first of its kind for Apple mobile devices was created in 2020.

The development of Sylheti Nagri Unicode continues, and reflects its speakers' sensitivities. As is the case in many other language communities trying to revitalise or support their language, issues of standardisation and representativeness arise. Lately, discontent has been expressed in the Sylheti community over the inclusion of Eastern Nagri (Bengali) script numerals instead of distinct Sylheti Nagri numerals in the original Sylheti Nagri Unicode proposal. In response, some Sylheti speakers have created their own fonts with the numerals stylised to look like Sylheti Nagri numerals, however this is only a temporary fix as devices without the font installed revert to displaying the numerals in the Eastern Nagri (Bengali) script. Some Sylheti speakers have also searched for similar-looking characters from other scripts to create a list of symbols for the numerals that can be used when typing Sylheti Nagri. While there may not be a consensus on the exact forms that the Sylheti Nagri numerals should have, which is not an uncommon situation,[7] a further step yet to be taken is updating the Sylheti Nagri Unicode proposal to include Sylheti Nagri numerals to better reflect Sylheti speakers' desired and actual usage and giving the Sylheti community the option of non-Eastern Nagri (Bengali) numerals.

The socio-cultural and political issues related to the status of Sylheti are complex and have already received some attention (see Lawson & Sachdev 2004; Zeitlyn 2008, inter alia). The papers presented in this volume highlight some of the striking structural differences between Sylheti and standard Bengali, in phonetics and phonology, lexicon, and grammatical structure, and challenge the view that Sylheti is merely a dialectal variation of Bengali.

[6] See https://bit.ly/2PYkHgf (accessed 2020-08-05)

[7] Even the Eastern Nagri had variations for certain numerals that were weeded out when one form was chosen for the Bengali script's Unicode characters.

1.2. Sylheti speakers in the UK and in London

According to the 2011 UK Census[8] the self-identified Bangladeshi community in the UK counted 451,529 people, of which just under half (49.2%) lived in the Greater London area. The 'heartland' of the London Bangladeshi community is to be found in the borough of Tower Hamlets where 22.8% of the UK Bangladeshi population live, with concentrations in the inner boroughs of Westminster, Islington, Camden, Hackney, Newham, and Southwark. Although predominantly Sylheti Muslim Bangladeshis, the Bengali community in London also comprises Indian Hindu Bengalis and non-Sylheti Muslim Bangladeshi Bengalis. These distinctions are important to the people themselves, but are mostly lost to the national and local governments in the dominant discourse in the British media which tends to refer to the whole Bangladeshi community, irrespective of language and religion, as 'Bengali' (Mookherjee 2014: 141-142).

The history of migration and formation of the UK Bengali community must be understood in the context of the British colonial past. Around 95% of the Bangladeshi community in the UK comes from the agricultural Surma Valley region of Sylhet (which formed a part of colonial Assam before Partition in 1947, see Section 1), where farmers could own small parcels of land, as opposed to being tenants on the large estates found elsewhere in East Bengal (Gardner 1995). The region was also home to lascars, sailors often employed in the British navy, some of whom settled in the UK throughout the 18th and 19th centuries. After the Second World War these first-generation migrants helped their relatives obtain employment vouchers for work in factories and textile mills in the UK. This migration affected almost exclusively men, who later asked their wives or fiancées to join them. The first wave of migration was followed by a second that was restricted to family members and to married partners. This process of Sylheti chain migration has had impacts both in Bangladesh and in the UK. In the former, it resulted in a certain localised prosperity, thanks to the remittances received from those settled abroad. In the latter, it contributed to the arrival of immigrants originating from the same *thanas* or administrative subdistricts, today called *upazillas* (Gardner 2002). For example, in the borough of Tower Hamlets, people who come from BeaniBazaar in Sylhet tend to live around Whitechapel and Shadwell, while families from Jagonnathpur, Sunamganj, and Bishwannath in Sylhet tend to live around Brick Lane. In Camden, many families tend to originate from MoulviBazar (Eade & Garbin 2006). This type of migration has had the effect of creating close-knit social networks within the London Sylheti community. Even today, socialisation and integration with

[8] https://bit.ly/3id1K5G (accessed 2020-06-19)

non-Sylheti people can be limited, leading to the continued existence of exclusively Sylheti social networks. It is not unusual to come across people who, after more than fifty years in the UK, have little or no fluency in English and rely on their children or grandchildren to communicate with non-Sylheti people in their neighbourhood, or with local or national institutions. Standard Bengali translations have limited usage and standard Bengali interpretation cause lack of understanding and mis-understandings (Comanaru & d'Ardenne 2018). Some UK institutions, like the National Health Service (NHS) and the Office of National Statistics (ONS), provide Sylheti interpretation and translation in order to functionally engage with Sylheti-speaking citizens who do not understand Bengali and have limited understanding of English.

Bangladeshis are said to be the most socially marginalised minority community in Great Britain (Gardner 2002; Eade & Garbin 2002; Platt 2007). According to Peach (2005: 23), they 'are poor, badly housed and poorly educated, suffer a high level of male unemployment and have a very low female participation rate in the labour market'. However, generational shifts are taking place. For instance, a study by Dench et al. (2006) found that third generation Bangladeshis achieve better education success, a trend that appears to be continuing today. In 2014 in the London borough of Camden, 69% of Bangladeshi girls and 52% of Bangladeshi boys got five or more A* on their GCSEs examinations (the highest qualification).[9] These results are higher than the overall scores for the borough – 68% for girls and 49% for boys – and belie the common assumption of low achievement or low aspirations among Bangladeshis.

The confused status of Sylheti as a language is also reflected in the UK immigrant communities. A study of the terms 'language' and 'heritage' with regard to language teaching and learning by Blackledge & Creese (2008: 535)[10] found diverging views from participants:

> all at first glance of the same 'ethnic' and 'linguistic' group, not only disagreed with each other about what constituted a 'language', they also disagreed with each other about where a 'language' began and ended, and about the value that could be assigned to a particular set of linguistic resources.

[9] General Certificate of Secondary Education, national examinations held at the end of high school in the UK.

[10] The study interviewed participants in Birmingham, England.

For many of the participants, speaking Sylheti rather than Bengali held negative associations, if not with regard to any linguistic differences in themselves, then in terms of social status and value, notably by being associated with the uneducated poor. This permeated the comments of all participants in the study, with one teacher quoted as saying: 'I talk posh Bengali, and the children can't understand me' (Blackledge & Creese 2008: 544). Although this research was not reproduced in Camden within the context of the SSP, ample anecdotal evidence corroborating these observations comes from interactions between SSP members and staff at the university with community members. In a shop on Tottenham Court Road (near SOAS in London), a young salesman of Bangladeshi background was thrilled to hear about the SSP, but then suggested we should really concentrate on Bengali, as Sylheti was not a 'very good language to study for a university'.

Sylheti speakers are not a monolithic block; there is variation between generations, for example, where older Sylheti speakers may be monolingual and younger speakers have only passive competence. There may be differences in fluency between men and women, with some of the latter having recently arrived through marriage and thus being more fluent compared to longer-established residents. The younger generations in Camden (London) display wide-ranging linguistic repertoires, as reported similarly by Blackledge & Creese (2008: 548) for their counterparts in Birmingham, who:

> [i]n addition to making use of linguistic resources of English, Sylheti, and Bengali, … watched Hindi films, read the Qur'an in Arabic, and listened to popular contemporary music in varieties of American English, and also Indian and Bengali pop music.

In the borough of Tower Hamlets in London, Rasinger (2007) noted the emergence of a 'Sylheti-Cockney' dialect among the youth of the second and third generation, composed of both East London adolescent vernacular and Sylheti elements. In both the Birmingham and Tower Hamlets studies, young people expressed multiple complex identities that intermingle 'home' as the UK with a strong sense of belonging to the 'Bengali' community.

Finally, the status of heritage languages in the UK, i.e. those languages spoken by immigrants or their children, does not simplify the situation for Sylheti speakers. The UK remains linguistically inward-looking, with an ideology that minority languages are a negative force in society that prevails in British political discourse and in the media (Blackledge & Creese 2008). The relationship between language and heritage is far from straightforward (Blommaert 2005), as demonstrated in the activities of the SOAS Sylheti Project, described below.

2. SOAS Sylheti Project

Since 2003 the Linguistics Department at the School of Oriental and African Studies (SOAS, University of London) has conducted postgraduate teaching and research on the theory and practice of language documentation and description, and language support and revitalisation, with the stated goal of developing the skills of those currently working on endangered languages or on minority/minoritised languages and training the next generation of language documenters.[11] The Department has a specialist MA in Language Documentation and Description whose core modules include Introduction to the Study of Language, Applied Language Documentation and Description, Language Support and Revitalisation, Field Methods, and Descriptive Linguistics.

In 2011 the Department started to collaborate with the *Surma Community Centre*, first set up in 1976 in neighbouring Camden by a small group of community activists as the base for the Bengali Workers' Association (BWA).[12] It initially provided a voluntary advice service on issues such as immigration, accommodation, and welfare for newly arrived migrants from Bangladesh. It is now a venue for a wide range of social, cultural, and community events, and is a focal point for the community in Camden. In spring 2011, Dr Mukid Choudhury, then-director of the *Surma Community Centre*, contacted SOAS linguists and explained that community members had expressed concern that the Sylheti language was being used less and less by younger generations, particularly those born in the UK, with the community undergoing a shift from being multilingual to English monolingual. For older Sylheti community members, the preservation of their identity, traditions, culture, and language is a priority, but younger people are under pressure, negotiating multilingual and multicultural identities. Such language shifts and concerns about the need to maintain heritage languages in diaspora communities is not unusual (Pauwels 2019). Dr Choudhury also mentioned the feeling of unease experienced by many younger community members when the variety of Sylheti they speak is deemed 'quaint' when they travel to the homeland because of the ways the language in Sylhet (Bangladesh) has changed, shifting more towards standard Bengali over the last 50 years. Older speakers find themselves disappointed that their efforts to transmit Sylheti to their children could be having a negative effect in the long term because of their children's negative experiences with language use during visits to Bangladesh. On the other hand, this diverging language

[11] https://www.soas.ac.uk/linguistics/programmes/malangdocdesc/ (accessed 2020-06-19)

[12] http://bwa-surma.org/about-us/ (accessed 2020-06-19)

evolution is not solely perceived as negative, as older speakers in Camden remember and use lexical items, idiomatic phrases, and constructions that are no longer in use in Bangladesh, making them precious keepers of linguistic and cultural knowledge.

The SOAS Sylheti Project (SSP) was set up in 2012 with the aim of documenting and describing the Sylheti language spoken in Camden, paying special attention to the usage of the elderly members of the community. The initial founders were Mike Franjieh (then lecturer in Field Methods), Tom Castle (then technician, SOAS Endangered Languages Archive), and Candide Simard (then post-doctoral fellow), with the support of the then Endangered Languages Academic Programme (ELAP), the Endangered Languages Archive (ELAR), and the SOAS Linguistics Department. While it was created as an extra-curricular and voluntary activity, it endeavoured to integrate its activities with the Field Methods, Applied Language Documentation and Description, and Language Revitalisation modules of the MA Language Documentation and Description programme. Participation was not restricted to MA students but was open to all those interested, including undergraduates. As it sits at the intersection of the students at SOAS, the surrounding metropolitan communities of Sylheti speakers, and those involved in developing and teaching the theory and practice of language documentation and description, the SSP has created potential spaces to experiment with pedagogical practices and approaches. It has been reflective at the same time, providing an opportunity to discuss ideas and perspectives outside the limits imposed by formal classes.

The SSP is a positive response to the institutional mission of UK universities to find effective ways of developing, transmitting, and applying knowledge through community engagement so that it benefits all. According to Laing (2016), community engagement should aim:

> to address social disadvantage and exclusion, to promote the idea of a fair society and should complement and collaborate with the university's service to business activities by focusing on all those areas of our daily lives that are of profound material and civic importance but which are typically seen as 'non-economically productive activity', such as caring, sustainable development, self-management of health and well-being, voluntary activity and the development of citizenship.

This is a challenge for many universities, particularly in a socio-economic context that is increasingly competitive and self-interested.

Within the discipline of linguistics, the SSP has challenged the non-realistic divisions between language documentation and description on the one hand, and language revitalisation or maintenance on the other. The models that have evolved over the last 25 years have resulted in a separation of these

areas that in practice is neither practical nor desirable (see Austin 2016; Austin & Sallabank 2018). Revised frameworks for language documentation have been suggested which argue for more participatory and politicised linguistics, making a place for activism. Penfield et al. (2007) (cited in Penfield 2018: 800) define a language activist as 'a person who focuses energetic action towards preserving and promoting linguistic diversity/supporting language rights'. Adopting such a perspective has direct implications: the power position of the researcher/linguist is altered, and the lines between linguist/researcher and community members are blurred, with projects becoming more clearly team efforts.

The SSP recognises that knowledge exists both at the university, including what the students themselves bring with them to the university, since SOAS students typically have rich and varied backgrounds of skills and experience, and outside the university in the community. Through their participation, lecturers, students, and community members are all contributing to the creation of new knowledge; such a perspective promotes ownership by all those involved. The SSP's major contributions have been:

1. to challenge the traditional exclusionary relationships between research, education, and the 'real world';
2. to enhance students' experiences by giving them training in authentic aspects of research, including community-based fieldwork and analysis; and
3. to demystify the role of universities in UK society by encouraging Sylheti speakers to get directly involved in research activities, thus breaking down the walls between academia and the 'outside world' and raising awareness of the work that linguists do and its relevance to the lives of community members.

While not attempting to strictly apply the model, the SSP is underpinned by the principles of constructivist pedagogy, which values learning that is initiated and directed by students. In the SSP, students build on current and past knowledge, applying it in a practical real-world context that is defined jointly, together with the community members at the *Surma Community Centre*. SOAS academic staff assist students in their discovery, but do not set the programme. These ideas stem from a long pedagogical tradition, going back to Dewey (1938), who coined the phrase 'learning by doing', where students needed to integrate skills and knowledge into their lives through experience, rather than being taught disconnected facts. The SSP is also informed by 'critical pedagogy', traditionally defined as teaching/learning practices designed to raise learners' critical consciousness regarding

oppressive social conditions (Freire 1970) by using meaningful dialogue to reconfigure the traditional teacher-student relationship.

Students are attracted to the SSP primarily through the opportunity to participate in a project that has practical relevance, where they develop transferable analytic, organisational, and communicative skills. Over the years, different sub-projects have been developed, some of which have reached completion, and others which have remained undelivered, often for a lack of funding to compensate consultants or because of the time-intensive work they demanded, given the otherwise heavy workload of students. In the SSP, a sub-project that does not reach completion is not a 'failed' project as long as all those involved feel they have learned something along the way.

As many as 50 students have participated in the SSP since its founding. It has yielded 37 hours of audio recordings relating to lexical and dictionary work, thousands of token recordings of words and example sentences, and many community-directed outcomes. Some initial efforts in 2012-2013 proved to be over-ambitious. For example, the *Tasty Tales* project aimed to video-record a cooking demonstration of a recipe that parents felt a young person leaving home for the first time would find useful. Resources ran short to cover the time-intensive work of translating what was being said in Sylheti, editing the video segments, and including subtitles in the edited outcomes. Subsequent initiatives started less ambitiously, and incorporated ways to address funding issues from the beginning. Examples of these include creating and carrying out a questionnaire about intergenerational communication, exploring teenage speakers' attitudes and knowledge of Sylheti. Another was the creation of animated videos with subtitles for some of the annotated stories collected during the Field Methods classes from academic years 2012-2013 to 2014-2015, made available via the SOAS Sylheti Project YouTube channel.[13]

An offshoot of the SSP is the *SOAS Sylheti Language Society*, created in 2013 to engage with non-linguistics students at SOAS, especially those with Sylheti heritage who are interested in learning the language. It also provides opportunities to teach the language, and has inspired the creation of tailor-made language-teaching materials, as existing resources are rare and difficult to obtain (see below). These are freely made available via the SOAS Sylheti Language Society Facebook page.[14]

The SSP has also had a long-standing dictionary project that has included lexical data collection through Field Methods classes and at the

[13] https://www.youtube.com/user/soassylhetiproject (accessed 2020-06-19)

[14] https://www.facebook.com/soassylhetilanguagesociety/ (accessed 2020-06-19)

Surma Community Centre's elderly Women's and Men's groups in 2013-2014, and 2014-2015. A database of more than 4,500 entries was created (growing to almost 14,000 entries in 2020) that led to the development of a first version of a Sylheti Dictionary Android mobile phone app,[15] sponsored by the SOAS Alumni & Friends Fund. The app was launched at a Sylheti Cultural Celebration public event in June 2015, along with the SSP's first lessons booklet, the result of collaborative teaching and learning by SOAS students (Sylheti-speaking and non-Sylheti) and Sylheti-speaking community members, supported by funds from the SOAS Students Union.

In 2015-2016, students worked on a storybook, *The Boy Who Cried Tiger, and Two Other Stories as Told in the Sylheti Language* (fuaTae sillaito bag aise aro duiTa kicca siloTi bashae) in hardcopy and in multiple digital versions.[16] The storybook is a good example of an outcome from the cumulative community-oriented activities of the SSP. It began with the goal of developing a high-quality Sylheti-language storybook, as young mothers remarked they did not have any books in Sylheti to read to their children. It came to fruition because of the students' professional experience in book publishing, project management, language teaching, and fund-raising.[17] It also involved sourcing and selecting submissions from professional illustrators, and tenders from publishing houses. The storybook's contents were selected from materials that had been previously recorded for language documentation purposes. The first story *The Boy Who Cried Tiger* (fueTae sillaito bag aise) had already been developed with Faruk Miah and presented as a short animated video, available on YouTube.[18] The third story *The Wind and the Sun* (boiar ar shuruz) was quickly elicited and recorded by Farhana Ferdous, then illustrated by the participants at the SSP's Cultural Celebration in 2016. It was also made available online as a short animated video.[19] The printed storybook became available in the autumn of 2017 and was freely distributed to the crowdfunding supporters, local community members, and local libraries (including the British Library). E-book versions are also freely available.[20]

---

[15] https://bit.ly/30LfIWj (accessed 2020-06-19)

[16] https://sylhetiproject.wordpress.com/storybook/ (accessed 2020-06-19)

[17] Funds for the storybook were raised through a crowdfunding campaign on the Internet – see https://soas.hubbub.net/p/sylhetistorybook/ (accessed 2020-06-19)

[18] https://www.youtube.com/watch?v=eCtMbiKgUMM (accessed 2020-06-19)

[19] https://www.youtube.com/watch?v=Zm0AbZCIzEk (accessed 2020-06-19)

[20] https://sylhetiproject.wordpress.com/ebook/ (accessed 2020-06-19)

The process of transposing these orally told stories into writing posed many challenges. For instance, there are multiple script and spelling choices available: Roman transcription for UK readers, regional Eastern Nagri, or Bengali-Assamese script transcription, for readers familiar with Bengali, and the revitalised historical script of Sylheti Nagri. Script and spelling selection required both an evaluation of the functional load of phonological features and the cognitive needs of speakers in the UK, many of whom range from partially literate to fluent in Bengali and English and able to read in both Bengali(-Assamese) and Roman scripts. The final decision to present parallel scripts in the storybook demonstrates the multiple ways different speakers can access written materials:

- Sylheti Nagri as an expression of a unique cultural heritage (although few potential readers would be literate in this script);
- Roman transcription as a gateway for heritage speakers and to facilitate learning the Sylheti Nagri script;
- Eastern Nagri transcription as a bridge for speakers literate in Bengali; and
- English translation for a London-based audience.

The decision was also made not to include a standard Bengali translation in order to underline the idea that Sylheti is distinct from Bengali. Since variation is inevitable in any language, and is perhaps more evident in one that is not usually written, the uniformisation of Sylheti's dialectal variation and grammar was also an issue. This is because the uniformisation needed to be done in a manner acceptable to most speakers, while respecting the originality of the two authors' tellings. Finally, issues of culture, identity, and authenticity needed to be taken into account, notably in the development of the illustrations.

As well as continuing with academic productions, the SSP is striving to maintain community involvement. In spring of 2018, lecturer Sheena Shah and the SOAS students in her Language Support and Revitalisation class visited the *Surma Community Centre*, talked to Sylheti language consultants Faruk, Farhana, and Nadia, and gained insights about what their part in the creation of the storybook meant to them. For them, it is a future heirloom when in 200-300 years Sylheti will have 'disappeared because of development'. It has inspired many conversations about life back in Sylhet and developed a greater awareness of code-mixing in their speech – working with linguists has helped them realise when they use Bengali instead of Sylheti, and it has encouraged them to seek and learn more Sylheti words from their parents and grandparents. Indirect gains of the storybook project

include an in-depth understanding of the importance of language documentation, and a contribution to the wellbeing of the community by instilling a sense of self-worth and confidence, as well as pride in their heritage and their native language.

Since its start, the SSP has also contributed to augmenting knowledge on Sylheti grammar, with an academic output that comprises:

- term papers on kinship terms (Bridle 2014), presented at the International Summer School in Language Documentation and Linguistic Diversity in Stockholm; and differential object marking (Laub 2014);
- MA dissertations on negation (Zambas), the additive particle (Brown), and converbs (Dopierala 2016);
- creation of a phonological inventory database (Eden);[21] and
- conference papers on aspects of morpho-syntax (Koumbarou & Thaut 2018), and documentation (Thaut 2019, 2020).

An unforeseen and very enriching addition to the linguistic documentation work of the SSP with the local diaspora community in London is the input and feedback received from participants on social media through the Facebook group created in 2013.[22] It now counts about 2,500 members, including speakers in other diaspora communities, in South Asia more broadly, as well as in the homeland of Sylhet, all participating virtually. By encouraging interdisciplinary collaboration, the SSP's online component has helped gain a greater picture of dialectal variation, worked on spelling conventions, and is having an increasing role in script revitalisation, as well as other linguistic and cultural revitalisation activities.

The SSP continues with community engagement, at the university and beyond. Here are a few examples of recent projects and activities. In 2017, the SSP was consulted in the development of a wellbeing survey by the UK Office of National Statistics (ONS). The project was engaged to compose a survey text in 'proper' Sylheti that elderly monolingual Sylheti speakers would understand and that could be presented in recorded audio and multi-script written forms.[23] In 2019, the SSP produced a Sylheti translation of the

[21] http://nidaba.co.uk/Contents/OriginalWordList/32 (accessed 2020-06-19)

[22] https://www.facebook.com/groups/sylhetiproject

[23] https://sylhetiproject.wordpress.com/ons-well-being-survey/ (accessed 2020-07-25)

InCommon group's intergenerational programmes material.[24] These translations were used when pupils in Tower Hamlets, London, visited Sylheti-speaking elderly women in a local retirement home. That same year at SOAS Languages Outreach Day,[25] the SSP led a group of GCSE pupils studying Bengali, most of whom were Sylheti(-heritage) speakers, in a discussion about Sylheti's place in the language plurality of 'Bengal', a region that has seen many diverse historical kingdoms and nations with various names. The SSP also presented a modern poem in the Sylheti language written in the Sylheti Nagri script, a surprise to the pupils who believed that Sylheti is 'slang' and cannot be written. Most recently, in spring 2020, the SOAS Sylheti Language Society's language lessons went online, and are now reaching a global group of learners, with editing continuing on the much expanded second edition of the Sylheti lessons book.

3. Conclusion

In this chapter we have all too briefly introduced the Sylheti language, the Sylheti community throughout the world with a focus on the diaspora in the UK, and the existing research on Sylheti language and use. We have endeavoured to include as many topical angles as possible. In the Introduction and Section 1.1 we described, among other things, Sylheti's linguistic classification as an Eastern Indo-Aryan language and its geographic distribution within and across current political borders, its history, and literary legacy. In Section 1.2 we explored the socio-cultural and political implications behind Sylheti's status (or lack thereof) as a language, and the complicated relationship Sylheti people, particularly second and third generation migrants to the UK, have with their linguistic heritage. In Section 2 we gave a detailed account of the creation of the SOAS Sylheti Project, its structure and goals, its activities past and present, and how these activities have and continue to positively contribute to both the academic and Sylheti-speaking communities.

[24] https://sylhetiproject.wordpress.com/projjects/incommon-sessions/ (accessed 2020-07-25)

[25] https://bit.ly/3fDDckE (accessed 2020-07-25)

References

Austin, Peter K. 2016. Language documentation 20 years on. In Martin Pütz & Luna Filipović (eds.) *Endangered Languages and Languages in Danger: Issues of ecology, policy and human rights*, 147-170. Amsterdam: John Benjamins.

Austin, Peter K. & Julia Sallabank. 2018. Language documentation and language revitalisation: some methodological considerations. In Leanne Hinton, Leena Huss & Gerard Roche (eds.) *Handbook of Language Revitalisation*, 207-215. London: Routledge.

Bhattacharjee, Nabanipa. 2012. We are with culture but without geography: locating Sylheti identity in contemporary India. *South Asian History and Culture* 3(2), 215-235.

Blackledge, Adrian & Angela Creese. 2008. Contesting 'Language' as 'Heritage': Negotiation of Identities in Late Modernity. *Applied Linguistics* 29(4), 533-554.

Blommaert, Jan. 2005. *Discourse: A Critical Introduction*. Cambridge: Cambridge University Press

Brandt, Carmen. 2014. Script as a potential demarcator and stabilizer of languages in South Asia. In Hugo C. Cardoso (ed.) *Language Endangerment and Preservation in South Asia*, 78-99. Honolulu: University of Hawai'i Press.

Bridle, Rose. 2014. A Typological Account of Sylheti Kinship Terminology in the UK Bangladeshi Diaspora. Paper presented at Young Linguists' International Conference, June 28, 2014, Stockholm, Sweden.

Comanaru, Ruxandra & Jo d'Ardenne. 2018. *Translation and cognitive testing of the harmonised personal well-being questions: Report on findings from the translation workshops and cognitive interviews for Sylheti and Urdu*. London: National Centre for Social Research (NatCen).

Chalmers, Rod. 1996. *Learning Sylheti*. London: Centre for Bangladeshi Studies.

Constable, Peter, James Lloyd-Williams, Sue Lloyd-Williams, Shamsul Islam Chowdhury, Asaddor Ali, Mohammed Sadique & Matiar Rhaman Chowdhury. (1 November 2002). 'Proposal for Encoding Syloti Nagri Script in the BMP'.

Dench, Geoff, Kate Gavron & Michael Young. 2006. *The New East End: kinship, race and conflict*. London: Profile Books.

Dewey, John. 1938. *Experience and education* (Vol. no. 10). New York: The Macmillan Company.

Eade, John & David Garbin. 2006. Competing visions of identity and space: Bangladeshi Muslims in Britain. *Contemporary South Asia* 15(2), 181-193.

Eade, John & David Garbin. 2002. Changing narratives of violence, struggle and resistance: Bangladeshis and the competition for resources in the global city. *Oxford Development Studies* 6(1), 137-149.

Eisenlohr. P. 2004. Language revitalization and new technologies: Cultures and electronic mediation and the refiguring of communities. *Annual Review of Anthropology* 18(3), 339-361.

Faquire, A. B. M. Razaul Karim. 2012. On the Classification of Varieties of Bangla Spoken in Bangladesh. *BUP Journal* 1(1), 130-139.

Freire, Paolo. 1970. *Pedagogy of the Oppressed.* New York: Herder & Herder.

Gardner, Katy. 1995. *Global Migrants, Local Lives: Travel and Transformation in Rural Bangladesh*. Oxford: Clarendon Press.

Gardner, Katy. 2002. *Age, Narrative and Migration: The Life Course and Life Histories of Bengali Elders in London*. London: Berg Publishers.

Kane, David. 2017. *Puthi-PoRa: 'Melodic Reading' and its Use in the Islamisation of Bengal*. London: Sylheti Research and Translation.

Koumbarou, Andriana & E. Marie Thaut. 2018. The marker -e in Sylheti: pragmatic, agentive and instrumental uses. Paper presented at 34th South Asian Languages Analysis Roundtable (SALA-34), Konstanz, Germany.

Laing, Stuart. 2016. Community engagement is what universities should be for. *Times Higher Education Supplement*. https://bit.ly/30wOUt2 (accessed 2020-06-19)

Lawson, Sarah & Itesh Sachdev. 2004. Identity, Language Use, and Attitudes: Some Sylheti-Bangladeshi Data from London, UK. *Journal of Language and Social Psychology* 23, 49-69.

Lewis, M. Paul, Gary F. Simons & Charles D. Fennig. (eds.) 2014. *Ethnologue: Languages of the World, Seventeenth edition.* Dallas, Texas: SIL International. http://www.ethnologue.com/17/ (accessed 2020-06-19)

Mookherjee, N. 2014. In pursuit of the 'authentic' Bengali: Impressions and observations of a Contested Diaspora. In M. Chakraborty (ed.) *Being Bengali: at home and in the world*, 140-159. Routledge: London.

Pauwels, Anne. 2019. Changing Perspectives on Language Maintenance and Shift in Transnational Settings: From Settlement to Mobility. In Gabrielle Hogan-Brun & Bernadette O'Rourke. (eds.) *The Palgrave Handbook of Minority Languages and Communities*, 235-256. London: Palgrave Macmillan.

Peach, Ceri. 2005. Muslims in the UK. In Tahir Abbas (ed.) *Muslim Britain: Communities under pressure*, 18-30. London: Zed Books.

Penfield, Susan. 2018. Funding the documentation and revitalization of endangered languages. In Kenneth L. Rehg & Lyle Campbell (eds.) *The Oxford Handbook of Endangered Languages*, 785-802. Oxford: Oxford University Press.

Penfield, Susan, Amelia Flores & Benjamin Tucker. 2007. The role of the language activist in documentation and revitalization. Paper presented at Conference on the Endangered Languages and Cultures of America (CELCNA 2007), Salt Lake City.

Platt, Lucinda. 2007. *Poverty and Ethnicity in the UK*. Bristol: The Policy Press.

Rasinger, Sebastian M. 2007. *Bengali-English in East London: A Study in Urban Multilingualism*. Bern: Verlag Peter Lang.

van Schendel, W. 2013. The Pakistan Experiment and the Language Issue. In W. van Schendel, & M. Guhathakurta (eds.), *The Bangladesh Reader: History, Culture, Politics*, 177-183. Durham, NC: Duke University Press.

Thaut, E. Marie. 2019. Language documentation and description with(in) a digital diaspora: The SOAS Sylheti Project. Paper presented at Digital Diasporas: Interdisciplinary Perspectives conference, School of Advanced Studies (SAS) University of London.

Thaut, E. Marie. 2020. Language documentation and description with(in) a digital diaspora: The Sylheti Project - SOAS in Camden. Paper presented at Disrupting Digital Monolingualism conference, King's College London.

Zeitlyn, Benjamin. 2008. *Challenging Language in the Diaspora. Bangla Journal* 6(14), 126-140.