

Shattered Dream of Purism: Sociolinguistic Dynamics of the Emerging Hybrid Social Media Language in Sri Lanka

Noel Dassanayake^a

^a*Sabaragamuwa University of Sri Lanka, Sri Lanka, noel@ssl.sab.ac.lk, <https://orcid.org/0000-0001-6041-5601>*

Abstract

The rapid digitization of communication in the last decade, especially in social networking spaces, has contributed to radical changes in the virtual language environment in Sri Lanka. It has not only reshaped the way individuals communicate, but has also fueled the emergence of new linguistic trends that reflect the dynamic nature of language in the contemporary virtual landscape. The present study seeks to explore the prospects of a pure language ideology in Sri Lanka with special reference to the use of Sinhala and English in the virtual space. The study used a corpus of memes extracted from social media and analyzed them from a sociolinguistic standpoint using Critical Discourse Analysis. The results show that social media have created a safe space for less competent language users to interact within their linguistic capacity. The adaptation of words and the invention of neologisms have occurred due to the unavailability of terms in users' L1 or L2, as well as to avoid taboos or prohibitions. Given the rapid evolution of communication technology, added to the dynamic changes that have already occurred within languages, aspirations to linguistic purism have become unrealistic optimism.

Keywords: *Linguistic Purism, Social Media, Language, Sri Lanka.*

Introduction

Linguistic purism has essentially been a controversial topic for over decades, and the new language attitudes brought about by rapid digitalization have made matters even more contentious between pro- and anti-purist factions around the globe. As claimed by Langer & Nesse, linguistic purism is related to the process of language standardization, the use of language as a building block in the creation of nations, and the stigmatization of linguistic varieties or cultures as undesirable or even a threat to one's identity (Langer, & Nesse, 2012, p. 607). Linguistic cleansing movements in a nation or region typically have their roots in the contact of local languages with other global or regional languages. As argued by Coperehewa (2011, p. 4), the relationship between language and nation in Sri Lanka became increasingly close in the early nineteenth century, and local elites pushed for defensive measures to preserve the purity of the local language against colonial languages. Apart from these ethnographic and xenophobic forms of purism, archaizing purism could be observed as the most dominant and challenged form of purism in the context of modern Sri Lanka. Since the archaic purist believes that the language of the past is fundamentally better, its modern modifications are despised as inferior (Ager, 1999, p. 92).

As Jernudd claims, voices of 'explicit Sinhala language purism' are still heard, although with much less intensity and power than a few decades ago (Jernudd, 1989, p. 08). The 'Hela Havula' movement founded by Cumaratuna Munidasa with its defensive stance against foreign impurities has found its way into modern Sri Lanka in more ethnographic and xenophobic approaches backed by nationalist political movements. Since language is a crucial tool for those in power and those who desire power (Joseph, 2006, p. 4), the purity of Sinhala, its symbolism, and its associations with other nationalist logos such as religion and culture have often been used for political purposes in Sri Lanka. The sociolinguistic dynamics of this web of circumstances have become more complex with the rapid popularization of social media in the decade since 2010, which has hardly been addressed in recent studies. The present study critically analyzes the radical shifts in social media language in Sri Lanka from a sociolinguistic standpoint, with special reference to the hybridization of Sinhala and

English terminology. Using a corpora of 1600 Facebook memes on YouTube, eventually filtered down to 431, the present study critically evaluates the dramatic shifts in social media language use and its sociolinguistic implications. This study examines the linguistic characteristics of Sri Lankan social media memes to explore how users navigate linguistic purism, hybridity, and evolving social norms. Using corpus and critical discourse analysis, the study systematically examines the intentional manipulation of language in social media that reflects broader sociolinguistic dynamics. This approach not only reveals how digital platforms are reshaping linguistic practices, but also addresses the lack of research on non-standard, hybrid language use in virtual contexts.

1. Review of Literature

1.1 Linguistic Purism and Language Change

Brunstad (2003, p. 52) identifies linguistic purism as "an ideology of language planning that includes resistance to foreign elements. In essence, however, purism is not limited to defending language only against incoming foreign elements, but also extends its defenses against changes within a language itself. Thus, from a directional perspective, purism could be labeled in two forms, internal and external, the former referring to impurities of native origin, such as neologisms and urban slang, and the latter referring to xenophobic impurities (Geers, 2005, p. 98). Linguistic purism interacts with many linguistic and extralinguistic factors, and purists are challenged with the arduous process of distinguishing between pure and impure content (Comrie & Thomas, 1994, p. 844). Since language itself is one of the most complex human tools, subject to arbitrary, unpredictable transformations and modifications, the purity or impurity of its content becomes a matter of controversy.

There are several forms of linguistic purism, such as archaizing, elitist, ethnographic, reformist, and xenophobic (Thomas, 1991). Archaizing purism refers to the idolization of early forms of language, such as the language of golden literary eras, and elitist purism is associated with more formal varieties of language, while ethnographic purism is associated with more vernacular forms of language, including folklore (Coronel-Molina, 2015, p. 29). Archaizing purists perceive the early forms of mother tongue with a reverent attitude, which they believe should not be challenged in any way. Reformist purism is associated with attempts to exclude earlier forms of language with the goal of reform and renewal, and xenophobic purism is usually associated with foreign elements such as loan and calque (Thomas, 1998, p. 67). Xenophobic purism is one of the most obvious forms of purism in the present, with its strong antagonism to code-mixing, switching, or any form of fusion with other languages.

The degree of purism is often driven by emotion, and can range from extreme liberalism to extreme purism. For example, while some see loanwords as an enrichment of the local lexicon, there are others who see them as evil and a threat to their native identity. Therefore, there are psychological factors involved in the concept of linguistic purism, which are often related to identity at both personal and social levels (Hansen et al., 2017, p. 1). Therefore, the extent of purism is also highly arbitrary due to many socio-cultural, religious, economic, geographical, etc. factors of individuals. For example, a person living in a highly developed metropolis may have a different opinion about linguistic purism than a person living in the countryside with more conservative thinking. Linguistic purism has often been associated with religious purism, especially in contexts where religion is highly politicized, such as South Asia, since both language and religion share symbolic values in the construction of group identity (Brass, 2005, p. 03).

Linguistic purism shares the tremendous power vested in language, and political movements around the world have often used linguistic purism, along with nationalist slogans, as a tool to secure political power. Political factions have been able to manipulate language and linguistic purism for both unity and division, depending on their motives. The interplay of language and power dynamics in the colonial and post-colonial eras of South Asia would provide evidence of this phenomenon. The motive of popularizing English in the colonies went far beyond communicative purposes to more ambitious politically driven motives of reshaping the indigenous vernacular, resulting in hierarchical and ideological divisions between groups of different linguistic and educational competence within the indigenous society (Naregal, 2001, p.4). All these hierarchical and ideological fractions have been in search of purism in its various forms during the colonial and post-colonial episodes of South Asian history.

Contemporary social behavior, complemented by language use, shows significant differences from those of previous generations, associated with time-space dimensions, interrelationship with the local and the global, geographical and social mobility, transcendence of territorial and social boundaries, technological breakthroughs, and intense focus on identity issues (Aronin & Singleton, 2012, p. 33). The dramatic increase in the use of the Internet and digital platforms for communication and learning has had a colossal impact on language use, linguistic diversity, and purism. While digital mobility has, on the one hand, provided increased opportunities for language contact, multilingualism and linguistic diversity, it is questionable whether every language has a fair chance to gain popularity and survive within digital platforms such as the dominant lingua franca, English. However, increased localization movements have reduced the presence of English on the Internet to less than 30%, down from nearly 80% in the mid-1990s (Mazrui, 2008, p. 193). Nevertheless, purists and purism have been significantly challenged by the advent and rapid popularization of digital media, mobilization, and migration, which have resulted in inevitable language change.

“In a world where humans grow old, tadpoles change into frogs, and milk turns into cheese, it would be strange if language alone remained unaltered” - (Aitchison, 2012, p. 4)

The increased influx of numerous loanwords and the creation of neologisms are two of the greatest threats to purist ideology. Digital communication platforms, devices, social media, and other applications have introduced neologisms into the language. As the language embedded in computer hardware and software, English continues to expand its far-reaching influence across the globe with even greater diversity (Enteen, 2009, p. 45). As a result, other languages are challenged with the arduous process of either borrowing from English or creating neologisms to represent new concepts. In this regard, the Chinese language has been one of the most successful in coining commonly used neologisms to replace borrowings. Since the late 19th century, Chinese language has undergone an almost complete reconfiguration under Western influences, reaching its peak in the 21st century, triggered by China's "reform and opening up" movement in the 1980s, which eventually initiated a global-level clash between native Chinese and global English (Liu, 2012, p. 2).

1.2 Language Politics and the Quest for Linguistic Purism in Sri Lanka

As in many other contexts around the globe, language movements in Sri Lanka, especially those associated with purism, have always been linked to politics and power. The departure of the British in 1948 marked the end of one geopolitical disruption of Sri Lanka's linguistic landscape, but ushered in another era of post/neocolonial global order with relatively unchanged power dynamics and a more dominant use of the former colonizer's language as a lingua franca (Liyanage, 2021). In Sri Lanka's post-independence history, political parties and leaders have frequently used language policies and planning movements to manipulate electoral outcomes. A notable case can be seen in the actions of Bandaranaike, who initially committed to a bilingual official language policy in 1944, but later contradicted his own words in 1956. Instead of sticking to his earlier position, Bandaranaike exploited the language issue for political gain (DeVotta, 2002, p. 86).

One of the main sources of tension between the Sinhalese and Tamil communities in Sri Lanka has been language policy. After independence, politicians from the Sinhala majority enacted laws and policies that favored Sinhala over Tamil, leading to a sense of marginalization and discrimination among the Tamil minority. This culminated in the introduction of the "Sinhala Only" Act in 1956, which made Sinhala the sole official language of Sri Lanka and effectively marginalized Tamil speakers from government and public life. This move fueled insecurity among the Tamil community, and both Sinhalese and Tamil political factions took advantage of the situation to manipulate votes in their favor. The coalition of Bhikkhus, Sinhala nationalists and the UNP protested against the B-C pact. Among the Tamils, the Ceylon Tamil Congress (CTC) and some Tamil extremists also opposed the agreement (DeVotta, 2002, p. 87).

During the most crucial phase of the national independence movement in the 1930s and early 1940s, Cumaratunga Munidasa spearheaded the 'Hela' movement, which aimed to purify the Sinhala language (Coperehewa, 2011, p. 13). His efforts went far beyond mere linguistic purification to the construction of an indigenous Hela identity, which was driven by nationalist motivations. Hela idealists believed that Sinhala was the key to Sinhala ethnic identity and this ideology influenced the contemporary use of the Hela concept (Koning, 2022, p. 126-127). The Hela ideologues believed that Sinhala should be free of all foreign matter, especially Sanskrit and Pali terminology, which formed a large part of the Sinhala lexicon. Hence, Cumaratunga and his Hela Havula comrades promoted the pure Sinhala register free from consonant clusters resembling Sanskrit or Pali (Henry, 2022, p. 92). In fact, the Hela movement aimed to revive Sinhala culture to its former glory by rejecting foreign influences such as Indian, colonial, and any other external elements (Lange, 2011, pp. 65-66).

After the demise of Cumaratunga in 1944, the Hela movement began to struggle, which survived only among a small group of intellectuals, but did not develop into a major nationalist movement (Rambukwella, 2018, p. 40). This is presumably related to the deviation of the Hela movement from its linguistic purification motives, along with its commercialization and politicization. Other nationalist symbols such as Ravana, the multi-headed demon king believed to be the protector of the island nation, were also associated with the Hela movement in the early 20th century (Koning, 2018, p. 19). These movements became increasingly commercialized towards the end of the 21st century, with many indigenous witch doctors and practitioners of indigenous medicine in Sri Lanka manipulating them to support a divine figure to validate their practice. Towards the first decade of the 21st century, linguistic purification became the least of the Hela movement's concerns as its new disciples were more concerned with ethnic and religious purification.

Despite the inward shift in language policy, the quest for English remained strong as it became the global lingua franca, while the Sinhala language remained the least involved in Sri Lanka's increasing participation in regional and global affairs (Liyanage, 2019, p.403). The colonial language policy established a new hierarchy of languages with English at the top, resulting in a social elite educated exclusively in English. Thus, English proficiency became highly valued not only for practical reasons but also for its social and cultural significance (Ramesh & Wong, 2020, p. 72). The dream of purism has been greatly diminished by the persistent changes in the language of digital media, especially the radical changes in the language of social media. According to the Asia Pacific Institute of Digital Marketing (2021, p. 10), by 2021, there will be 6.4 million social media users in Sri Lanka, representing 29.94% of the total population. This dramatic increase in the use of social media has led to the creation of a new linguistic environment in the digital space where linguistic purism is the least of concerns.

2. Methods

2.1 Theoretical Background

The present study is based on a sociolinguistic approach that incorporates two primary theories regarding the language dynamics of social media. The first and foremost is Communication Accommodation Theory (CAT), which posits that individuals adapt their language and communication style to fit the social context in which they communicate. CAT suggests that users adapt their language and communication style to fit the norms and conventions of the particular platform they are using, as well as the audience they are communicating with. It provides a useful framework for understanding how individuals adapt their language use in order to communicate effectively and build social relationships with others.

Second, the theory of language convergence (LC), which is also related to CAT, suggests that social media users may engage in language convergence or divergence (Giles & Ogay, 2007, pp. 293-294). Language contact and the resulting variation and change, which can lead to both convergence and divergence, are widely observed phenomena in language use (Braunmüller & House, 2009, p. 1). Convergence occurs when individuals adopt a communication style that is similar to the person they are communicating with, while divergence occurs when individuals intentionally use language or communication styles that are different from the person they are communicating with. These concepts are used to analyze how social media users innovate new forms of language use, how social media language practices vary across contexts and communities, and how social media shapes attitudes and beliefs about language use and linguistic norms.

2.2 Materials and Analysis

Corpora and screening

Present Study is a corpus-based study which uses a corpora of Facebook memes posted by Sri Lankan users that demonstrate radical linguistic features deviating from authentic language use. Apart from browsing for memes on Facebook pages, YouTube meme compilations were combed with the hashtags #bukiyerasakatha, #SinhalaFBmemes and #SriLankanmemes. A total of 1600 memes were manually collected within a period of three months from 2nd June 2022 to 2nd September 2022. Following Sanei (2021, p. 459), all data were systematically collected through non-participatory observation. The criteria for selection of memes comprised of three primary measures as,

- Intentional mistakes (Kostadinovska--Stojchevska & Shalevska, 2018, p. 166) such as non-standard spelling, and grammar.
- Lexical borrowing - Borrow language from other languages and sources, such as popular culture references, catchphrases, or slang.
- Intertextuality (Vukadin, 2019) - Incorporate text from other sources, such as popular culture references or quotes from movies or TV shows.

The initial collection of 1600 memes was then screened for duplicate content, offensive or inappropriate content, political bias, identifiable personal information, and low engagement. After screening, the final set of memes consisted of 431 memes.

Analysis

The present study triangulates statistical analysis, content analysis, and critical discourse analysis (CDA). CDA views language as a type of social activity, and its focus is on conducting methodological investigations of hidden power dynamics and ideologies embedded in discourse (Johnson & McLean, 2020). The analysis examines the extreme linguistic features of the social media memes and their sociolinguistic implications, juxtaposed with the purist ideology. The following criteria were used to analyze the language of the memes.

- Abbreviations and Acronyms: Popular acronyms used in social media and new acronyms created within social media.
- Code Mixing and Code Switching: Code switching and code mixing between two or more languages found within memes
- Catchphrases: Catchphrases and hashtags that categorize and organize content, fostering trends and discussions around specific topics.
- Slang and Neologisms: New slang and neologisms that quickly spread among users, reflecting current trends and shared experiences.
- Influencer Language: The language used by influencers, including unique catchphrases and expressions, can influence the broader online community.
- User-Generated Content Influence: Popular phrases or linguistic styles adopted by users.

Semiotic factors such as symbols, images, and artwork were not included in the analysis. The study focuses on sociolinguistic aspects of language use, such as intentional linguistic deviations and code-switching, which are primarily textual manifestations. The exclusion of semiotic factors avoids the complexity of multimodal analysis and the need for additional theoretical frameworks. This focus ensures greater depth and rigor in the analysis of the linguistic dimensions of social media memes. Common linguistic features were categorized and their frequencies statistically analyzed. For ease of use and data friendliness, statistical data was entered into Microsoft Excel and formulas were used for analysis.

3. Results and Discussion

The corpora of 431 memes consisted of 7499 words in total with a mean of 17.3 words in a meme. Table 1 demonstrates the word frequencies of the corpora.

Table 1. Rudimentary word frequencies of the corpora.

Category	f	%
No of memes with code-mixing or code switching	391	90.71
No of non-Sinhala words	1949	25.99
No of English words	1686	22.48
Abbreviations and Acronyms	70	0.93
Hashtags and Neologisms	23	0.30

While the number of memes with code-mixing itself is over 90% of the total corpora, the percentage of code-mixing within memes is also significantly high. The word 'like' (n=137) and 'share' (n=102) had the highest frequencies out of the total number of words, followed by the word 'comment' in both its Sinhala and English forms with a total of 85 occurrences, the catchphrase 'my time has come' (n=62), 'කුෂ්ඨක/krʌʃ ekə/ (n=54), which in Sinhala refers to 'the crush or infatuation someone has for another', msg' (n=48), which is an acronym for the word 'message'. In addition to English words, a number of words from other languages were also found in the memes. 42 occurrences of Hindi words were found, largely due to a popular catchphrase and 11 Chinese characters, and 26 occurrences of Tamil words and the French word 'le'. Table 2 shows the frequency of code-mixing within the memes.

Table 2. Code mixing within memes.

% of code-mixing within a meme	75	60	40	20	10
<i>n</i>	18	31	75	259	248
% out of the total no of memes (n=431)	4.18	7.19	17.40	60.09	80.74

Table 3 shows the distribution of themes across the corpora, and only the most obvious and prominent theme of each meme was considered in the analysis. It could be inferred that the theme 'politics and social issues' has less popularity within the corpora, while 'love and romance' and 'social media' are the most popular themes. Another significant finding is that 'religion and faith' has the lowest frequency in the distribution, which could be related to the recent legal action taken in Sri Lanka based on the ICCPR, which has significantly restricted the mockery of religion and popular beliefs on social media. The large number of memes that are contextualized in classrooms, schools, universities and learning indicates that the creators of these memes are mostly adolescents

Several significant sociolinguistic implications were elicited through the content analysis of the corpora.

3.1. . Social media as a safe space for less competent language users

A considerable number of intentional and unintentional errors in spelling, grammar and word order were found in both the Sinhala and English corpora. The errors in English, in particular, highlight an important sociolinguistic factor. As revealed in the literature review, English education in Sri Lanka is associated with the imperialist ideology of the colonizers, which has given it a higher social value than the local vernacular. Low English language proficiency in Sri Lanka was often associated with learner anxiety caused by the fear of making mistakes. The prevalence of intentional and

unintentional language errors in the memes suggests that social media users are less anxious about language accuracy. See the examples in Table 4.

Table 3. Thematic distribution of the corpora.

Theme	f	%
Sports	14	3.25
School and Education	44	10.21
Technology	24	5.57
Love and Romance	126	29.23
Politics and Society	25	5.80
Social Media Presence	79	18.33
Workplace	9	2.09
Faith and Religion	4	0.93
Abstract humor	106	24.59

Table 4. Language mistakes in the corpora.

Extract	Language	Correct form	English meaning
රෙස්පට් බුදුර්	English (transcribed in Sinhala)	රෙස්පෙක්ට් බුදුර්	Respect brother
පයන	Sinhala	ප්‍රශ්න	Issues/problems
ගර්ල්පෙන්	English (transcribed in Sinhala)	ගර්ල්ෆ්‍රෙන්ඩ්	girlfriend
පොස්	English (transcribed in Sinhala)	පොෂ්	posh
ෆේස්බුක්	English (transcribed in Sinhala)	ෆේස්බුක්	Facebook
brathar	English	brother	brother
Resalts	English	Results	Results

From a linguistic point of view, the pronunciation of consonant clusters in English, as in the word 'girlfriend' in the above example, is challenging for Sri Lankan learners, and they often tend to substitute the sound 'f' with 'p' for ease of pronunciation. Such mistakes, which are often subject to ridicule in formal contexts, are commonly observed on social media platforms.

3.2. Invention of new forms of words to avoid taboo and ban.

Although several decades have passed since Sri Lanka's ratification of the ICCPR in 1980 (Hyndman & Human Rights Watch, 1992, p. 3), there has been a significant shift in its application over the past decade as the government has taken steps to legally address the spread of hate speech and abusive language against groups and individuals on social media. The government has also taken steps to negotiate with social media companies to restrict or ban the presence on social media of individuals and groups who disseminate hate speech. However, what is considered hate speech or taboo is still a matter of serious debate and an effective content moderation database on such grounds is still far from reality in the case of Sri Lanka.

Social media users in Sri Lanka are using creative methods to produce content that bypasses content moderation databases. Two of the most common methods are:

- Blending of characters of L1 and L2 – eg. ‘මරනවා (means ‘to kill’)
- Substitution of characters of similar phonemes - eg. ‘ses’ to mean ‘sex’

These words have become very popular among social media users in Sri Lanka, and many swear words and other offensive words have been modified to avoid being banned or restricted on social media. Users have increasingly embraced these modified versions of offensive words and phrases, as the modifications have effectively reduced their level of offensiveness.

3.3 Borrowing, code-mixing and code switching

Numerous borrowings and code-switching were identified in the corpora, with many gaining popularity among users of social media. Words such as ඉමෝජිය (emoji), ප්‍රශ්නපත්‍ර (question paper), කමෙන්ටුව (comment), ප්‍රතික්‍රියකරුවන් (reactors), මිමර්ලා (memer), ලෙජන්ඩ් (a legend), යූසර්ස් (users), කස්ටමර්ස් (customers), චාජර් (charger), බ්‍රෑන්ඩ්

(to be dumped/to break up), ජොබ් (job), බංගලිශ් (Bangladeshi), ස්ටෝරි (facebook story), ගැලරිය (picture gallery on mobile) are frequently used borrowings in the corpora. Most of these are phonetic borrowings which involves adopting words based on their phonetic attributes rather than their semantic content.

There were other forms of peculiar adaptations where a neologism has been formed based on the phonetic attribute of a foreign word. The masculine form for the word ‘bestie’ which means ‘the best buddy’ in English has been formed as බෙස්ටා /besta:/ in Sinhala on social media. The word ‘Genius Club’ has been phono-semantically matched as දිනියස් එකතුව where the first word is a phonetic borrowing and the second is a translation. There were other direct translations such as the word ‘black prince’ for the Sinhala word කළු කුමාරයා¹. A key question is how much these words have penetrated into Sinhala. Although these words are largely used on social media at present, it is a matter of time till these words are fossilized within the local vernacular.

As claimed by Haspelmath (2009, p. 42), the extent of adaptation often fluctuates based on factors such as the age of a loanword, the level of familiarity with the donor language among recipients, and their overall attitude toward the donor language. In terms of the attitude towards the donor language, it is evident that the modern social media users in Sri Lanka have less anxiety in using words, utterances and sentences of the donor languages which stands highly in contrast with classroom language learning where there is a high level of learner anxiety, especially towards English. The extent of code-switching and code-mixing found within the memes are in evidence of the fact that the memers have created a new hybrid language which is free from anxiety towards making errors, borrowing, code-mixing and even creating their own neologisms.



Figure 1. An example for extensive code-mixing.



Figure 2. An example for extensive code-mixing.

Apart from the extensive code-mixing within the meme in Figure 1 and Figure 2, they both have some specific linguistic characteristics. Aside from the impurities of code-mixing, the structural integrity of the sentence in Figure 1 is acceptable, suggesting that it was composed by a user with a high level of English proficiency. In Figure 2, the sentence at the top is grammatically correct, while the sentence at the bottom has an intentional error added to the code-mixing. Second, all three sentences are English sentences with Sinhala slang infused into them, whereas the reverse is typically the case with English words infused into Sinhala sentences. This suggests that social media now serves as a unifying platform for language users with different levels of proficiency, unlike earlier times when there was a significant divide between high and low proficiency users. Most importantly, linguistic purism is one of the least of these language users' concerns, where different forms of language contact, blending and hybridity exist.

3.4. Hybridized urban slang on Sri Lankan social media.

In addition to the prevalence of urban slang in the corpora, there are numerous emerging trends worthy of exploration. The analysis observed that urban slang used on social media undergoes swift transformations, with newly coined words and phrases evolving within a short timeframe. Old phrases in urban slang such as ‘රෙදි නෑ’ has been modified as ‘රෙදි ඉල්ලේ’ of which the second word has been replaced with its Tamil translation. The phrase ‘confused unga bunga’ featuring a screenshot from the movie *The Last Hunter* (2010) was localized as ‘චික්චින්න unga bunga’ in

¹ The word කළු කුමාරයා which is translated in this example as ‘black prince’ refers to a demon in Sri Lankan folk culture who is said to have an influence on young girls who travel at dusk.

Shattered Dream of Purism: Sociolinguistic Dynamics of the Emerging Hybrid Social Media Language in Sri Lanka
Dassanayake, N.

which the first word ‘confused’ has been translated into Sinhala. This catch phrase appeared 23 times within the corpora which accounts for the second highest frequency of all catch phrases.

Social media users have also started to coin neologisms, often adapting from English words and in some cases generating within local social media discourse itself. Words such as විපර් වෙලා (to be annoyed by something) ලෙජන්ඩ් කෙනෙක් (someone who possesses exceptional skills in something), සුදාම්මි (girlfriend), සුදාන්ති (boyfriend) and නිබ්බි (newbie) are few frequently used neologisms. New abbreviations and acronyms have been generated in Sri Lankan social media some of which are demonstrated in Table 5.

Table 5. New acronyms and abbreviations extracted from the corpora.

Item	Type	Expansion	Language	Meaning
kmli	abbreviation	Kammaliyi (කම්මලියි)	Sinhala transcribed in English	lazy
mkd	abbreviation	Mokada (මොකද)	Sinhala transcribed in English	why
wdk	abbreviation	Wedak (වැඩක්)	Sinhala transcribed in English	some work
khmd	abbreviation	Kohomada (කොහොමද)	Sinhala transcribed in English	how
ss	acronym	screenshot	English	screenshot
33kdp	acronym	නිස්සන්කෝටියක් දෙවි පිහිටයි	Sinhala transcribed in English	330 million gods' blessing ²
msg	acronym	Message	English	message
app	acronym	Application	English	application
ගුණ බුස	acronym	ගුණ නයිට් බුස සරණයි	English transcribed in Sinhala	good night
DP	acronym	Display picture	English	display picture

Some peculiar alterations in words were found in the corpora that have gained extensive popularity on social media in Sri Lanka. One such case is the substitution of the sound ම/m/ for other sounds i.e. එමන for එන්න (to come), දෙමන for දෙනන (to give), කෙමලෙක් for කෙල්ලෙක් (a girl), කමන for කන්න (to eat), රිමදෙනවා for රිදෙනවා (it hurts). Similarly, there was a trend of substituting the sound න for ද in words as in the words හරින for හරිද , යමුන for යමුද, කරමුන for කරමුද. Nevertheless, it has been noted that these trends are transient, arising primarily for amusement and holding minimal potential of fossilization and integration into local vernacular. A notable evolution in language usage on social media in Sri Lanka over the past decade is the extensive adoption of taboo, lewd, and explicit expressions. What holds greater sociolinguistic significance is that these expressions have transcended boundaries of gender, age, and social status.

The analysis clearly shows that social media has given rise to a dynamic, distinctive and diverse linguistic environment where conventional structural and lexical rules have little influence. From a sociolinguistic perspective, the language used in social media has penetrated the traditional frameworks of dominant cultures and languages, bypassing established belief systems, religions, or other norms that have controlled language use in the offline world. The question of whether changes in the virtual realm can infiltrate local vernaculars and, more importantly, how these changes might affect standard language, remains controversial. However, as Crystal (2011, p. 57) asserts, history is no guide when it comes to electronic technology, and unlike the past, when it took a considerable amount of time for a newly coined word to gain "high community profile," the creation of a neologism can now occur in a matter of hours. Moreover, an additional disadvantage for indigenous languages facing a dominant lingua franca is the lack of newly coined terms to articulate emerging concepts in the global context, especially those related to technology and virtual presence.

Under these circumstances, dreams of linguistic purism amount to speculative optimism. In the virtual world, geographical boundaries do not neatly coincide with linguistic boundaries. Online participants who belong to more than one language community share highly dynamic sociolinguistic norms regardless of their geographic location (Danet & Herring, 2007, p.7). In addition, the virtual world offers much more versatile and rapid means of intercultural communication, where multiple languages and cultures become accessible to its participants. According to Datareportal's Digital Global Overview reports for the years 2022 and 2023, the number of active Internet users has grown from 4.95 billion to 5.16 billion (a growth of 98 million users) in just one year, and the number of active social media users has grown from 4.62 billion to 4.76 billion (a growth of 137 million users) in the same period (Kemp, 2022), (Kemp, 2023). These data envision a human race with a more widespread universal virtual presence, where both tangible and intangible boundaries have little control over robust linguistic and cultural norms as in the past.

² 330 million gods' blessing – There is a deep-rooted belief in Sri Lankan Buddhist culture that there are 330 million gods protecting them and this is an utterance used for obtaining their blessing.

4. Conclusion

The modern digital world has catalyzed a rapid evolution of language, introducing new terms, abbreviations, and expressions that reflect the dynamic nature of communication in online spaces and social media platforms. As users engage in real-time conversations, share content, and navigate the vast expanse of the Internet, a linguistic metamorphosis unfolds, providing a fascinating lens through which to understand the rapid transformation of language in the digital age. In considering today's evolving linguistic environment, the question arises as to whether linguistic purism is an elusive ideal, given the constant flux of language influenced by technological, cultural, and global dynamics. For these reasons, the present study assessed the sociolinguistic implications of the new social media language in Sri Lanka through content analysis of a corpora of social media memes.

The study revealed several significant shifts in the new language environment of the cyber age in Sri Lanka, especially triggered by social media. Local languages, along with English as a second language, have undergone remarkable transformations in the virtual realm, which are gradually fossilizing within the local vernacular. Under these circumstances, aspirations for language purification face challenges, and linguistic purism as a political concept seems to be losing influence over time. From a broader perspective, as Langer & Nesse (2012, p. 610) argue, linguistic purity itself does not exist, firstly because there is no way to trace the exact pure origin of a language, and secondly because every language is a product of some contact with other languages.

The analysis of memes also highlights the dynamic and hybridized nature of language in Sri Lankan social media, challenging traditional notions of linguistic purism and standardization. This study highlights the importance of social media as a platform for linguistic innovation and cultural negotiation, and provides a framework for further exploration of the impact of digital communication on language use and identity formation. As noted in the introduction, there is a significant research gap in studies of the emerging language landscape in Sri Lanka. Studies on second language learning are largely anchored in the classroom, often overlooking the significant impact of real language situations outside the classroom on the teaching-learning process. Further studies of social media discourse on topics of debate such as politics, culture, and religion would shed much light on the evolving language dynamics in Sri Lanka.

References

- Ager, D. E. (1999). *Identity, insecurity and image: France and language*. Bristol: Multilingual Matters.
- Aitchison, J. (2012). *Language change: Progress or decay?* Cambridge: Cambridge University Press.
- Aronin, L., & Singleton, D. (2012). *Multilingualism*. Amsterdam, Netherlands: John Benjamins Publishing.
- Asia Pacific Institute of Digital Marketing. (2021). *Digital Outlook Sri Lanka 2021. Annual Market Insight Report*, Retrieved from: <https://apidm.lk/apidm/Digital%20Outlook%202021.pdf>
- Brass, P. R. (2005). *Language, religion and politics in North India*. Bloomington, USA: iUniverse.
- Braunmüller, K., & House, J. (2009). *Convergence and divergence in language contact situations*. Amsterdam, Netherlands: John Benjamins Publishing.
- Brunstad, E. (2003). Standard language and linguistic purism. *Sociolinguistica*, 17(1). <https://doi.org/10.1515/9783110245226.52>
- Comrie, B., & Thomas, G. (1994). Linguistic purism. *Language*, 70(4), 844. <https://doi.org/10.2307/416339>
- Coperehewa, S. (2011). Purifying the Sinhala language: The HelaMovement of Munidasa Cumaratunga (1930s–1940s). *Modern Asian Studies*, 46(4), 857–891. <https://doi.org/10.1017/s0026749x11000291>
- Coronel-Molina, S. M. (2015). *Language ideology, policy and planning in Peru*. Bristol: Multilingual Matters.
- Crystal, D. (2011). *Internet linguistics: A student guide*. London: Routledge.
- Danet, B., & Herring, S. C. (2007). Introduction: Welcome to the Multilingual Internet. In B. Danet & S. C. Herring (Eds.). *The multilingual internet: Language, culture, and communication online*. Oxford: University Press.
- DeVotta, N. (2002). South Asia Faces the Future: Illiberalism and Ethnic Conflict in Sri Lanka. *Journal of Democracy*, 13(1).
- Enteen, J. B. (2009). *Virtual English: Queer internets and digital creolization*. London: Routledge.
- Geers, M. (2005). A Comparative Study of Linguistic Purism in the History of England and Germany. In N. Langer & W. Davies (Eds.). *Linguistic Purism in the Germanic Languages* (p. 98). Berlin: Walter de Gruyter.
- Giles, H., & Ogay, T. (2007). Communication Accommodation Theory. In B. B. Whaley & W. Samter (Eds.), *Explaining communication: Contemporary theories and exemplars* (pp. 293–310). New Jersey: Lawrence Erlbaum Associates Publishers.
- Hansen, K., Wypych, M., Bańko, M., & Bilewicz, M. (2017). Psychological determinants of linguistic purism: National identification, conservatism, and attitudes to Loanwords. *Journal of Language and Social Psychology*, 37(3), 365–375. <https://doi.org/10.1177/0261927x17737810>
- Haspelmath, M. (2009). *Lexical borrowing: concepts and issues*. In M. Haspelmath & U. Tadmor (Eds.). *Loanwords in the world's languages: A comparative handbook*. Berlin: Walter de Gruyter.
- Henry, J. W. (2022). *Ravana's kingdom: The Ramayana and Sri Lankan history from below*. Oxford: University Press.

- Hyndman, P., & Human Rights Watch. (1992). *Human rights accountability in Sri Lanka*. Library of Congress Catalog No. 92-72416, Retrieved from: <https://www.hrw.org/reports/pdfs/s/srilanka/srilanka925.pdf>
- Jernudd, B. H. (1989). *The Texture of Language Purism: An Introduction*. In B.H. Jernudd & M.J. Shapiro (Eds.). *The politics of language purism*. Berlin: Walter de Gruyter.
- Johnson, M. N., & McLean, E. (2020). *Discourse Analysis*. In A. Kobayashi (Ed.). *International encyclopedia of human geography*. Amsterdam: Elsevier Science.
- Joseph, J. E. (2006). *Language and politics*. Edinburgh, UK: University Press.
- Kemp, S. (2022). Digital 2022: Global Overview Report. Datareportal. Retrieved from: https://datareportal.com/reports/digital-2022-global-overview-report?utm_source=Global_Digital_Reports&utm_medium=PDF&utm_campaign=Digital_2023&utm_content=Digital_2023_Global_Overview_Promo
- Kemp, S. (2023). Digital 2023: *Global Overview Report*. Datareportal. Retrieved from: <https://wearesocial.com/wp-content/uploads/2023/03/Digital-2023-Global-Overview-Report.pdf>
- Koning, D. D. (2018). The ritualizing of the martial and benevolent side of Ravana in two annual rituals at the Sri Devram Maha Viharaya in Pannipitiya, Sri Lanka. *Religions*, 9(9), 250. <https://doi.org/10.3390/rel9090250>
- Koning, D. D. (2022). *"Ravanisation": The revitalisation of Ravana among Sinhalese Buddhists in post-war Sri Lanka*. Münster: LIT Verlag.
- Kostadinovska-Stojchevska, B., & Shalevska, E. (2018). Internet Memes and Their Socio-Linguistic Features. *European Journal of Literature, Language and Linguistics Studie*, 2(4). <https://doi.org/10.5281/zenodo.1492894>
- Lange, M. (2011). *Educations in ethnic violence: Identity, educational bubbles, and resource mobilization*. Cambridge, UK: Cambridge University Press.
- Langer, N., & Nesse, A. (2012). *Linguistic Purism*. In J. M. Hernández-Campoy & J. C. Conde-Silvestre (Eds.) *The handbook of historical sociolinguistics*. New Jersey: Blackwell Publishing Ltd.
- Liu, J. (2012). *Chinese under globalization: Emerging trends in language use in China*. Singapore: World Scientific.
- Liyanage, I. (2019). *Language Education Policy in Sri Lanka*. In A.Kirkpatrick & A. L. Liddicoat (Eds.), *The Routledge international handbook of language education policy in Asia* (pp. 399-413). New York: Routledge.
- Liyanage, I. (2021). *Geopolitics and trilingual media of instruction in post-conflict Sri Lanka*. In R. Arber, M. WiennMann & J. Blackmore (Eds.). *Rethinking languages education: Directions, challenges and innovations* (pp. 148-163). New York: Routledge.
- Mazrui, A. (2008). *Globalization and the Sociolinguistics of the Internet: Between English and Kiswahili*. In C.B. Vigouroux & S.S. Mufwene (Eds.). *Globalization and Language Vitality: Perspectives from Africa*. London, UK: Bloomsbury Publishing.
- Naregal, V. (2001). *Language politics, elites, and the public sphere* (p. 4). New Delhi: Permanent Black.
- Rambukwella, H. (2018). *The politics and poetics of authenticity: A cultural genealogy of Sinhala nationalism*. London, UK: UCL Press.
- Ramesh, S., & Wong, M. P. (2020). *Sri lankans' views on English in the colonial and post-colonial eras*. Newcastle, UK: Cambridge Scholars Publishing.
- Sanei, T. (2021). Normativity, power, and agency: On the chronotopic organization of orthographic conventions on social media. *Language in Society*, 51(3), 453-480. <https://doi.org/10.1017/s0047404521000221>
- Thomas, G. (1991). *Linguistic purism*. Harlow, UK: Longman Publishing Group.
- Thomas, G. (1998). *The Development of Slovak Purism*. UK: Oxford University Press.
- Vukadin, A. (2019). *Metadata for Transmedia resources*. UK: Chandos Publishing.