

Nativising Strategies in Mixed Contexts

C. Dilkushi Senarathne

Abstract

In nativization, other language items are integrated into the borrowing language with those items assuming similar phonological, morphological, syntactic and semantic features. This paper analyses nativizing strategies of 63 selected single English lexical items in Sinhala newspapers that have undergone immersion due to contact with the Sinhala language. The words were chosen from an initial survey of three broadsheet and two tabloid Sinhala newspapers. In addition, websites of electronic media channels were surveyed to substantiate data gathered from the newspaper survey. The selection of words was based on the Markedness Model. Research conducted by Kachru on nativization and Muysken's CM theory were used to identify patterns of integration of single word items. The items were categorized as marked and unmarked items based on Myers-Scotton's Markedness Model. From the selected words, 46 were identified as marked items revealing specific nativizing patterns and 17 were unmarked items revealing insertional patterns. The findings reveal nativization as a productive strategy used in written contexts to facilitate inclusion of single English words. The analysis sheds light on differentiating borrowings from code mixes. The findings reveal the morphological processes used to nativize and integrate English lone words in mixed contexts.

Keywords: Nativization, Insertion, borrowing

Introduction

Most Sri Lankans are either bilinguals or multilinguals and many can speak more than two languages. Sinhala is the language spoken by the majority and is also the most widely used (Fernando 1977). The Department of Official Languages in its Official Languages Policy ⁴ states that Sinhala and Tamil are both official languages while English is recognized as a link language in the country. Sinhala is an Indo-Aryan Language and according to Gair (1998) has had significant contact with Dravidian languages (of South India) and colonial languages (Portuguese, Dutch and English). Hence, its word-stock is enriched with borrowings or occasional mixes from all the languages it has come in contact with. English, as the link language, enjoys considerable prestige and is a marker of social class. English used in Sri Lanka is a variety and is referred to as Ceylon English, Lankan English and Sri Lankan English (Fernando 1985). The context of mixing between Sinhala and English has given rise to many innovative strategies in language use both in formal and informal domains. Research conducted in Sinhala and English code mixing and the resultant discourse strategies reveal the working of four functions that the bilingual employs namely foregrounding, neutralization, nativization and hybridization (Senaratne 2011). This study deals mainly with English lone lexical items collected through an initial survey of Sinhala newspapers. Items that reveal the process of nativization and integration in matrix Sinhala sentences were selected for the study.

Literature review

The process of mixing is a highly debated topic. It is evident that mixing and borrowing are two different phenomena. Many terms are used to refer to mixed data found in a bilingual corpus. They can be borrowings or nativizations, 'loans', established loans', switches,

⁴http://www.languagesdept.gov.lk/web/index.php?option=com_content&view=article&id=38&Itemid=163&lang=en accessed on May 6, 2019.

'mixes or even 'nonce borrowings' (Weinreich 1953). Hocket (1958), as quoted by Hoffer (2002), mentions that speakers of languages have a variety of options when confronted with other language items. Accordingly, mixed items are analyzed as loanwords (using the grammatical processes of the new language such as nouns taking on the plural or possessives of the new language), loan-shifts (adapting native words to the new meanings), loan-translations (item for item native version of the original) or loan-blends (one element is a loanword, and the other is a native word). It is argued that these options are used by bilingual and multilingual speakers when bringing a new word to the language. In essence, it is plausible to state that all borrowings at some point during their life span, started their journey into the other language as a code mix. Hence, understanding the relationship between a mix and a borrowing is crucial when attempting to define borrowings or nativizations in bilingual and multilingual contexts.

Nativization or borrowing employs complete integration of other language items in the borrowed language (Poplack and Meechan 1995). Due to the frequency of use, the subsequent integration they undergo in the borrowed language and the functional uses, borrowings become fossilized in the recipient language differentiating them from other 'spontaneous' or less reputed counterparts.

However, many researchers consider borrowing as Code Switching (Myers-Scotton 1993) or a strategy that is immersed in Code Mixing (Muysken 2000). It is clear that linguists have always tried to define the presence of other language items irrespective of their life span in the borrowed language. This study will use the terms nativization and borrowing interchangeably to refer to the same phenomenon.

The motivations to nativize or borrow other language items can be many. According to Grosjean (1995), borrowing takes place when a 'word or a short phrase' (usually phonologically or morphologically) is

borrowed from the other language or when the ‘meaning component’ of a word or an expression in the foreign language is expressed in the base language. This expression can take various forms, either as borrowings, or their less reputed counterparts, code mixes. Borrowings are usually integrated lexical items as opposed to code switches (Grosjean 1982). However, the complexity of identifying the two is acknowledged.

Nativizations in mixed data

Proposing a hierarchical model to describe mixing patterns, Kachru (1983) ranks mixing of simple lexical items as the lowest (NP insertion, VP insertion, Unit hybridization) and mixing of entire sentences as the highest (sentence insertion, idiom and collocation insertion, inflection attachment and reduplication) in bilingual contexts. The analysis emphasizes the social significance of mixing where CM is used for register identification as a ‘foregrounding’ strategy to ‘attract attention’, and a discourse strategy for specific communicative purposes such as style, elucidation, neutralization and interpretation. A function is assigned for each code (Kachru 1983). Kachru’s analysis reveals CM as a discourse strategy used by educated Indians.

His analysis is based on the Indian context of CM where three strategies of mixing are identified namely Englishization (where English is mixed), Sanskritization (where Sanskrit is mixed), and Persianization (where Persian is mixed). In these three strategies, the base language is always one of the many Indian languages, mixed with three other languages: English, Persian and Sanskrit (Kachru 1983).

Nativization, according to Kachru (1978), is a result of extensive contact between languages where elements from the donor language are integrated into the base language and the donor language acts as an additive source of linguistic material in the development of a specialized register. Motivations for the adoption of

these mixed codes by a speaker will be different. A highly Englishized code is used for political and administrative purposes which may signal aspirations of upward mobility of the speaker (Kachru 1978/1979). Englishized Hindi is used with the family and non Englishized Hindi with domestics. The Englishized mixed code, typical of the Indian Middle class, is used as a marker of high social class. The code is also a symbol of elite membership and an expression of power and prestige. Instances where one of the local languages will be mixed with Sanskrit, which is labelled as Sanskritization, can be a marker of caste or religious identity used in philosophical, literary or religious discourse (Kachru 1978) whereas the third type of mixing, Persianization, is associated with Muslim culture symbolizing the Muslim religious identity. Englishization is also used to re-express and re-define what has already been stated in the native language. Another positive attribute of Englishization is that it does not reflect religious connotations and hence 'cuts across religious boundaries'.

Hierarchical models to explain well-formedness in mixed utterances encounter difficulties as pointed out by Sankoff (1998) and Muysken (1995) since the models cannot ensure compatibility of the same-language parts of neighbouring constituents.

Insertions in mixed data

Muysken's (2000) theory of CM describes borrowings, nonce borrowings and constituent insertions as part of insertional code mixing where the order of the sentence displays a nested a b a structure. In insertional CM, a dominant language is always at work. Insertions are usually single, selected content words.

Nonce borrowings in mixed data

The theory of nonce borrowing is used to explain the vast majority of spontaneous other language items or speech borrowings in mixed data. Accordingly, Poplack and Meechan (1995) differentiate

the two categories of borrowings as (a) established borrowings, defined as lexical items that are morphologically, syntactically and often phonologically integrated into the borrowed language and (b) as 'nonce' borrowings (Poplack et al 1988; Sankoff et al. 1990). A 'nonce' borrowing is defined as an 'incorporation' of a singly uttered word from another language by a single speaker in some reasonably representative corpus. These are mostly content words, which display similar morphological, syntactic and phonological features as their established counterparts, borrowings, the only difference being that they are neither recurrent nor widespread. In an elaborate analysis of the two processes, Sankoff et al. (1990) suggest that the two processes are best distinguishable by the degree of syntactic and morphological integration of the loanword into the host language.

Poplack and Meechan (1998) contend that in lone word integration or switching, the items are usually positioned syntactically in the borrower language and often appear to retain the phonological and morphological properties of the donor language. In essence, the theory proposes a linear model to describe mixing patterns. Integration is defined under the Equivalence Constraint (switching that does not violate syntactic rules of either language) and the Free Morpheme Constraint (switching that does not occur word internally). They argue that CS implies alternation and hence, single word code-switches should therefore show 'less integration' into the other language (Poplack and Meechan 1998:129). The presence and degree of phonological and morpho-syntactical integrations accordingly are the main features that enable the distinction of a borrowing from a code switch. A small number of lone other language items are categorized as code-switches, as they do not pattern with the recipient-language whereas 'most' lone other-language items are defined as borrowings. Words that do not fall into both categories carry a different label. Mixes that violate the constraints are treated under a special borrowing system labelled 'nonce borrowings'. The constraints apply only to CS. Nonce borrowings display total embedding in the Matrix Language (ML) and display morphological integration.

Poplack and Meechan (1998) argue in favour of nonce borrowing, proposing that however different linguistic properties are in language pairs, and however typologically distant they are, lone other language items can always resurface. This accordingly, is evidence that these single word elements are borrowed into the recipient language. While differing from the relevant patterns of the donor language, the lone items can be considered borrowed, as only the grammar of the recipient language is operative.

Markedness in mixed data

The Markedness Model (MM) developed by Myers-Scotton (1993) explains speaker code choice and motivations for Code Switching. Myers-Scotton's (1993) MM proposes that speakers have a 'sense of markedness' in the use of the linguistic codes available to him/her and that all code choices can be explained in terms of speaker 'motivations' (Myers-Scotton 1993:109). The model explains that in any given linguistic context where mixing takes place a code remains 'unmarked' or in other words 'neutral' whereas in other cases the codes are 'marked' and carry linguistic or 'meta- messages'. Myers-Scotton explains social identity negotiation as the major motivation for the variety of choices made by the speakers. It is further stated that speakers are well aware of the consequences of the marked or unexpected choices. Accordingly, speakers prefer the 'unmarked' choice for safer and simpler reasons. The unmarked choice contains no surprises and sustains interpersonal relationships. The 'unmarked' code generally acquires 'fewer distinctive features' and represents 'greater frequency' (Myers-Scotton 1993). In essence, the language that is less expected to be used at a given interaction with an interlocutor represents the 'marked code choice'. Both cases involve costs and rewards, according to the MM, as language is used to communicate 'much more than referential meanings' (Myers-Scotton 1993: 96). Accordingly, it is plausible to conclude that CS or CM is the unmarked choice of bilingual speakers.

Sinhalizations in mixed data

Analyzing the phonological integration of nativized elements poses many difficulties in bilingual contexts due to the influence of L1 when pronouncing English words. As Sankoff (1998) points out, bilinguals, in contrast to monolinguals, have linguistic awareness of the etymology of the words they nativize and therefore, will reflect this knowledge when they pronounce the words. Another reason is the difference in the learning contexts of the languages concerned that results in individual speaker variations in pronunciation.

However, Senaratne (2009) defines phonological integration of English items in the Sinhala and English mixed corpus as Sinhalizations. Unlike borrowings, Sinhalization gives rise to phonological deviations that results in errors in utterances where the matrix language is English. However, in Sinhala matrix contexts, the deviations are simply results of the influence of L1 where the speaker makes use of the internalized grammar to accommodate foreign items in speech (Senaratne 2012). Structural features of Sinhalizations are vowel prefixed consonant clusters, word-final fricative deletion and items followed by the mixed nominalizer. In essence, Sinhalizations project the working of two grammars that result in unexpected deviations based on the speaker's L1. Senaratne (2011) finds four discourse strategies used in bilingual discourse of Sri Lankan Sinhala-English bilinguals such as foregrounding, neutralization, hybridization and nativization. The analysis identifies nativization as the most significant strategy, as attitudinal and contextually neutral English insertions were found in the corpus. In addition, it is identified as a productive strategy where items are adapted in a contextual manner without violating the rules of the languages concerned. Hence, nativization is a productive grammatical process used by the bilingual to accommodate English insertions in matrix Sinhala utterances.

Methodology

The 63 English lone lexical items were selected through a survey of more than 100 words found in popular Sinhala newspapers published in 2019. Words were chosen from an initial survey of three broadsheet and two tabloid Sinhala newspapers. The selection of the three broadsheet newspapers was based on Sri Lanka (SRL) Media Facts newspaper readership survey and LMRB NDM survey. In addition, websites of electronic media channels were surveyed to substantiate data gathered from the newspaper survey. The analysis was based on both online and printed versions of the selected media.

The study selected only those words written in Sinhala using the Sinhala alphabet. The study excluded words written in English in the Sinhala newspapers (mostly in the online versions) as phonological deviations were not explicit. Repetitions and direct insertions with no significant integration pattern were excluded from the analysis as the study was conducted mainly to identify patterns of integration. The selection excluded company names, brand names and other acronyms since their usage in the content of news is unavoidable. The contexts in which the words occurred, in other words, the surrounding environment and the words preceding and following the mixes, were also considered for categorization and exclusion.

English insertions were then categorized as marked and unmarked using Myers-Scotten's Markedness Model. The criteria to categorize the marked items were determined by the morphological processes involved in the integration of the selected words in the borrowing language. The unmarked items were categorized based on insertional CM patterns proposed by Muysken (2000). Unmarked lexical items were mostly single word code mixes that depicted nested a b a structures based on Muysken's CM theory. Kachru's theory on nativization and Muysken's typology of CM were used to identify patterns of integration in borrowings and insertions. Senaratne's (2011) theory on Sinhalizations was used to identify patterns depicting explicit phonological deviations in English lone lexical items written in Sinhala.

Analysis

Accordingly, the analysis revealed the following distinctive structural characteristics in the marked and unmarked English inclusions in matrix Sinhala sentences.

Patterns found in 'marked' English words written in Sinhala

- a. English inanimate singular nouns with Sinhala suffixes and Sinhala long vowels: the insertions depicted phonological integration in the borrowing language with the use of Sinhala suffix *aya*⁵. The pattern reflects word-internal mixing with the use of the bound morpheme from Sinhala in words such as *kolaraya* (colar), *loriya* (lorry), *rivolvaraya* (revolver), *pistolaya* (pistol) *packajaya* (package), *balloonaya* (balloon), *leegaya* (league), *boombaya* (bomb), *persiyaanu bokka* (Persian bay), *yookeraya* (the yoker), *parsalaya* (parcel). Note that when the English borrowed word ends with the alveolar nasal /n/, the Sinhala bound morpheme */anu/* facilitates integration as in *persiyaanu* in 'Persian'. Note that the same process is used to nativize English items such as Indian *Indiayanul* into Sinhala (11)⁶.
- b. English inanimate singular nouns ending in /r/ combined with Sinhala suffix *aya*: Examples *giyaraya* (gear), *tractoraya* (tractor), *galariya* (gallery), *thiyariya* (theory). Since the words show complete integration, it is plausible to analyze them as established borrowings. This is compatible with NP insertion (4).
- c. English inanimate singular nouns combined with Sinhala suffix *aya*: Example, *blousaya* (blouse). The word is a borrowing and is compatible with NP insertion. (1).
- d. English inanimate singular nouns combined with Sinhala suffix *u+va*: words such as *gazett+u+va* (gazette), *esthameent+u+va* (estimate) show complete integration in the matrix language. In these cases, /u/ is the suffix and /va/ create the singular form.

5. Sinhala words and bound morphemes are italicised.

6. Number of examples from the data

The words are phonologically and morphologically marked with the Sinhala long vowels and bound morphemes facilitating integration. The words show complete integration and can be analyzed as established borrowings. (2)

- e. English animate plural nouns with Sinhala suffixes and stress on word final /r/: the Sinhala plural marker suffix *waru* which is suffixed with Sinhala plural nouns like *aemathi+waru* (ministers), carry on the same process when confronted with English animate nouns in words such as *brigadier+waru*, *driver+waru*. These borrowings reveal the process of hybridization based on Kachru's theory and is compatible with NP insertion. The data from the written domain reveals contrasting strategies when compared with spoken code mixed data. For example, in mixed spoken contexts, the productive pattern of incorporating English plurals will be 'driver+s+/a, where both plural markers from the two languages facilitate mixing (Senaratne 2009). Significantly, mixes in spoken contexts project insertional CM strategies as opposed to the written context where the items are nativized (2).
- f. English inanimate plural nouns combined with Sinhala suffixes and stress on word final /r/: words such as *computer+wala*, *car+wala*, *video game+wala*, *tractor+wala*, *tyre+wala* show complete phonological and morphological integration in the written domain. Note that the English plural marker 's' is not retained in the data. The integration of the words is facilitated wholly by the Sinhala suffix. The words are borrowings due to complete integration, although part of the word is from English. The mixing process can be analyzed as hybridization, compatible with Kachru's unit hybridization pattern. Hence, here too a distinction can be drawn between spoken and written contexts where two mixing strategies are employed (5).
- g. English inanimate singular nouns with Sinhala long vowels and stress on /r/: the Sinhala long vowels are used to facilitate complete integration in the base language and create borrowings in words such as *kolaraya* (colar) (1).

- h. English item and Sinhala nouns as head: These hybrid compounds in Sinhala reveal complete phonological and morphological integration using the Sinhala long vowels and stress on word-ending /r/ facilitating integration in examples such as *cab rathayak* (a cab), *Interpol delak* (an interpol net), *twitter paniwidayak* (a twitter message), *twitter satahana* (a twitter message), *concrete kanuwak* (a concrete pole), *chocolate dawatanaya* (a chocolate wrap), *cyber prahara* (a cyber-threat), *make up wiyaparaya* (the make-up industry), *cab maaraya* (a cab killer), *police gnathiya* (a relative who is a police officer), *container maaraya* (a container killer), *public gatumak* (a public confrontation), *fancy karabu* (fancy earrings). (13)
- i. Nativized hybrid compounds: Words that revealed both phonological and morphological integration with Sinhala long vowels and stress on /r/ facilitating integration are *droonerr thaakshanaya* (drone technology), *tractoorr rathayak* (a tractor), *yooda balloonaya* (a huge balloon), *komis geahilla* (taking commissions), *premierr leagayata* (the premier league), *tractoraya* (hand tractor). (6)
- j. Sinhalizations with vowel prefixed consonant clusters: Sinhalizations usually occur in spoken data. However, there was evidence of Sinhalizations in the newspaper corpus. Observe the phrase *liskootiye beheth aedala* (has transported drugs in a scooter). The noun 'scooter' is Sinhalized. Sinhalizations are a result of nativization. Due to the significant influence of the L1, Sinhalized items create deviations in pronunciations and result in errors. In this instance, the item appears in written data. (1)

Patterns found in ‘unmarked’ English words written in Sinhala

- k. English pronunciation retained The corpus included English inanimate nouns with Sinhala plural markers where the pronunciation of the borrowed word was retained indicating code mixing strategies. In these instances, it is plausible to conclude that code mixing is used as a strategy since the relevant English item does not require nativization. Examples: bike+wala , race +wala, bag+wala, salon+wala, data+wala, site +wala, showroom +wala, video game +wala. (8).
- l. English animate nouns with Sinhala colloquial pronoun *kenek*: Generally, in Sinhala, *kenek* is used to convey respect as in *nilame kenek* (Gunasekera 1891). The data from the newspaper corpus too revealed words such as rider+*kenek*, traffic *mama+kenek*, colourful batsman +*kenek*, long-time member *kenek* (a long-time member). These can be analyzed as hybrid compound nouns indicating CM strategies. Examples are compatible with unit hybridization based on Kachru’s analysis. (4)
- m. English inanimate nouns with *eka* nominalizer: whenever the nominalizer *eka* follows a mixed English insertion, the structure reveals code mixing strategies. For the purpose of this study a few examples were selected to identify the common pattern in all the data. current *eka* (electricity) , game *eka* (the game), airline *eka* (the airline), free hit *eka* (a free hit), official ground *eka* (the official ground). (5)

Summary and conclusions

The analysis shows nativization as an extremely productive process when incorporating single English words in matrix Sinhala sentences in written contexts. In order to maintain monolingual standards in the mainstream media, many strategies have been consistently used in the process of integration. When adopting nativizing strategies, the writers of the news content have intentionally

reverted to Sinhala suffixes and plural markers to facilitate mixing. The Sinhala suffixes, stress on word-final /r/ and long vowels appear to be productive strategies employed to nativize English animate and inanimate items. The structural components of the hybrid nativized compounds contain an English element and Sinhala noun as the head. Even though single English words are a part of the hybrid, the influence of the L1 is visible in the stress on /r/ and the presence of long vowels. The occurrence of Sinhalizations appears to be minimum with just one example in the data. In addition, the Sinhala long vowels explicitly reveal variation in pronunciation patterns when incorporating single English words in matrix Sinhala sentences. The findings reveal the presence of a marked code when integrating single English words in matrix Sinhala sentences in written contexts. The marked code employs nativizing strategies to facilitate mixing. The unmarked code reveals code mixing strategies. Distinctive characteristics of the unmarked code are the use of the *eka* nominalizer, the retention of the original pronunciation of the mixed word and the use of the Sinhala pronoun. The study highlights the differences in mixing patterns used to integrate single English words in matrix Sinhala sentences and reveals nativization as a productive strategy in written contexts.

References

- Fernando, S. (1977) English and Sinhala bilingualism in Sri Lanka. *Language in Society*, 6: 341- 60.
- Fernando, C. (1985) *Changes in Sri Lankan English as reflected in phonology*. University of Colombo.
- Gair, J.W. (1998) *Studies in South Asian linguistics - Sinhala and other South Asian languages*. Oxford: Oxford University Press.
- Grosjean, F. (1982) *Life with two languages. an introduction to bilingualism*. Cambridge, Mass.: Harvard University Press.
- Grosjean, F. (1995) A psycholinguistic approach to code-switching. In Milroy and Muysken, 259-75.
- Gunesekara, A. M. (1891) *A comprehensive grammar of the Sinhalese language*. Colombo: Sri Lanka Sahitya Mandalaya.
- Hockett, C. F. (1958) *A course in modern linguistics*. New York: MacMillan Company.
- Hoffer, B. L. (2002). Language borrowing and language diffusion: An overview. *Intercultural communication studies*, 11(4), 1-37.
- Kachru, B. B. (1978) Code-mixing as a communicative strategy. In J. Alatis ed., *International dimensions of bilingual education*. Washington, D.C: Georgetown University Press, 107-24.
- Kachru, B. B. (1979) The Englisization of Hindi: language rivalry and language change. In I.Rauch, and G. F. Carr eds., *Linguistic method. Essays in honor of Herbert Peritz*. The Hague: Mouton, 199-211.
- Kachru, B. B. (1983) *The Indianization of English: The English language in India*. Oxford: Oxford University Press.
- Muysken, P. (1995). Code-switching and grammatical theory. In L. Milroy & P. Muysken (Eds.), *One Speaker, Two Languages: Cross-Disciplinary Perspectives on Code-Switching* (pp. 177-198). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511620867.009

- Muysken, P. (2000) *Bilingual speech: a typology of code mixing*. Cambridge: Cambridge University Press
- Myers-Scotton, C. (1993) *Social motivations for code switching: evidence from Africa*. Oxford: Clarendon Press.
- Poplack, S., & Meechan, M. (1995). Patterns of language mixture: Nominal structure in Wolof- French and Fongbe-French bilingual discourse. Cambridge University Press.
- Poplack, S., & Meechan, M. (1998). Introduction: How languages fit together in code mixing. *International journal of bilingualism*, 2(2), 127-138.
- Poplack, S., Sankoff, D. and Miller, C. (1988) The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics*, 26: 47-104.
- Sankoff, D. (1998). The production of code-mixed discourse. In *Proceedings of the 17th international conference on Computational linguistics-Volume 1* (pp. 8-21). Association for Computational Linguistics.
- Sankoff, D., Poplack, S. and Vanniarajan, S. (1990) The case of the nonce loan in Tamil. *Language Variation and Change*, 2: 71-101.
- Senaratne, C.D. (2009) *Sinhala-English code mixing in Sri Lankaa sociolinguistic study*. The Netherlands Graduate School of Linguistics. LOT publications.
- Senaratne, C.D. (2011) *Discourse Strategies of Sri Lankan Bilinguals*. Godage Publishers.
- Senaratne, C. D. (2012). The dominant language and mixing patterns. *Vistas Journal*. Open University of Sri Lanka.
- Weinreich, U. (1953) *Languages in contact, findings and problems*. The Hague: Mouton.