Code Switching in the Lexical Corpora of Children

B. A. Mahalakshmi Prasad
All India Institute of Speech and Hearing, Mysore, India
Email: machiprasad@gmail.com

Prema K. S. Rao
All India Institute of Speech and Hearing, Mysore, India

Abstract—Code switching is the hallmark of communication in a multilingual society. Code switching occurs in everyday speech at words, morphemes sentence and discourse levels. There are varied perspectives on the phenomenon of code switching where linguists have viewed code switching as inevitable in a plurilingual situation. Few believe that code switching helps to express meanings more precisely, others believe that code switching vitiates a language, rather than enhance the communication between individuals. In the Indian context extensive research on code switching in children has been undertaken which focus on how children use languages in the home setting with adults/parents or their siblings and suggest that bilingual children switch between language/s according to the cognitive demands of the tasks and the contextual demands such as participants and topics. The present study is designed to address code switching in children’s speech corpora. In the present study, the speech corpora elicited from six to eight-year-old children who are native speakers of Kannada language living in Mysore city but exposed to other languages by virtue of the multilingual environment both in their neighbourhood as well as in school, showed instances of code switching in their speech sample. The children switched languages while describing certain aspects of language like categories of nouns referring to vegetables, furniture, verb categories, and identification of colours. In the present study, the base language is Kannada, a language that shows highest number of morphemes in an utterance (Scotton, 1993). The study gives evidence for children’s ability from very early age to use linguistic codes from different languages as an equivalent of native language code depending on the context.

Index Terms—code switching, Kannada, English, spoken language, lexical corpora, children

I. INTRODUCTION

Code switching (CS) is one of the phenomena of spoken language that is observed in a multilingual society. It can be defined as the appearance of language contact and mixing, which include borrowing on the lexical and syntactic levels, language transfer, etc.

Early works in the area of code switching tended to consider sociolinguistic factors such as setting, topics, domain, participants, and language function etc. (Fishman 1967) proposes a domain-specific code-switching model wherein bilingual speakers choose which code to speak depending on where they are and what they are discussing. (Grosjean 1982) observed that the choice of a code used by a speaker in a particular context signals about group solidarity, ethnic identity etc. Tay (1989) observes that in multilingual societies common communicative strategies have evolved and speakers are not aware of their code switching behaviour during discourse. Myers-Scotton (1995) observes that speakers are aware of the range of codes that are appropriate in a particular context and the choice of codes used by these speakers.

A. Theoretical Framework for the Study

Linguists and Speech Language Pathologists have been fascinated with the phenomenon of code switching and have carried out vast research in this area. (Lance1975) was the first to question the linguistic constraint in the phenomenon of code switching and concluded that there was none. Lance’s theories were refuted by studies done by (Gumperz 1976), (Timm 1975), (Kachru 1977), (Pfaff 1976 & 1979), (Gibbons 1979) and (Woolford 1983).

B. Code Switching vs. Borrowing.

It is undisputable that the early traders and settlers from different regional backgrounds who settled in the former British colonies played a crucial role in the phenomenon of early borrowings from languages spoken locally into English and vice versa. These early lexical borrowings have established themselves in the native language and continued to be used by the native speakers rampantly. Borrowing, has been defined as "a performance phenomenon, not a learning process, a feature, therefore, of language use and not of language structure" (Corder 1992: 26). In other words, borrowing is a communication strategy, which compensates for missing knowledge during conversational
speech production. Researchers such as (Tarone 1983) have classified borrowing as a language learning strategy. Borrowing can be used by a single individual or by a speech community, and it usually affects the lexicon only; when the whole speech group incorporates a word from one language into the other, the borrowed word becomes firmly established in the speech. (Haugen 1950) in his model of lexical borrowing defines the resulting linguistic changes resulting from this phenomenon as 'importation' where the original form and pronunciation of a particular feature are retained as part of the transfer process from source language to the recipient language. 'Substitution' where the borrowed features changes in form or pronunciation utilising the morphemic and phonemic features that have been adapted leading to nativization of the borrowed word to make the borrowing conform to the native language restrictions. However, several researchers like (Poplack 2000), (Myer-Scotton 2004) and (Myers 2000) draw distinction between code switching and borrowing citing the reason that borrowed words are more likely to be integrated into the recipient language. (Myers-Scotton 2002) tries to give an explanation to this criterion by enumerating that if a borrowed word or switched word has a dictionary entry i.e. if the word is listed in the speaker’s mental lexicon then its status is undisputed. Borrowed words are assumed to be listed in the mental vocabulary of the speaker where as switched words are not, as the phenomenon of borrowing of words is more frequent and predictable than switching words.

Alternatively, (Poplack 1981), (Sankoff, and Poplack 1981), proposed two major syntactic constraints of code switching, i.e., 'Equivalence Constraint' and the 'Free Morpheme Constraint' on code switching. The equivalence constraint states that the switched sentences are made of integrated fragments of alternating languages, each of which is grammatical in the language of its origin. The boundary between adjacent fragments occurs between two constituents that are ordered in similar way in both languages, ensuring the linear coherence of sentence structure without omitting or duplicating lexical content (Lipski1977), (Pfaff 1979), (Myysken 2000). The equivalence constraint has been verified as a general tendency by innumerable studies conducted in Spanish-English -(Poplock1978; 1980; 1981); Finnish-English -(Poplock et al., 1978), Arabic-Finnish-(Barek & Sankoff, 1988), Tamil-English- (Sankoff et al., 1990), Fongbe-French-(Meechan & Poplock, 1995), Wolfo French- (Poplock & Meechan, 1995), Igbo–English -(Eze,1998), French- English -(Turpin1998) and Ukaranian-English-Budzhak-(Jones 1998).

Several studies have been conducted to specifically explore code-switching behaviours in children e.g., (Bauer & Montero, 2001); (Fantini, 1985); (McClure, 1977); (Saunders, 1982). Findings suggest that bilingual children switch languages according to the cognitive demands of the tasks and the contextual demands such as participants and topics. However, these case studies usually focus on how children use languages in the home setting with adults/parents or their siblings.

The present study is a part of the study taken-up for post-doctoral work that is aimed at establishment of spoken language corpus of children between the ages of six to eight, in Kannada language residing in the city of Mysore. It examines the phenomenon of code switching in the speech of six to eight year old children who are native speakers of Kannada language living in Mysore city. Spoken language data of six children was elicited through the elicitation of a story using standardised pictures, (Nagapoornima, 1990), narration and through administration of Computerized Linguistic Protocol for Screening (CLIPS), (Anitha & Prema, 2008). The speech samples were recorded on a digital recorder, transcribed, and coded for analysis. The data was analysed using Systematic Analysis of Language Transcripts (SALT) (Miller and Chapman 1982). In the present study the base language considered is the mother tongue of the children, a language in which the highest number of morphemes in an utterance is present (Myers-Scotton 1993).

The spoken language data of six children, whose mother tongue is Kannada and who go to Kannada medium schools was elicited through pictures in different contexts of story elicitation, narration and through administration of Computerized Linguistic Protocol for Screening (CLIPS), (Anitha & Prema, 2008). The following informant criterion was adhered to:

Criteria for selecting the children for the study
• Native speakers of Kannada language (Mysore dialect)
• Age range of 6-8 years
• Residents of Mysore
• Children attending Kannada medium school
• Children of parents who speak the same language / dialect

II. METHOD

The data was collected in the month of April 2010. The data was collected in the participant’s school. The school’s permission was sought before collecting the language samples from the participant. For the purposes of administering the test to elicit language data, each participant was seen individually in a room. The data was elicited in three steps.
• Picture description - CLIPS, (Anitha & Prema, 2004)
• Story telling - Standardised pictures, (Nagapoornima, 1990)
• Narration of experiences in school (Spontaneous)

1 Tarone’s Classifications of Communication Strategies refers to the strategy of “borrowing” as the process, which involves the learner either translating word for word from the L1 to express his meaning, or using the L1 to convey his meaning without bothering to translate.
The participant’s class teacher assistance was procured to fill out a questionnaire to collect the demographic details of the participants. The data was recorded on an Olympus digital voice recorder. The language samples were transcribed and morphological units were marked. The transcripts were analysed with the aid of SALT (Miller & Chapman 1981).

The story narration task involved showing the participant a standardised pictures, (Nagapoornima, 1990) that has a series of related pictures that can be used to narrate a story. The narration task involved asking the child the games he likes to play. (CLIPS) Computerized Linguistic Protocol for Screening (CLIPS), (Anitha & Prema, 2008) is a comprehension expression task that has pictures used in different linguistic variables such as semantics(lexical categories, antonyms, polar question, syntagmatic relationship, paradigmatic relationship, semantic similarity, semantic anamoly, semantic contiguity) Syntax(plurals, sentence types, PNG markers, transitives, intransitives, causatives, conditionals, conjunctions, comparitivites, quotatives, case markers, tenses, participial constructions)

The children switched languages while describing certain aspects of language like categories of nouns referring to vegetables, furniture, verb categories, and identification of colours. They also used words that have been nativised by the addition of the vowel ‘u’ to the borrowed word2.

The spoken language data was transcribed and marked according to the protocols of SALT software Systematic Analysis of Language Transcripts (Miller & Chapman 1981). The following linguistic measures were procured from the analysis of the data. The Mean Length Utterance (MLU) which is a measure of linguistic productivity in children. In the present study the base language is Kannada, the mother tongue of the children, a language that shows highest number of morphemes in an utterance (Scotton, 1993). Type-token ratio (TTR) is a measure of vocabulary variation within a written text or a person’s speech. The type-token ratio is a helpful measure of lexical variety within a text. Number of different words is the number of different root words the speaker has used in his speech. Code switching (CS) the number of instances of code switching seen in a transcript, and nativisation the number of instances of nativisation of words seen in a particular transcript.

<table>
<thead>
<tr>
<th>Sl. No</th>
<th>No.of Total Words</th>
<th>MLU</th>
<th>No.of Diff Words</th>
<th>TTR</th>
<th>CS</th>
<th>Nativisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transcript 1</td>
<td>389</td>
<td>31.9</td>
<td>242</td>
<td>0.62</td>
<td>67</td>
<td>6</td>
</tr>
<tr>
<td>Transcript 2</td>
<td>438</td>
<td>3.13</td>
<td>236</td>
<td>0.54</td>
<td>71</td>
<td>12</td>
</tr>
<tr>
<td>Transcript 3</td>
<td>364</td>
<td>3.22</td>
<td>218</td>
<td>0.60</td>
<td>42</td>
<td>8</td>
</tr>
<tr>
<td>Transcript 4</td>
<td>358</td>
<td>3.20</td>
<td>206</td>
<td>0.58</td>
<td>38</td>
<td>14</td>
</tr>
<tr>
<td>Transcript 5</td>
<td>354</td>
<td>3.25</td>
<td>188</td>
<td>0.53</td>
<td>42</td>
<td>6</td>
</tr>
<tr>
<td>Transcript 6</td>
<td>362</td>
<td>3.22</td>
<td>220</td>
<td>0.53</td>
<td>42</td>
<td>8</td>
</tr>
</tbody>
</table>

(MLU) Mean Length Utterance, (TTR) Type-token ratio, (CS) Code switching

The children code switched or borrowed words from English language while naming vegetables and fruits, furniture, profession, automobiles and certain verbs.

The participants used terms from English to describe the linguistic categories given in the tables below in table 1

<table>
<thead>
<tr>
<th>Address</th>
<th>Aunty, Miss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relation</td>
<td>Friend</td>
</tr>
<tr>
<td>Colour</td>
<td>Black, White</td>
</tr>
<tr>
<td>Noun</td>
<td>Pencil, Scale, Bed sheet, Book, Bucket, School, Tub, Bottle, Glass, Subject, Market, Bat, Kitchen, Bathroom, Room, Gym</td>
</tr>
<tr>
<td>Automobiles</td>
<td>Bus, Car, Train, Cycle, Auto, Scooter</td>
</tr>
<tr>
<td>Clothing</td>
<td>Sweater, Shirt, Pant</td>
</tr>
<tr>
<td>Profession</td>
<td>Tailor, Police, Doctor, Nurse, Teacher</td>
</tr>
<tr>
<td>Furniture</td>
<td>Table, chair</td>
</tr>
<tr>
<td>Food</td>
<td>Tea, Carrot, Papaya, Ice cream, apple, grapes</td>
</tr>
<tr>
<td>Verb</td>
<td>Cutting, Dance</td>
</tr>
</tbody>
</table>

A list of words that are nativised are given below in table 2

---

2 The rule of Kannada language is that words do not end with a consonant sound, all the words end with a vowel sound and therefore, people generally add ‘u’ to the English word and incorporate it into their speech. E.g Chairu, tabelu, glassu, pencilu etc.
The following types of code switching have been observed in the children’s speech.

Borrowing - is the use of a word from another language, which demonstrates morphological/phonological adaptation to the matrix language. Often it represents the appropriation of a term not available in the matrix language. (Myers and Scotton 2004) “dictionary criterion” which avers that if a borrowed word or switched word has a dictionary entry i.e. if the word is listed in the speaker’s mental lexicon then its status is undisputed. Borrowed words are assumed to be listed in the mental vocabulary of the speaker where as switched words are not as the phenomenon of borrowing of words is more frequent and predictable than switching.

A careful glance at the data reveals that the children switched code while speaking or naming certain categories. There exists an equivalent Kannada word for some of the most of the words used by the children. The children were aware of the particular equivalent Kannada word and yet chose to use the English words.

Eg: Iru cutting madtavare (a response to the picture of a barber cutting hair)
Cutting is an English word in Kannada the equivalent word is katarsu
The sentence is something like Iru kudlu katurstuiddare.
Yaaru hidyalla avaru outu (describing a game)
The one who doesnot catch, he is considered out.
The sentence in Kannada would be yaaru hidyalla avaru atadinda ache.

Equivalence constraint introduced by Poplack and her associates (Poplack 1980, 1981); (Sankoff and Poplack 1981) The Equivalence constraint has this form in (Poplack 1980). “Code switches tend to occur at point in discourse where juxtaposition of L1 and L2 elements does not violate a syntactic rule of wither language, i.e. at points around which the surface structure of the two languages map onto each other”. According to this simple constraint, a switch is inhibited from occurring within a constituent generated by a rule from one language which is not shared by another” However, the syntactic structure of Kannada and English languages are different from each other. Kannada language belongs to the family of Dravidian language and follows the SOV(subject-verb-object) sequence and English belongs to the Indo-European following the SVO(subject-verb-object) sequence. The Dual Structure Principle proposed by (Sridhar and Sridhar 1980) who studied the phenomenon of code switching between the Kannada and English languages produced their own constraint on linear ordering, restricting the point at which code switch constituent may being, but allowing for the possibility that a constituents internal structure differs from that of the host language. The Dual Structure Principle is as follows “the internal structure of the guest constituent (Embedded Language) need not conform to the

### Table 2

<table>
<thead>
<tr>
<th>Lexical Category</th>
<th>Nativised word</th>
<th>Transliteration</th>
<th>Transcription</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automobile</td>
<td>Basu Coru Railu</td>
<td>bassu koru ralu</td>
<td>bassu koru ralu</td>
</tr>
<tr>
<td>Profession</td>
<td>Teachers Taliwu</td>
<td>Tifsunu</td>
<td>tifsunu</td>
</tr>
<tr>
<td>Furniture</td>
<td>Tableki Chairu</td>
<td>TEBallu cErwu</td>
<td>tebalu tfuru</td>
</tr>
<tr>
<td>Noon</td>
<td>Scalei Pencilu Roomu</td>
<td>sKelnu pensilnu runu</td>
<td>skelu prinsalu rumu</td>
</tr>
<tr>
<td>Food</td>
<td>Iku creana, appleu, grapeyu</td>
<td>jis krimu, applelu, grapeyuA</td>
<td>jis krimu, applelu, grapeyu</td>
</tr>
<tr>
<td>Clothing</td>
<td>Shuru Pano</td>
<td>sharTsu pyaMtu</td>
<td>faru penatu</td>
</tr>
<tr>
<td>Colour</td>
<td>Blacku, whienu, greenu</td>
<td>blaku vaIlu grihu</td>
<td>blacku, waji, gri</td>
</tr>
<tr>
<td>Games</td>
<td>Volleybaliu</td>
<td>vaalibaalu</td>
<td>valaliu</td>
</tr>
<tr>
<td>Verbs</td>
<td>outu</td>
<td>ouTu</td>
<td>avtu</td>
</tr>
</tbody>
</table>
constituent structures rules of the host language (Matrix language). So long as placement in the host language obeys the rules of the host language.

1. Train inda biltovane
   He is falling from the train,
2. Dance madtavre
   They are dancing

It has been hypothesized that, if code-switching and borrowing are basically the same thing, it should be possible to formulate principles that are valid for both language contact phenomena. Well-known 'hierarchies of borrowability' (Haugen, 1953); (Muysken, 1981) can be applied to show that some types of constituents are switched more often, and more easily than others like the constitutes that represent food, furniture etc. Data also revealed that nouns constitute to a large number of switches. This finding corroborate to the findings of various studies involving different language pairs, for example (Pfaff 1979), (Poplack 1980,81) and (Timm 1975) studies on English-Spanish where nouns were found to comprise the highest number of switches (Berk-Seligson1986)

The reasons for borrowing generally can be ascertained to prestige as indicated by the presences of the words in the native language for the borrowed item the prestige perspective gives explanation to the borrowings of basic vocabulary items into the native language such as words related to categories of noun, verbs, profession etc. Therefore, the words that are being used by children in their speech is well established, rampantly used nativised structures of the borrowed words.

The demographic data revealed that children belong to the lower socio economic group and are seldom exposed to external influences of different languages and styles. The parents of most of the children are daily wagers and the sole motivation of the parents to send the child to school is the ‘mid-day meal’ scheme initiated by the Government of Karnataka.

**IV. SUMMARY**

In the present study, the examples demonstrate that code switching occurs in children’s speech. It also demonstrates that in a plurilingual country like India, pure monolingual speakers do not exist as evident from the instances of code switching seen in the language of children who go to Kannada medium schools, whose mother tongue is Kannada and are living in predominantly Kannada speaking neighbourhood.

The children switched codes from Kannada to English and rarely in Hindi. Further research is required in the code switching phenomena in children involving syntactically divergent languages like Kannada and English.

In the process of establishing a spoken language corpora of children more instances on code switching may emerge over time throwing light on the codes and the motivation for code switching by children.

**APPENDIX**

Apple, grapes
Kannada equivalent (Sebu, drakshi)
Hajam
(for a picture of a barber, Kannada equivalent shoura)
Bhaia
(for a picture of boy, Kannada equivalent anna)
Intrasentential- switching at the clause, phrase level, or at word level.
Yaaru hideyallu avaru out
(The one who does not catch, he is considered out)

**REFERENCES**


© 2011 ACADEMY PUBLISHER


B. A. Mahalakshmi Prasad is a student at the Department of Speech Language Sciences in the All India Institute of Speech and Hearing, Mysore. She was born in Mysore in 1979. She holds a Ph.D in Linguistics titled, ‘A study of communication through theatre, Scio Linguistic Perspective from Prescott- University, U.K in 2008 and a M.A in English Literature, University of Mysore 2002. Articles published by her- B.A. Mahalakshmi Prasad (2010) Instances of Code Switching in Indian Television Serials. Language in India; B.A. Mahalakshmi Prasad (2010) Code Switching in Kailsam’s Play-Poli Kitty. Language in India; B.A. Mahalakshmi Prasad (2010) Corpus Linguistics as a Methodology for Studies in Sociolinguistics. The Fifth Inter-Varietal Applied Corpus Studies (IVACS) group International Conference 2010. She is currently pursuing her POST-DOCTORAL FELLOW 2009-2011 at All India Institute of Speech and Hearing, Mysore. She also served as a RESOURCE PERSON at the Central Institute of Indian Languages, Mysore between 2003-2007. FACULTY DEPT. OF ENGLISH at the Mahajana Degree College, Mysore between 2007-2008. Her research interests are in the field of language teaching and language technology. Dr. Prasad is a member of Linguistic Society of India.

Prema K. S. Rao is a Professor at the Department of Speech Language Sciences in the All India Institute of Speech and Hearing, Mysore. Her academic achievements are as follows: Post-doctoral Study (Senior Research) on Prevention of Literacy Failures (Language-based learning disability) On Fulbright Fellowship by The Univ. of Virginia and Univ. of Toledo, Ohio, USA in 2005-2006; Advanced study in ‘Community Health Care and Research’ Kon Kaen Hospital, Thailand Awarded WHO Fellowship in 2005; Ph.D (Speech & Hearing) All India Institute of Speech and Hearing, Mysore Title of thesis: Reading Acquisition Profile in Kannada in 1998; She has over 26 years of experience in clinical, teaching, research and community service. Currently she is serving as the HEAD OF DEPARTMENT OF SPECIAL EDUCATION at All India Institute of Speech and Hearing, Mysore and PROFESSOR -DEPARTMENT OF SPEECH LANGUAGE SCIENCES at the All India Institute of Speech and Hearing, Mysore. She has published vastly in the area of Speech Language Sciences to list a few. Karanth P., Manjula R., Geetha Y. V. & Prema K. S. (2007). ‘/Asamartharigagi ondu tarabathi sadhana: ondu kiri sahayadondige/’ (With a Little Bit of Help… Early Language Training Manual’) in Kannada, Disability Training Aids, Books for Change, Action Aid India, India; Prema K.S (2007). Organization and Planning (course IV). Self learning material for Diploma in Hearing, Language and Speech, All India Institute of Speech and Hearing, Mysore, India; Prema K.S. Rao. , Prasitha, P. , Savitha, S. , Purushotham, P. , Chitra, R. , & Balaji, R. ‘Clinical Markers for Identification of Children with SLI’. In D. Vasanta (ed.) Special Issue on Clinical Linguistics in the Indian Context: Assessment of Language Disorders, Indian Journal of Applied Linguistics-Special Volume, 2010, Bahri Publications, pp.181-192. Her research interests include language acquisition, emergent literacy. Dr. Rao is an active member of Indian Speech and Hearing Association International Reading Association, USA, Dravidian Linguistics Association Fulbright Alumni Association Women in Science Indian Association for Preschool Education AIISH Alumni Association.