Coding Manual: Linguistic Hybridity

Linguistic Hybridity in Singapore Primary School Classes

National Institute of Education
RITA ELAINE SILVER and SITI AZLINDA AMASHA

Updated 13 August 2015
Contents

Introduction 3
Coding Development 4
Procedures 5
Coding Categories (i.e. ‘codes’) 6
Language 6
Chinese 7
English 7
Chinese and English 8
Codeswitching Analysis 8
Where (Silver & Bi, 2008) 8
When (Silver & Bi, 2008) 11
Function 12
Topic (Silver & Bi, 2008) 14
Issues for uncodable points in codeswitching 16
Localism/Globalism 16
Localism 17
Globalism 19
Singlish (Silver, 2008) 21
Count_Non-count noun 21
Null copula 21
Null subject 21
Null verb auxiliary 22
NULL/Plural marking 22
NULL Object 22
Uncodable Error! Bookmark not defined.
Acknowledgements 23
References 23
## Appendix A

<table>
<thead>
<tr>
<th>Relation (Ncoko, Osman &amp; Cockroff, 2000)</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>Establish rapport/Disrupt rapport</td>
<td>25</td>
</tr>
<tr>
<td>Include/exclude participation</td>
<td>26</td>
</tr>
<tr>
<td>Cooperation</td>
<td>26</td>
</tr>
<tr>
<td>Establish cooperation</td>
<td>27</td>
</tr>
<tr>
<td>Interfere cooperation</td>
<td>27</td>
</tr>
</tbody>
</table>
**Introduction**

This coding manual has been designed for the project “Linguistic Hybridity in Singapore Primary School Classes” (RS 6/11 RES) and refined for the project “Building a Corpus of Multilingual Language Use During Peer Interactions (RS 8/13 RES) (referred to as ‘this project’ in this manual since the same coding was used and the projects were linked in purpose), both funded by Office of Graduate Programmes and Leaderships, National Institute of Education. It was adapted and refined from earlier work done by Silver & Bi (2008). The examples featured here are drawn from the project CRP 48/03 ES which was funded by the Centre for Research in Pedagogy and Practice, National Institute of Education. This manual is a working copy used by the team for their analysis.

This project uses existing data of interactions among primary school students (peer) and student-teacher interactions in classrooms to investigate 'linguistic hybridity'. All of the data are from grade 3 students in Singapore, participating in extra school lesson after school for one week: one lesson per day, one hour per lesson.¹ Based on preliminary analyses and prior research on language use in multilingual contexts, the project examines the premise that language use in classrooms with multilingual students often cannot be categorically determined as one distinct variety (e.g., English/Chinese/Singlish) but as a hybrid variety with its own interactional competence.

What is 'linguistic hybridity'? Although, to date, it has not been well-defined, the term has been loosely used to signal the creativity of English in multilingual contexts, partly in contrast to (and as critique of) a World Englishes approach that relies on Ferguson’s (1959) notion of diglossia, or the separation of language varieties based on functional and contextual domains. Hybridity allows for a plural view of English, a creative blending of languages. (See, e.g., Rubdy and Alsagoff [2014] and Manyak [2001] on hybridity; cf., Garcia [2009] on translanguaging; Piccardo [2013] on plurilingualism.) Our assumption is that language use is finely-tuned to a variety of social and situational factors, thus fluid in its representations.

**Objectives:**

The purpose of the project “Linguistic Hybridity in Singapore Primary School Classes” (RS 6/11 RES) was two-fold. Firstly, to explore classroom language use in Singapore primary classrooms in order to better understand how students use their linguistic repertoires to represent themselves and their meanings as they adjust to a variety of social-contextual factors. Secondly, to investigate

---

¹ A comparable set of day for students in primary grade 1 exists. Students in primary grade 1 used the same tasks with the same teacher and same time frame. Those data have not been analysed for this project due to the complexities found with analysing for hybridity in the grade 3 data set. An analysis of grade 1 has been done using corpus software developed for CRP 48/03 ES.
the relevance of a framework of linguistic hybridity for Singapore's multilingual/multidialectal environment. The purpose of the follow up project, “Building a Corpus of Multilingual Language Use During Peer Interactions (RS 8/13 RES), was also two-fold. Firstly, to finalize coding of a data set of 31 transcripts from P3 children engaged in peer work using the concepts in this manual, and to prepare a set of transcripts for submission to an international database of child language (CHILDES, http://childes.psy.cmu.edu/).

Coding Development

Throughout the coding process, the first coder and the principal investigator (authors of this manual) discussed the scope and definitions of the codes. The coding system was first used to code a sub-set and to establish draft definitions and examples. Subsequently, a second coder independently coded the same sub-set of data using the same coding system. An inter-coder agreement report using Cohen’s Kappa for each code in each transcript was generated and problematic codes were discussed with the principal investigator and both coders. Coding agreement was re-checked and found to be sufficiently high (.70 or higher) for most instances.2 ‘Instances’ refers to coding at the level of overarching categories (e.g., Codeswitching) as well as embedded levels of coding (e.g., Codeswitching\Topic\Comments about the task). Any remaining disagreements, at any level of coding, were resolved through discussion between the PI and the first coder.

Notes about codes which were problematic due to the specific context of the study (e.g., ‘Globalism/Western forms of address’) were retained if possible to verify whether these would be useful in the analysis. Where necessary, coding definitions were adapted. For example, Codeswitching/Personalisation and objectivization, ‘personalization’ and ‘objectivization’ (Gumperz, 1982) were adapted as we found that students were not using the language at this level of

2 Landis and Koch (1977) list .61 - .80 as ‘substantial agreement’. Fleiss (1981) suggests that .40-.75 is ‘fair to good’ and .75 and above is ‘excellent’. However, we note that for this data set the number of samples is important. In some cases there are very few instances to be coded leading to either very high or very low agreement (e.g. if there are only two instances and the coders disagree on one of the two, agreement is very low; if coders agree on both, agreement is very high). One other issue is that NVivo calculates Cohen’s Kappa for each code and sub-code for each transcript based on the number of characters — not the number of codes and not computing the overall coding accuracy (across transcripts). Therefore, we also consider not only the .70 ‘sufficient agreement’ but looking across transcripts whether that level of agreement is usually established. For example, for Codeswitching > When > After there was some variation:
• on the transcript P3.D3.C1.WH+ZY.final, Cohen’s Kappa was .9761
• on the transcript P3.D3.C1.XN+RT.final it was .5783
Judging by the scores from Cohen’s Kappa overall and the fact that P3.D3.C1.WH+ZY had only one instance of codeswitching, the agreement was considered to be sufficient.
sophistication. Coding for ‘Singlish/Count and non-count nouns’ was revised to include only instances when non-count nouns are used as count nouns e.g., furnitures. In addition, the definition for ‘Codeswitching/Qualification’, was amended to better represent what we found in our data considering use of Singlish. The original definition by Gumperz (1982) was problematic as the copula verb is often dropped in Singlish. Two proposed categories were found to be unsuitable for our data set and were removed from the coding scheme. These are discussed in Appendix A. For codes which continued to show considerable disagreement due to different interpretations by coders, definitions were discussed and refined with examples added to the manual.

Finally, to facilitate coding accuracy, two coders (first and second coder) proceeded with independent coding in specialized areas: the first coder – a Singaporean, Malay-English bilingual, coded for Speaker, Language, Globalism/Localism, and Singlish. She also did some coding on Codeswitching/Where and Codeswitching/When which are relatively self-evident. The second coder – a Singaporean, Chinese-English bilingual, did all other coding for Codeswitching as well as verifying translations in the transcripts. The coding was finalized by the PI, in discussion with the first and second coders.

**Procedures**

Coding of data is carried out using NVivo qualitative research software (QSR International, 1999-2013). All coding is done following the categories as described below. Questionable points in the data and/or points that might be of interest but did not fall within the coding scheme were annotated for discussion by members of the research team.

The aim of the coding is to surface instances that best represent the codes. Hence, a conservative, rather than comprehensive, approach is adopted. This means that we do not attempt to code all utterances; instead, the focus is on coding clear and relatively uncontroversial instances of each code.

The unit of analysis varies across the individual codes due to their differing nature. We base our decision on the unit of analysis according to what seems most commonsensical for each code. For example, for ‘Codeswitching’, whole turns are coded; on the other hand for ‘Localism_Use of pragmatic particles to express interpersonal meanings’, only the pragmatic particles (e.g., ‘lah’ and ‘meh’) are coded. Details are given in the explanation for each code.

Multiple coding is applied to the data. In other words, for the same turn, more than one code is applied if appropriate. Coding is applied to the sub-categories and not at the main category. For
example, a whole turn that is in English is coded at the sub-category ‘English’, not at the main category ‘Language’. The sub-categories under each main category are then aggregated.

**Coding Categories (i.e. ‘codes’)**

**Speaker**

Each turn is coded according to who the speaker is. For example, all turns by the teacher are coded as ‘Speaker: Teacher’³ and the entire turn is selected for coding, as shown in Example 1:

Example 1

```
T: Ok, so you can get started. And the time is now twelve-forty so fifteen minutes later I will stop you. (Pause)⁴
```

(P3.D1.C1.TW_TJ_XN_RT_XL.final)

If the transcript notes XXX or (inaudible), the turn should still be coded for speaker as this indicates that the speaker said something, although it was unclear. So, Example 2 would be coded as Speaker > TW.

Example 2

```
TW: (Unclear) XXXXXX go after this.xxx (TW speaks so softly)
```

(P3.D1.C1.TW_TJ_XN_RT_XL.final)

Similarly, when the transcript shows that a response was elicited from a specific person, but the person did not reply, the turn can be coded. In Example 3, turn 2 by XN as coded as Speaker > XN.

Example 3

```
1. T: So what do you think, XN, of the idea? You are very happy, whatever they’re doing?
2. XN: (No audible response)
```

(P3.D1.C1.TW_TJ_XN_RT_XL.final)

³ Although NVivo software can automatically code for speaker, our transcripts were originally created for a different project using different software. Therefore the auto-code feature of NVivo could not be used and manual coding was done.

⁴ Highlighted text indicates the area for which coding was applied.
Where the students reply in unison, the transcripts is not coded for speaker as there are no individual speakers in these turns and these are usually replies to teacher instructions.

**Language**

Each turn is coded according to which language is used, as explained below. Where an entire turn is marked as ‘inaudible’ or with ‘X’ (or multiple XXX) in the transcript, the turn can be marked for speaker, as above, but not for language. (Since the turn is inaudible, we cannot be sure which language was used.) Similarly, if the entire turn is (Sigh) or (laughter) or some other audible signal that is not language specific, the turn is not marked for language.

‘Huh’ and ‘ok’ are ignored for coding language as it is not possible to determine if ‘huh’ is distinctly Chinese or English. ‘Ok’ has become so ubiquitous that Chinese speakers report it is used even among those who are not English-users. Therefore it is also not considered. In a turn, code for whatever language is used along with ‘Huh’ or ‘Ok’. Following this logic, Example 4 is coded as Chinese.  

Example 4

TJ: Huh? 什么? (What?)

The entire turn is selected for coding.

**Chinese**

This refers to instances when only Chinese is used within one turn (Example 5).

Example 5

JG: 一直讲好好好，给你气死了。(Keep saying you are done done done, you’re making me mad.)

**English**

This refers to instances when only spoken English is used within one turn (Example 6).

Example 6

ZY: (Chuckles) Eh! Ssssoccer! Where you?!

---

Arguably this misses some potential reiteration as part of code-switching.
Chinese and English

This refers to instances when both Chinese and English are used within one turn as in Example 7.

Example 7

TJ: sleeping hai shi (or) swimming? 6

Codeswitching Analysis

This analysis refers to instances of using multiple languages within or across turns.

Codeswitching is defined as “The juxtaposition within the same speech exchange of passages of speech belonging to two different grammatical systems or sub-systems” (Gumperz, 1982, p. 59). For this analysis, the two grammatical systems or sub-systems are the languages English and Chinese as used by the students in their peer interaction.

Only peer-to-peer interactions are coded – interactions involving the teacher are not included. 7 In addition, no distinction is made between codeswitching and codemixing. For each of the sub-category of Codeswitching, entire turns were coded.

Where (Silver & Bi, 2008)

This code marks codeswitching within a turn or at the turn boundary. This prevents selection errors or seeming coding disagreements (such as one coder selecting a ‘space’ as well as a word while the other coder selects the word without the space – NVivo software ‘counts’ this as a coding disagreement).

Turn Boundary

This is when codeswitching happens when there is a new turn, i.e. at a turn boundary with a new speaker as in Example 8.

6 Some original versions of the transcripts used hanyu pinyin. In the final versions of the transcripts, all have been changed to Chinese characters with English glosses. We have not made similar changes in this working manual.

7 There were no instances of use of Chinese or codeswitching in interactions between teacher and student(s).
Turn 4 by JG is annotated for ‘codeswitching/turn boundary’ as the preceding spoken turn (Turn 2, by WH) ends in English. TW’s chuckle (Turn 3) is not considered as it is not a spoken turn and is not included in the analysis. Turn 4 is in Chinese, following Turn 2 (since Turn 3 is not a spoken turn) and so is coded as ‘codeswitching/turn boundary’. At Turn 5, ZY speaks in English. Since this follows JG’s Chinese speech, this is also ‘codeswitching/turn boundary’. So, in this excerpt there are two instances of ‘codeswitching/turn boundary’: Turn 4 and Turn 5.

Within Turn
This related to codeswitching which occurs within a single turn (Example 9).

Example 9

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>TJ: Number 1, computer, ok, start ah, xian zai (now)...</td>
</tr>
<tr>
<td>2.</td>
<td>JG: I don’t have blue sticks leh?</td>
</tr>
<tr>
<td>3.</td>
<td>TJ: Huh?</td>
</tr>
<tr>
<td>4.</td>
<td>JG: Wo mei you (I don’t have) blue stick lah, ni chau si ren (you’re being too noisy)...</td>
</tr>
<tr>
<td>5.</td>
<td>TJ: He says he don’t have blue...tag</td>
</tr>
<tr>
<td>6.</td>
<td>JG: I don’t have blue stick.</td>
</tr>
<tr>
<td>7.</td>
<td>T: Yeah, yeah...I’ll give to you...</td>
</tr>
</tbody>
</table>

In turn 1, TJ uses English initially and then switches to Chinese. This is codeswitching/within turn. At Turn 4, JG begins in Chinese then switches to English, then switches to Chinese again. Thus, in Turn 4, there are two instances of codeswitching/within turn.

---

8 In a few cases transcribers tried to capture pronunciation through spelling. This was not done consistently and is not a factor in the analysis.

9 In the final transcripts, all Chinese has been changed to Chinese characters, although this manual shows a mix of characters and pinyin. Chinese in pinyin are written in italics.
**Turn Boundary and Within Turn**

There are also instances when one turn by a speaker is coded for both codeswitching/within turn and codeswitching/turn boundary (Example 10).

**Example 10**

1. JG to XL: Oval... *wo you leh* (oval... I have)
2. XL to JG: Ah nah, *ni xie oval ni you table...* (Write down that you have an oval table)
3. JG to XL: *Wo you liao.* (‘I’ve got it’)
4. XL to JG: *Oval hai shi rectangle? Bu yi yang de leh!* (Is it oval or rectangle? It’s different!)

In Example 10, when XL speaks to JG in Turn 2, ‘ah nah’ is not considered as either English or Chinese since it could be used as part of the localized variety of English. (See the section on Globalization/Localization.) Therefore, there is no codeswitching/turn boundary from Turn 1 to Turn 2. However, there is codeswitching/within turn with ‘*ni xie*’ in Chinese, followed by ‘oval’ in English, followed by ‘*ni you*’ in Chinese and then ‘table’ in English.

Turn 3 then is codeswitching/turn boundary because the final word of the previous turn was in English (‘table’) while JG’s utterance is in Chinese ‘*wo you liao*’. There is another codeswitch/turn boundary in Turn 4 as XL replies in English (again the word ‘oval’). Since XL continues in Chinese (‘hai shi’), then English (‘rectangle’), then Chinese, Turn 4 is coded for codeswitching/within turn and for codeswitching/turn boundary (from Turn 3 to Turn 4).

In Example 11, we see Codeswitching/turn boundary at Turn 2 because TJ ends with English and JG starts with Chinese, ignoring ‘huh?’ which is uncodable. Turn 2 is not coded for ‘within turn’ because ‘huh’ cannot be clearly established as English or Chinese. Similarly, Turns 3 and 4 are not coded for code switching/turn boundary. Because ‘huh’ can be either English or Chinese – it is not clear where the switch occurs (Turn 2 to Turn 3 or Turn 3 to Turn 4?).

**Example 11**

1. TJ: Number 3. Number 3 she is also swimming
2. JG: *Huh?* 什么? (*What is it?*)
3. TJ: Huh?
4. JG: A girl is swimming ah?

(P3.D2.Cycle3.JG+TJ.final)

---

10 In the final transcripts the translations are more exact as to which words where in Chinese or English.
Uncodable Utterance for Codeswitching

In some instances, there is codeswitching which cannot be coded due to issues in the transcription. In Example 12, Turn 2 is marked as (Non-English) by the transcriber. This might have been Chinese but of a different variety (e.g. Hokkien) or it might have been a different language such as Malay. On the other hand, it might be a transcription error – an instance which should have been transcribed and translated for Chinese but was missed. Since information on the exact language is not given, we cannot effectively code for codeswitching. Another issue is when there is codeswitching but part of what is said, or the subsequent turn, is unclear/inaudible. In these cases, we cannot accurately code for codeswitching. These turns are marked as ‘uncodable’ so a count of such turns can be calculated.

Example 12

TJ: Ok. 好了。 (It’s done)
JG: (Non-English) One...boy is...uh...one boy is...eating ice cream

When (Silver & Bi, 2008)

This deals with codeswitching instances that occur at different stages of the activity within the lesson, i.e. before the activity commences, during the actual activity and after the activity has been completed or has stopped. For this analysis, the coder must read through the entire transcript to understand what the activity is and where it begins/end. Typically, annotation for this will be in single or multiple turns depending on the codeswitching patterns. This category has to do with the timing of the codeswitching in relation to the task and the lesson as a whole.

Before
The codeswitching turns that occur before the task set by the teacher begin (as shown in Example 13, where part of the transcript that has been coded for ‘before’ is reproduced).

JG: Jiao ni bu yao wan, deng yi xia lao shi zai gei . (told you not to play with it, the teacher will give it to you later)

During
Example 14

JG: Camp好了, camp好了才去 hotel 冲凉吗. (Camp finish, camp finish then go hotel and bath mah.)
After

Example 15

JG:  Wo qu check-up ah, deng zhe me jiu, wo ma ma hua zhe duo qian ah. Zhen de leh, ni bu zhi dao leh. (I went for a check-up, waited for a long time, my mother has to spend a lot of money. It’s true, you know.)

Transcripts are organised based on the lesson cycles where one lesson cycle comprise one transcript. However, for some of the transcripts, two consecutive lesson cycles have been included in one transcript. For these transcripts, students completed the task of the first lesson cycle followed by a lag time where they waited for the next lesson cycle to commence. During this time, they usually engaged in small talk of a social nature. This will be coded as ‘after’. When the next lesson cycle commences, this will then be coded as ‘during’.

Coding for CS/When

<table>
<thead>
<tr>
<th>First lesson cycle</th>
<th>Lag time</th>
<th>Second lesson cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td>During</td>
<td>After</td>
<td>During</td>
</tr>
</tbody>
</table>

Where the lesson cycle is the same throughout the transcript, any ‘lag times’ occurring in the middle of the lesson should be treated as ‘during’. After the teacher has ended the lesson and collected the materials, then the ‘lag times’ would be treated as ‘after’.

Function

Gumperz claims that “switching serves roughly similar functions in different situations, so that a single preliminary typology can be set up which holds across language situations” (1982, p.75). We adopt four categories of the communicative functions from Gumperz (1982) as these four categories are the most frequently attested by the current data from our preliminary analyses. These four functions are Reiteration, Qualification, Personalization and Objectification. Instances of codeswitching that cannot be categorized into these functions but seem to be functionally-oriented are coded as ‘Other’ in the initial analysis, for re-analysis later.

Personalization and Objectivization

“The code contrast here seems to relate to such things as: the distinction between talk about action and talk as action, the degree of speaker involvement in, or distance from the message; whether a statement reflects personal opinion or knowledge, whether it refers to specific instances or has the authority of a generally known fact.” (Gumperz, 1982, p. 80)
For the purposes of this analysis, we have adapted Gumperz’s definitions.

**Personalization** refers to instances when personal opinions and experiences are stated (e.g., I think..., I went to ...), as in Example 16.

**Example 16**

JG: [After dis] -- Hotel then go home.  
TJ: 我擦指甲. (I applied nail polish.)  

**Objectivization** refers to instances when objective information is shared (e.g., teacher says..., the map says...), as in Example 17.

**Example 17**

JG: Wo mei you (I don’t have) blue stick lah, ni chau si ren (you’re being too noisy)...
TJ: He says he don’t have blue...tag ['blue tac']

**Qualification**

“A large group of switches consists of qualifying constructions such as sentences and verb complements or predicates following a copula” (Gumperz, 1982, p.79). For example,

The oldest one, *la grande la de once anos* (the big one who is eleven years old). (ibid)

Applying this definition to our data is problematic as Singlish constructions are often without copulas, as seen in the Example 18 given, or without subjects. The definition has therefore been adapted to suit our data. This category will only apply to constructions that qualify or extend an idea presented in a noun of one language, using another language for the qualification. The qualification could be sentences or phrases. Further, in cases where the subject is implicit, coders will take into consideration the implicit presence of a noun or noun phrase which is retrievable from context – where it is being qualified by way of extension of an idea and the speaker codeswitches, the utterance will be coded as ‘qualification’.

**Example 18**

JG: ah, roller-blading...hao liang ah ... eh...eh...eh (It’s cooling)

---

11 A common classroom material for temporarily sticking items together or sticking things on the wall.
Reiteration

“A message in one code is repeated in the other code, either literally or in somewhat modified form. In some cases, such repetitions may serve to clarify what is said but often they simply amplify or emphasize a message” (Gumperz, 1982, p. 78). In Example 19, JG first called XL ‘stupid’ in English in turn 2. In turn 4, he amplified this by calling XL ‘you stupid fool’ in Chinese. Turn 4 is therefore coded as ‘reiteration’.

Example 19

1. JG to XL: Write down lah
2. JG to XL: Eh, write down lah, stupid!
3. XL to JG: What are you saying?

For our analysis, codeswitching/function/reiteration can be a language switch by the same speaker or a different speaker, as long as the message conveyed is the same but in different languages.

Other functions

All the turns coded for codeswitching are also coded for the sub-category ‘function’. However, some of these do not fit into any of the categories of function as stated above. The sub-category ‘other functions’ is therefore created to account for such instances. In many cases, utterances coded as ‘other functions’ are distinguished in the analysis for ‘Topics’. For example, our preliminary analysis suggests that Chinese is preferred when discussing social topics or talking about how to do the task, while English is used when students are actually performing the task or discussing school matters. Subsequent analyses might attempt to tease other ‘other functions’ distinct from ‘topic’ but for the initial analysis, codeswitching which is not coded for one of the functions above should be coded as function/other.

Topic (Silver & Bi, 2008)

Coding is done based on the topic of discussion in the codeswitching segment vis-à-vis the assigned task.

Procedures for tasks

This refers to discussions of how to do, carry out, complete the task (Example 20).

Example 20

TJ to JG: JG 你要 describe 给, 你要 describe .... (JG you must describe to, you must describe...)
JG to TJ: Eh, you have a ... you have a lamp?

Comments about performing the task
This refers to talk about performing the task that occurs before or after actually enacting the task (Example 21).

Example 21

JG: *Wei shen me bu yao gei wo kan ni de?* (Why don’t you let me see yours?) (pause) I do finish. I do finish.

JG made this comment after the group had finished doing their task. More importantly, it is not intended to facilitate enacting or completing the task. Rather it is JG’s comment about (another student’s) performance of the task.

Enacting tasks
This refers to codeswitching used in order to accomplish the academic task set by the teacher (Example 22).

Example 22

JG: Ok. One girl is...drawing

TJ: Huh? Er, *gang cai ni shuo de computer de laeh? Computer de?* (What about the one on the computer?)

School-related topics
This refers to any other discussion related to the class or school, but not specifically about the task. In Example 23, XL’s utterance is about the regulations of class behaviour in general (sit properly) and rules specific to the task (don’t try to cheat).

Example 23

XL to JG: Ok lah, why like that one?

JG to XL: *Hei, hen hao xiao leh, (it’s funny) I tell you.*

XL to JG: *Zuo hao ah ni, bu yao zuo bi.* (Sit properly, and don’t try to cheat)

JG to T: Teacher, this chair.

Social topics
This refers to topics of personal or social nature including jokes and insults (Example 24).
Example 24

JG: *Wo mei you* (I don’t have) blue stick lah, *ni chau si ren* (you’re being too noisy)...

(P3.D2.Cycle2.JG+TJ.final)

**Issues for uncodable points in codeswitching**

As above, interjections such as ‘huh’ and ‘ah’ are not considered for codeswitching. In Example 25, a snippet of conversation between two students, Turn 1 is coded as ‘codeswitching/within turn’. The second turn ‘Huh?’ is not distinctly English or Chinese and therefore is uncodable. Thus, Turn 2 would not be coded for codeswitching.

Example 25

TJ: Haven’t. *Kuai dian*. (hurry up)
JG: Huh?

(P3.D2.Cycle1.JG+TJ.final)

Similarly, pragmatic particles (e.g., ‘lah’, ‘lor’, ‘meh’) are not coded for codeswitching because they can be part of Chinese but also part of Singapore English. Therefore in Example 26, JG’s turn has been coded for codeswitching/within turn but the ‘lah’ is not a significant utterance for determining the codeswitch boundary.

Example 26

JG: *Wo mei you* (I don’t have) blue stick lah, *ni chau si ren* (you’re being too noisy)...

**Localism/Globalism**

This category has to do with English language use only. While some researchers have identified localised features of English in Singapore as ‘Singlish’, this tends to consider localization as non-standard (e.g., Platt & Weber, 1980). Other analyses of Singlish see Singlish as being diglossic – non-standard but used in functional contrast to a standard variety (e.g., Gupta, 1994). Our analysis of the category ‘Singlish’ takes into account some of the most commonly documented exemplars, as above. However, in keeping with our hypothesis that student language use represents functional competence in a hybridized variety we look beyond exemplars of Singlish and consider possible implications of localism and globalism as posited by Alsagoff (2007; 2010a, 2010b).

Alsagoff (2010a) hypothesizes that speakers of English in Singapore vary their use of the language by shifting stylistically between global and local orientations. This shifting is described as “fluid and
hybrid” (2010a, p. 115) as speakers move along the plane of variation in degrees rather than fixed points at the extremes in negotiating their identities. She suggests further that speakers orientating towards the global forms use linguistic features that are affiliated with global characteristics such as “formality, educational attainment, institutionalism and authority” (Alsagoff, 2010a, p. 117).

Speakers orientating towards the local forms use linguistic features associated with group membership and building rapport. Alsagoff’s original conception was intended to rely on accurate description without comparisons with other norms. However, in some instances (e.g., Globalism_Absence of endogenous lexical items or meanings/ Localism_Inclusion of endogenous lexical items) contrasts are built into the analysis.

As with all coding for this project, coding is done at the level of sub-features and aggregated to ‘Localism’ or ‘Globalism’. Only the relevant feature is selected which in some cases is an individual word, a phrase, or a full term (details below). However this means that it is possible for utterances which are primarily ‘global’ (i.e. in syntax) to be coded as ‘local’ due to isolated items (lexis, use of pragmatic particles). For example, ‘What is that lah’ is coded as ‘localism/pragmatic particle’ though it shows the global features of wh-fronting.

Localism

Localism_Inclusion of endogenous lexical items or meanings

Speech that uses local slang, colloquialisms or lexical items. This excludes pragmatic particles, which are coded separately, and proper nouns. Coding is done at the level of the lexical item, whether a word or a phrase. In Example 27 the pragmatic particles (‘orh’ and ‘neh’ in this example) are not coded for ‘Localism_Inclusion of endogenous lexical items or meanings’. However, the phrase ‘Die Alamak die’ is.

Example 27

H: Orh, next morning next morning. Ah, ten.
ZY: Eh ten neh?
JG: (Mimics) Ten neh?
ZY: (Chuckles)
WH: Forgot. Die Alamak die.

(P3.D1.C1.ZY_WH_JG.final_1)

Localism_Use of pragmatic particles to express interpersonal meanings

Examples of these are oi, lah, lor, leh, neh, nor, nah, orh hor, mah?? and meh (Example 28). Coding is done at the level of the lexical item.
Example 28

JG: (Defensively) I also cannot see lah!

(P3.D1.C1.ZY_WH_JG.final)

When ‘lah’ is used with English, we consider this to be localism. When it appears with Chinese (Example 29), we consider it to be Chinese so it is not coded for Localism/Globalism.

Example 29

ZY: 不要讲华语lah. (Don’t speak Chinese lah.)

(P3.D1.C2.all.final_2)

Localism Use of ethnic Asian forms of address

Due to the nature of the data (peer-peer, in classroom settings), there are few forms of address overall. This category could include ‘Teacher’ or ‘Cher’ to address the teacher. However, since the focus of our analysis is on peer-peer talk exclusively and talk to/from the teacher is excluded, this category is unlikely to occur in our data. Coding is done at the level of the form of address (i.e. word or phrase).

Localism Conditional clauses expressed without subordinating conjunctions

Conditional clauses expressed without explicitly stated subordinating conjunctions. Coding is done at the level of the sentence. No examples found in our data.

Localism In-situ wh-question words in interrogative clauses

Interrogative clauses with wh-words in-situ (Example 30). Wh-words positioned at the end of the clause should not be coded for this category (e.g. The television have what?). Coding is done at the level of the clause.

Example 30

TW: The boy hiding behind the sofa wear what shirt?


Localism Reliance on adverbs or contextual information to indicate time of event

Use of adverbs or contextual clues to indicate time in conjunction with unmarked verbs (Example 31). This is coded at the level of clause.

Example 31

RT: 我带两百块就够了. (I’ll bring two hundred dollars and that’s enough already.)
Globalism
This has to do with linguistic features that can be characterised as “global” English.

Globalism_Absence of endogenous lexical items or meanings
Speech that does not make use of local slang, colloquialisms, lexical items (Example 32).

Example 32

JG to TJ: Eh, you have a ... you have a lamp?

The coding process for this category is unique since endogenous items are coded under ‘Localism_Inclusion of endogenous lexical items’. This means, by definition, any items which are not endogenous would fit in this category. Rather than manual coding, results on this category are found by using the query function of NVivo after coding for ‘Localism_Inclusion of endogenous lexical items’, ‘Localism_Use of pragmatic particles’ and ‘Localism_Asian forms of address’ are complete.

Note that informality is not a consideration – informal lexical items can be considered as part of Globalism_Absence of endogenous lexical items or meanings. Grammatical structure is also not considered for this category. In Example 33, the phrase structure suggests localism but the lexical items do not. Therefore, Example 33 would be considered as Globalism_Absence of endogenous lexical items or meanings.

Example 33

ZY: Want to Go to Merlion.

Globalism_Conditional clauses linked by subordinating conjunctions
Conditional clauses explicitly linked by conjunctions like ‘if’, ‘unless’, ‘provided (that)’ and ‘as long as’ (Example 34).

Example 34

X: half an hour five dollar.
ZY: No need
TW: ZY, if you do, then you, then I will have ...
ZY: (unclear).
Globalism. Fronted wh-question words in interrogative clauses

Wh-question words at the start of interrogatives with subject-verb inversion (i.e., the verb is positioned before the subject) or with use of an auxiliary verb without inversion (e.g., ‘do’ ‘What do you have?’). Constructions that do not include either a subject or a verb or both would not be coded for this category. In Example 35, the clause ‘What color is your chair?’ has been coded for this category as it starts with ‘what’ and the verb ‘is’ comes before the subject ‘your chair’.

Example 35

TW: Your ... your ... your chair what color? What color is your chair?

Globalism. Use of disjuncts and parenthetic comments to express interpersonal meanings

These are words or asides that convey attitude/mood e.g., ‘honestly’, ‘clearly’, ‘in my opinion’, see Example 36.

Example 36

ZY: XXXX turkey one hundred dollars?
WH: All our money, all our money (chuckles)
WH: Wah, my god! [One hundred dollars, one fif --!]
T: Do you all agree, [to what they are saying?] is that ok?

Globalism. Use of interrogative mood for expressing questions

These are questions in the interrogative mood, where there is an inversion in the subject and auxiliary verb (Alsagoff, 2010a), as in Example 37.

Example 37

TW: Is your television on or off?

Globalism. Use of verbal inflections (tense) to indicate time of event

Use of present, past or perfect tense to indicate time of event (code verb forms with inflections), as in Example 38. Coding is at the level of word.

12 Alsagoff (2010a) also includes ‘Use of declarative mood for expressing questions’. We have not included this category as part of our analysis as interrogatives expressed orally with intonation rather than grammaticization are features of informal, rather than localized, English.
Example 38

G: wo mei you (I don’t have) blue stick lah, ni chau si ren (you’re being too noisy)...
TJ: he says he don’t have blue...tag
JG: I don’t have blue stick.

Globalism_Use of Western forms of address
This includes ‘Mr’ or ‘Madam’ and other Western forms of address. As with ‘Localism_Use of Asian forms of address’, this category does not seem to apply to this data set as our focus is on peer-peer talk and the peers do not address each other formally. No examples found to date.

Singlish (Silver, 2008)
Discussion of language-in-education in Singapore makes frequent reference to the localized variety, sometimes referred to as ‘Singlish’. There are multiple theoretical interpretations of Singlish vis-à-vis other English varieties, and notions of ‘standard’ and ‘non-standard’. Alsagoff (2010a) has a summary of some of the most well-known models for analysing Singlish including critique of those models. For our purposes, given the educational context and local discussions of ‘Singlish’, we identify four linguistic features that are commonly considered to be evidence of Singlish. These features have been described the PWPT Anotation Manual: Singlish (Silver, 2008) which is drawn from multiple sources. Because of the overlaps with the global/local features posited by Alsagoff (2010a), only four features are coded for Singlish (as distinct from ‘localization’).

Count_Non-count noun
For the purposes of this analysis, coding is done only for non-count nouns used as count nouns as in. This is coded at the level of word. No examples found in our data.

Null copula
The copular verb ‘be’ is unstated (Example 39). Coding is at the level of sentence.

Example 39

ZY: (Chuckles) Eh! Ssssoccer! Where you?! (P3.D1.C1.ZY_WH_JG.final)

Null subject
A grammatical subject is not explicitly stated (Example 40). Where there is ellipsis, the utterance will not be coded for ‘null subject’. Coding is at the level of sentence.
Example 40

ZY: Want to Go to Merlion.

Null verb auxiliary

An auxiliary verb such as ‘do’ (Do you want to go?), ‘have’ (Have you got the paper?) or ‘be’ (I am eating or I was given a book) is not used (Example 41a). Use of the perfective ‘have’ is quite tricky since simple past can often be used. Note 41b in which the auxiliary is not stated: ‘have finished’ would be one option for expressing verb tense; another alternative would be ‘finished’. Since the latter is possible, 41b cannot be coded for ‘null verb auxiliary’. Utterances with ellipsis are also not coded for ‘null verb auxiliary’. Coding is at the level of sentence. (Note also that our analysis does not take into account morphology for past tense.)

Example 41a

WH: No, wait wait, wait
TW: I ask you, I tell you, I, I suggest this: After [at lunch time] we go to Hotel for our lunch right!
WH: [What you suggest?] Lunch ah? (Animatedly) Very gooood!
TW: (Chuckles)

Example 41b

TJ: We aready finish.

Null Plural marking

The morpheme ‘-s’ to mark plural nouns is not used. Coding is at the level of sentence.

Example 42

TW: Then there is two drink ... then there is a cup of drink on the floor

Null Object

No explicit object is stated for active sentence structures.
Example 43

TW: Jane is using a computer

XL: Number 2 hah?
TW: Huh?

XL: You must tell the number

In the above example, XL’s utterance is missing the object ‘me’ (You must tell me the number)

Acknowledgements

Funding provided by the Centre for Research in Pedagogy and Practice (CRP 48/03 RS) and Graduate Programmes & Leadership Office (RS 7/11 RS; RS 8/13 RES), National Institute of Education, Singapore. Our thanks also to FU Ying Wei (Anna), Galyna KOGUT, TAN Lee Choo, WONG Angelia and WONG Hui Xuan for assistance with the data, and to Wendy BOKHORST-HENG for discussion of the coding. Final interpretation of the codes are those of the authors.

References


Appendix A

This appendix provides information on codes which the authors believe could be useful, but which seemed to be too subjective and idiosyncratic in the actual coding.

Relation (Ncoko, Osman & Cockroff, 2000)

Ncoko, et al (2000) studied a group of multilingual students in South Africa to understand their use of and reasons for codeswitching in formal and informal contexts, i.e. the classroom and playground respectively. As part of that analysis, they concluded that different aspect of ‘relation’ among the students were important. The basic definitions for aspects of relation were drawn from their work.

Establish rapport/Disrupt rapport

“Codeswitching can occur when a speaker tries to change the tone of the conversation. In this case codeswitching has a phatic function, which can be positive or negative. It is positive (establishing rapport) if it servers to narrow social distance, or if it is indicative of a relationship of solidarity. It is negative (disrupt rapport) if it serves to increase social distance” (Ncoko, Osman & Cockroff, 2000, p. 232).

We were able to identify a few potential examples:

Establish rapport

TJ: 你覺得好玩嗎? (Do you think it’s fun?)

( P3.D1.C2.all.final)

Disrupt rapport

JG: Eh 不要打我的頭. (Eh don’t hit my head.)


However, these examples seems to be highly subjective. In addition, there are instances an utterance could be interpreted as ‘establish’ or ‘disrupt’ depending on the coder’s perspective. For example,

JG: Everything hao le (ready), gei ni qi si le. (You are driving me mad.)

( P3.D2.Cycle1.JG+TJ.final)
In this case, JG’s comment can be considered as impolite and increasing social distance as it chastises the other recipient (Brown & Levinson, 1987). On the other hand, bald on record comments of this sort can be indicators of close social relationships (Wolfson, 1988).

It seems that it is insufficient to rely exclusively on the speech content in order to make a decision on whether codeswitching is used to establish/disrupt rapport. It would seem that other features such as non-verbal communication and speech tone play important roles in making that decision. (There is no indication in the study by Ncoko et.al. whether they had considered tone of the speakers in their analysis.) A more thorough analysis might make it possible to code for establish/disrupt rapport in ways that are theoretically valid and consistent; however, as of this writing we found the coding to be subjective and inconsistent.

**Include/exclude participation**

**Include**
The speaker switches to the language that a monolingual who joins the conversation knows and understands; the addressed may even be invited to participate in the conversation. (Ncoko, Osman & Cockroff, 2000)

**Exclude**
This switching is often conscious and contains negative comments about those excluded. It indicates that the speakers share the language and the social distance between them is narrowed at the expense of the others who do not speak the same language. (Ncoko, Osman & Cockroff, 2000)

Include/exclude participation does not seem to be appropriate for our data set because of the different contexts of the present study and the original. The students in the original study spoke many different African languages and English and therefore, could use language to include or exclude others in their conversations. All the students in our study are bilingual in Chinese and English.

**Cooperation**
This is a new category based on the idea of cooperation and collaboration promoted through peer work, especially for theories of cooperative learning (Cohen, 1994; Kagan, 1990; Jacobs, 1994; Slavin, 1990). Because the tasks were set up for peer-to-peer interaction involving cooperation, we also considered whether it was possible to operationalize instances of establishing/disrupting cooperation in relation to codeswitching.
Establish cooperation

This is when codeswitching is used to allow for the smooth completion of the task at hand.

TJ: Ok, never mind.  

Bu kuo zhe ge zhao bu dao jiu gen wo shuo ok hah? (If you can’t find this, you let me know) (P3.D2.Cycle2.JG+TJ.final)

Interfere cooperation

This is when codeswitching is used to disrupt the collaboration between the group members whilst doing the task.

JG:  

Eh his turn he don’t want.  

到你的 turn liao leh. (Eh his turn he don’t want. Your turn liao leh.)

As with establish/disrupt rapport, we found that the actual coding was highly subjective and inconsistent. Therefore, these codes were not used in the analysis.