Text as I Say not as I Do?

Comparing Code-Switching Behaviour in Swiss-German SMS Data with its Metadata

Claudia Bucher
Cédric Krummes
★ Beat Siebenhaar (thank you!)
“Although people who live in bilingual communities are, generally, aware of the existence of CS and of the fact that they themselves sometimes switch, their level of awareness of their own CS behaviour seems to lag far behind their practice.” (Gardner-Chloros 2009: 15)
Overview

• Language Situation in Switzerland
• Corpus Data
• Corpus Metadata
• Corpus Annotation & Code-Switching
• Results
• Discussion

SMS Users
1. What they say.
2. What they do.
3. Linking what they say with what they do.
Language Situation in Switzerland

Official in CH
63.7 % DEU
20.4 % FRA
6.5 % ITA

+ National in CH & Official in CH-GR
0.5 % ROH

9.0% other

Lüdi & Werlen
(2005: 12)
Diglossia in German-Speaking Switzerland
Swiss-German (GSW) & High German (DEU)

• Classical diglossic situation (Weinreich 1953)
• Medial diglossia (Kolde 1981) with few exceptions in oral communication
• Altered medial diglossia with a new dialectal literacy in CMC
  – mainly informal use of written dialect
  – no orthographic standard
  – regional and even local dialects can be distinguished in written dialects (Siebenhaar 2003, 2005)
Corpus Data
www.sms4science.ch

- data collected: November 2009 – July 2011
- 25,947 SMS
- ~517,557 tokens
- 2,784 users
- 80% of SMS with metadata
- 47% of users with metadata
Our GSW-Subcorpus

- 1,771 SMS
- 2,473 code-switches ("CS")
- 53,907 tokens
- 521 users
### Metadata

- age
- gender
- mother tongues
- other languages spoken
- education
- occupation
- postcode
- SMS usage
- CMC usage
- reading & writing habits
- switching and mixing languages
20. I chum au nüme drus. Aber eis isch sicher: mir hend hüf vom 2-
viertelvorf 4 in vorlesig im kol-f-NNN, ob matley oder dä ander. Morn
hemen ka ta, stoht: "no session"

21. Ne, don't panic, i hans ganz sicher nöd als pflichtmodul
ufgschribe. Aber hani die info nöd vo dir, bzw vom baptiste? KuBi und
aa, find da vom tweeber nöd :-( :$

24. Oh schön, do wäri doch gab au dri gsprungä. I bis au chli am
chille. A dä Bar mit dä Mom, ihräre Cousine, dären Ma und Tochter.
Han gend en Sangria hinder mir ;-) 

25. Salüü madame :) schöön, dases dur guet gaht .. ich bin à paris
und gnüeses sehr ( und jaa, es gaht nüme lang "juhu") bisoous

32. Je viens de rater mon train ce qui veut dire :) dàu du chli uf
mich wartä muasch ... Bin wahrschindli so uf di 10 ab zwölf det .... Bis
dänn et dsl pour le retard ...

52. Hoi Lari! Chunsch morn ad vorbesprechig vom seminar übungen
am BGer? Lg Delphine

61. Danke NNNN mal für dini SMS! Chunt sho guet mite MA-Arbeit!
Jüre hed Fun mite Cam, abr nöd mitm B. Mr sind lang shwijing

62. Liebs Löl, shöne Sonntig! M. really sucks! Abr jo, Jüre
amusiert sich glich, ez chin, ässe, mjam =) i zeig dr den Föteli. fri mi
übri dini liebe SMS! Düsbs fest! morn Musical

63. Publikation, Pris: go 4 it! Gib ales, i drukdr fest de Dume! Jüre
Booffahrspaß uf Themse & StarbuckCookie & chli Sonza & CamFun =)
GutN6 /Gut/ /N8/ my dearest Super-Leli, i dänk fest a di!

64. Mer sind ins da Cambridge! Ish mis Mail anchet. Lavern, a beach =)
wie gohts mite MA-Arbeit? ez suechedmir de Fluß zum chille
Code-Switching ("CS")

- alternations ("ALT")
  - between utterances in a turn or between turns
  - change of matrix language

- insertions ("INS")
  - embedding of a word or a phrase into the morphosyntactic structure of a matrix language
  - no change of the matrix language

(Muysken 2000, Poplack 2004)
Examples of CS

ALT
Ja hani, na rechtviel. Danke also, I try to write you a Postcard honey! Und du pass auf dich auf und bis i 3 Wuche. (SMS ID 3933)
(Yes I have, pretty much. So thank you, I try to write you a postcard honey! And you take care and see you in 3 weeks.)

INS
He wemmer öpis choche oder hesch scho gesse?
Mini parents sind au weg (SMS ID 1393)
(Hey, do we want to cook something or have you already eaten? My parents are also away)
ANNIS Online Corpus Platform for Multi-Level-Annotated Data Analysis with Export Functions
Research Questions

1. What do they do?
   observed CS

2. What do they say?
   reported CS

3. Do they do what they say?
   comparing observed with reported CS
What they do
For the total number of CS, there is not significant effect for:

- education
- occupation
- age group
- gender (unless you ask the one-tailed question “Do women switch more than men?”)
What they do
The number of High German ALT significantly rises with age.

High German ALT by Age Group
What they do

The number of English ALT significantly decreases with age. There is a significant difference between the 0-17 age group and the other age groups.

![English ALT by Age Group](image-url)
What they say
Almost 2 users out of 7 in our subcorpus claim they do not code-switch. Even though they do indeed code-switch!

<table>
<thead>
<tr>
<th>Do you CS?</th>
<th>N</th>
<th>%</th>
<th># of CS</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>366</td>
<td>70</td>
<td>1937</td>
<td>78</td>
</tr>
<tr>
<td>no</td>
<td>149</td>
<td>29</td>
<td>522</td>
<td>21</td>
</tr>
<tr>
<td>no answer</td>
<td>6</td>
<td>1</td>
<td>20</td>
<td>1</td>
</tr>
</tbody>
</table>
Linking what they say with they do

Users who claim code-switching use more CS than those saying they don’t code-switch.

Number of CS by Claim to Switching in SMS
(one outliner with 100 CS not in the figure)
Linking what they say with they do

Users who claim code-switching use more CS in their SMS (containing CS) than those saying they don’t code-switch.

Number of CS per SMS (with CS) by Claim to Switching in SMS
Discussion 1/3

- CS in general is not affected by socio-demographic factors
- Alternations into High German and into English vary by age group
- 2 SMS users out of 7 are not aware of their own CS
- awareness of CS correlates with producing CS
Discussion 2/3

What are the reasons of non-awareness?

• Negative attitudes towards mixing languages?
• Integration of words
• CS-Mode in its own right?
  - Users perceive no CS when languages flow into another
  - Researchers perceive/annotate distinct CS/Codes
Discussion 3/3

(Non-) awareness in written CMC

• Asynchronous CMC allows more planning

• Does CS always correlate with more planning, drafting and rewording?

• How should Corpus Linguistics deal with language awareness and text use in general?

(Androutsopoulos 2011:17)
01 June 2014: our WhatsApp project launch

What's up, Switzerland?

www.whatsup-switzerland.ch