

The user, the switch, and the SMS: Reported and observed code-switching behaviour in CMC

CÉDRIC KRUMMES
CLAUDIA BUCHER
BEAT SIEBENHAAR
Universität Leipzig

One part of the future of dialects lies in its use in non-formal written communication that emerged with the development of CMC (computer-mediated communication, e.g. SMS, email, chats). Dialectal, standard, and foreign language variants are used side by side, resulting in numerous forms of code-switching in written communication, which promote and reflect multilingual practices and identities.

Although code-switching in written computer-mediated communication has been investigated since the 1990s (see Sebba et al. 2012), scholarship has only focused on authentic code-switching data, but failed to adequately complement this perspective by focusing on the users' view on their own code-switching as well.

The paper uses data from the Swiss SMS corpus sms4science.ch, which was collected between November 2009 and July 2011. It comprises almost 26,000 SMS, of which the main languages are Swiss German Dialects (41.2%), Standard German (28%), French (17.8%), Italian (5.7%), and Romansh (4.3%). Besides sociodemographic data metadata about the SMS users include answers to two open questions about the users' manner of writing SMS and switching languages in SMS, which are in the focus of this paper.

This paper compares the behaviour of code-switching patterns found in a sub-corpus of Swiss German SMS (1,927 text messages) with the metalinguistic reflections of its users.

Clustering analysis agglomerate users in different SMS-writing and code-switching types. Other multivariate analyses take into account formal aspects of code-switching which structure the messages. Comparing the metadata clusters and the annotation clusters shows correlations between SMS users' reported code-switching behaviour and their observed behaviour.

Overall, this paper argues that while SMS users may report one type of code-switching behaviour when completing a metadata questionnaire, their data suggests different patterns of behaviour. This ultimately feeds into the adequacy of querying code-switching behaviour in metadata questionnaires and the general psycholinguistic motivations behind code-switching.