Language, Mobile Phones and Internet: A Study of SMS Texting, Email, IM and SNS Chats in Computer Mediated Communication (CMC) in Kenya

Published by
LOT
Janskerkhof 13
3512 BL Utrecht
The Netherlands

phone: +31 30 253 6006
fax: +31 30 253 6406
e-mail: lot@uu.nl
http://www.lotschool.nl

Cover illustration: Boy with mud-phone. Courtesy of the Entrepreneurial Programming and Research on Mobiles (EPROM), Kenya. http://www.media.mit.edu/ventures/EPROM/

ISBN: 978-94-6093-044-7
NUR 616

Copyright © 2010: Sandra Nekesa Barasa. All rights reserved.

Language, Mobile Phones and Internet: A Study of SMS Texting, Email, IM and SNS Chats in Computer Mediated Communication (CMC) in Kenya

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. P.F. van der Heijden,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 11 November 2010
klokke 15.00 uur

door

Sandra Nekesa Barasa

geboren te Bungoma, Kenia
in 1975

Promotiecommissie:

Promotores:	Prof.dr. M.P.G.M. Mous Prof.dr. M. van Oostendorp
Overige leden:	Prof.dr. H. Bennis (Meertens Instituut, Amsterdam) Prof.dr. A. Bodomo (University of Hongkong) Prof.dr. M.E. de Bruijn

Acknowledgements

The realization of this endeavor has been a major challenge encountered and this offers me a great opportunity to display my deep and sincere gratitude to all those who contributed to make it a success.

I am indebted to Prof. Vincent van Heuven who set off my journey after accepting my research proposal for admission into the PhD programme at Leiden University Centre for Linguistics (LUCL). In the same breath, my deepest gratitude goes to my two promotors Prof. Maarten Mous and Prof. Marc van Oostendorp for their persistent help and invaluable guidance in shaping this book to a full realization. The many meetings we held both in bad and good weather days proved worthwhile. Prof. Marc, thanks for supplying the Dutch summary. I would also like to thank Prof. Hans Bennis for his commitment and guidance in embettering this dissertation.

Next, my gratitude goes to my LUCL colleagues for a congenial, cheerful and supportive work environment; Heleen Smits, Mercy Lamptey, Ongaye Oda, Victoria Nyst, Felix Ameka, Connie Kutsch Lojenga, Anne-Christie Hellenthal, Jenneke van der Wal Maggie Konter, Thilo Schadeberg, Crit Cremers, Hilke Reckman, Kristina Riedel, Christian Rapold, Jean Chavula, Kofi Dorvlo, Azeb Amha, Tolemariam Fufa, Ramada Elghamis, Stanly Oomen, Khalid Mourigh, Martine Bruil, Margarita Gulian, Marijn van 't Veer, Kathrin Linke, Michaël Peyrot, Kalinka Timmer, Rinus Verdonschot, Pepijn Hendriks, Frank Landesbergen and Vivian de la Cruz. I also thank the LUCL 'gals' Rebecca Voll, Kateřina Součková, Camelia Constantinescu, Juliette Huber and Allison Kirk for the nice get-togethers. From the phonetics lab, I am grateful to Jos Pacilly, Franziska Scholz, Willemijn Heeren, Jurriaan Witteman and Yiya Cheng.

I am sincerely grateful to my two paranimphs, Maarten Hijzelendoorn, and Liza van der Aa. Maarten, without your *eDatax.mdb* program, I would still be doing research counts. You reduced my data analysis work by half, thanks. Also to your family Els, Merel & Jeroen for the lovely dinners they shared with me and for allowing me to win the Sjoelen game on a new year's eve. And Liza who was the first Dutch friend I made and she literally became my younger sister after she introduced me to her whole clan. I will always be grateful to the generosity of Marieke and the Van der Aa family: Wietze, Bella, oma, Liza, Jessica, Saskia, Hielke and of course Sammy and Simba who adopted me into their warm and cozy family and made sure that I never lacked nor spent any Christmas holiday by myself. I enjoyed all your extended family gatherings where you taught me to pronounce the Dutch tongue twister *gefeliciteerd*.

This research would not have matured without the unending support from Ngundo, Rhoda, Meja and Laura who coordinated my data collection and remuneration. *Thanks kwa kunichujia data, mlinisort vipoa mpaka I'm bila words*! Ngundo, thanks for the Sheng summary. Loads of gratitude go to Dr. Gertrud Schneider-Blum for her tireless commitment in editing the manuscript ;-)

The fun gals, Els, Mo, Kat, Fem, Anna, Esmeralda, thanks for all the fun we had together;-) Mark, Maurizio and Arthur, you always made my day with your never ending curiosity about the Kenyan culture. Benneth, you turned out to be a true brother through thick and thin.

Many thanks are conveyed to the families that provided a home away from home, Anita & Frank, Yola & Bert, Sander & Suzanna, Jimmy & Godlif, Jan & Alice, Cecelia & Peter. I'm grateful to my Kiswahili group made up of Myrte, Annemieke, Frank, Sander and Christian for the warm evening meetings. You kept my Kiswahili spirit alive.

It is an honour for me to thank my CLIC colleagues at Eindhoven University (TU/e), Vincent, Fran, Astri, Lettie, Steven, Desiree and of course my Dutch teachers Elly & Leonie- Dank u!

I will not forget the support from my Kenyan friends who kept me in their prayers, Lau & Davy Mamuli Barasa, Mama Teddy, Joan & Mwambura, Joyce & Mutisya, Milka Aswa, Bob Mbori, Richard Onchiri, Elizabeth Abenga, Grace Koteng, and Jacquiline Ondimu,

My heartfelt thanks go to my parents Barasa Wangila & Violet Nafula who instilled in me from an early age the desire to obtain a PhD and for taking over my responsibilities while I was away. I would like to also thank my siblings, Fe & Rose, Nelima & Myk, Fid & Judy, Janet and Liz. Of course, last but not least it is my pleasure to thank Leo, Louisa, and Ngundo for the unending patience and support. Leo and Loui, this book is dedicated to you.

I wish to thank the respondents of my study, who remain anonymous for confidentiality purposes, and for all others that I have not mentioned, you know who you are, and I am deeply grateful for your contribution in making this dissertation a reality.

Finally, I thank the almighty Father and our lady Mary for all the blessings bestowed upon me.

TABLE OF CONTENTS

LIST OF FIGURES ..... **VI**

LIST OF GRAPHS ..... **VII**

LIST OF TABLES ..... **VIII**

LIST OF CHARTS ..... **X**

LIST OF ABBREVIATIONS ..... **XI**

CHAPTER 1. INTRODUCTION **1**

1.1. COMPUTER MEDIATED COMMUNICATION (CMC) 1
1.1.1. Challenges of CMC *4*
1.2. CMC GENRES UNDER STUDY 5
1.2.1. Short Message Service (SMS) *5*
1.2.2. Electronic Mail (Email) *11*
1.2.3. Instant Messages (IM) *12*
1.2.4. Social Network Sites (SNS) *15*
1.3. COMPUTER MEDIATED COMMUNICATION AND LANGUAGE 26
1.3.1. CMC and Discourse *28*
1.3.2. CMC and Sociolinguistics *29*
1.3.3. CMC Register *30*
1.3.4. Language Change and Variation in CMC *32*
1.3.5. CMC and Multilingualism *34*
1.3.6. Other Languages in CMC *37*
1.4. THE CURRENT RESEARCH 39
1.4.1. Research Objectives *39*
1.4.2. Rationale of the Research *40*
1.5. HYPOTHESES 42
1.5.1 Principle of Rapid Communication *42*
1.5.2 Principle of Least Effort *43*
1.5.3 Principle of Mode Limitation *44*
1.5.4 Principle of Informal Communication *44*
Codeswitching 45
Peer Communication and Identity 45

CHAPTER 2. LANGUAGE AND COMMUNICATION IN KENYA **47**

2.1. LANGUAGE IN KENYA 47
2.1.1. Indigenous Vernacular Languages *48*
2.1.2. Kiswahili *50*
2.1.3. English *52*
2.1.4. Sheng and Engsh *54*
2.1.5. Conclusion *59*
2.2. COMPUTER MEDIATED COMMUNICATION IN KENYA 62

2.2.1. Mobile Phones and Networks ... *63*
2.2.2. Safaricom ... *66*
2.2.3. Zain Kenya ... *67*
2.2.4. Orange Mobile ... *68*
2.2.5. YU ... *68*
2.2.6. Internet ... *69*

CHAPTER 3. METHODOLOGY ... 75

3.1. Population Group ... 75
3.2. Scope and Limitations ... 76
3.3. Ethical Considerations ... 77
3.4. Data Collection ... 78
3.4.1. Emails ... *78*
3.4.2. SMS ... *79*
3.4.3. IM ... *79*
3.4.4. Social Network Sites (SNS) ... *79*
3.5. Variables ... 80
3.5.1. Phonological Spelling ... *81*
3.5.2. Pronounceable Letters ... *82*
3.5.3. Pronounceable Numericals ... *83*
3.5.4. Abbreviations ... *84*
3.5.5. Acronyms ... *85*
3.5.6. Exclusive Consonants ... *86*
3.5.7. Contractions ... *86*
3.5.8. Misspelling and Typographic Errors ... *87*
3.5.9. Capitalisation ... *88*
3.5.10. Punctuation ... *89*
3.5.11. Graphics: Smileys and Emoticons ... *91*
3.5.12. Other Symbols ... *93*
3.5.13. Salutations ... *94*
3.5.14. Language Choice ... *95*
3.5.15. Interword Codeswitching ... *95*
3.5.16. Intraword Codeswitching ... *96*
3.6. The eDatax.mdb Program ... 97
3.7. Procedure of Data Analysis ... 98

CHAPTER 4. TECHNOLOGICALLY MOTIVATED FEATURES OF CMC ... 103

4.1. Phonological Spelling ... 103
4.1.1. Determiners and Demonstrative Pronouns ... *110*
4.1.2. Use of /z/ and influence of Engsh in CMC ... *110*
4.1.3. Conclusion ... *117*
4.2. Pronounceable Letters ... 118
4.2.1. Pronounceable Vowel Letters ... *122*
4.2.2. Pronounceable Consonant Letters ... *123*
4.2.3. Conclusion ... *129*
4.3. Pronounceable Numericals ... 130

4.3.1. Bilingual use of Numericals *134*
4.3.2. The L33tspeak Phenomenon *144*
4.3.3. Conclusion *146*
4.4. PRONOUNCEABLE SYMBOLS 148
4.5. LEXICAL COMPRESSION 149
4.5.1. Abbreviations *151*
4.5.2. Exclusive Consonant Spelling *159*
4.5.3. Acronyms *161*
4.5.4. Contractions *164*
4.5.5. Other Forms of Lexical Compression *167*
4.5.6. Conclusion *167*
4.6. RELAXING SPELLING STANDARDS IN CMC TEXTS 168
4.6.1. Missing Characters *174*
4.6.2. Exclusion of Spaces *175*
4.6.3. Letter Repetition *176*
4.6.4. Capitalisation *178*
4.6.5. Variants of English *187*
4.6.6. Conclusion *188*
4.7. PUNCTUATION 190
4.7.1. Hyphens (-) *193*
4.7.2. Exclamation Marks (!) *196*
4.7.3. Question Marks (?) *198*
4.7.4. Ellipsis (...) *200*
4.7.5. Comma (,) *204*
4.7.6. Quotation Marks (") *206*
4.7.7. Colon (:) and Semicolon (;) *207*
4.7.8. Apostrophe (') *207*
4.7.9. Other Symbols *212*
4.7.10. Termination Marks *213*
4.7.11. Conclusion *217*
4.8. GRAPHICS 218
4.8.1. Smileys *222*
4.8.2. Emoticons *225*
4.8.3. Explained Motions *226*
4.8.4. Conclusion *226*

CHAPTER 5. SOCIALLY MOTIVATED FEATURES OF CMC 229

5.1. SALUTATIONS IN KENYAN CMC 230
5.1.1. Greetings *232*
5.1.2. Valedictions *240*
5.1.3. Language Choice for Openings and Closings *246*
5.1.4. Conclusion *250*
5.2. LANGUAGE CHOICE IN KENYAN CMC 251
5.3. CODESWITCHING IN CMC 258
5.4. INTERWORD CODESWITCHING 263
5.5. INTRAWORD CODESWITCHING 267

5.5.1. Intraword codeswitching in Verbs *269*
5.5.2. Kiswahili Inflectional Prefixation *270*
5.5.3. Kiswahili Derivational Suffixes *272*
5.5.4. English Affixation *276*
5.5.6. Affixation in Sheng *276*
5.5.6. Affixation in Engsh *278*
5.5.7. Grammar in Engsh *278*
5.5.8. Noun plurals as markers of Engsh *282*
5.6. Conclusion 284
5.7. Motivation of CS in CMC 285

CHAPTER 6. CODESWITCHING (CS) AND CODES 289

6.1. Properties of Kiswahili – English CS in CMC 289
6.2. Properties of English-Kiswahili CS in CMC 291
6.3. Challenges to the Matrix Language Framework (MLF) Hypothesis 295
6.4. New Motivations of CS 301
6.5. CS as an Unmarked Choice 303
6.6. Codes and CS 304
6.6.1. Sheng *304*
6.6.2. Engsh *308*
6.7. CS, CMC and Orality 310
6.8. Conclusion 310

CHAPTER 7: GRAMMAR IN CMC 313

7.1. Function Words and Grammatical Parts 314
7.2. Articles 314
7.3. Pronouns 315
7.4. The Auxiliary Verb *be* 317
7.5. Conjunctions 318
7.6. Conclusion 319

CHAPTER 8. GENERAL CONCLUSION 321

8.1. Conclusions on CMC Genres 326
8.1.1. SMS *326*
8.1.2. Email *326*
8.1.3. IM *327*
8.1.4. SNS *327*
8.2. Conclusions from Hypotheses 328
8.2.1. Principle of Rapid Communication *328*
8.2.2. Principle of Least Effort *329*
8.2.3. Principle of Mode Limitation *330*
8.2.4. Principle of Informal Communication *331*
(i) Codeswitching *331*
(ii) Peer Communication and Identity *333*
8.3. Research Contribution 333

8.4. RECOMMENDATIONS FOR FURTHER RESEARCH 335
REFERENCES 337
SUMMARY IN ENGLISH 355
SAMENVATTING IN HET NEDERLANDS 358
SAMA YA SHENG 361
CURRICULUM VITAE 365

List of Figures

Figure 1: An SMS Transmission Flow.....7
Figure 2: Traditional mobile phone Keypad.....10
Figure 3: iPhone touch screen keypad.....10
Figure 4: Map showing the 8 provinces in Kenya.....51
Figure 5: The Main Window of the eDatax.mdb Database Input Program.....99

List of Graphs

Graph 1: Mobile Subscription in Africa (2002-2012) 65
Graph 2: % Growth of Mobile Phone Subscription in 2008 in Africa 66
Graph 3: Growth of Mobile Phone Subscription in Kenya 67
Graph 4: Percentage of Internet Penetration in Africa 70
Graph 5: Percentage of Internet Users in Africa 71
Graph 6: List of the Top 10 African Countries in the Use of Internet 72
Graph 7: Percentage of Internet Growth in Kenya 73

List of Tables

Table 1: Sheng Lexemes with their origin 58
Table 2: List of Recurring Emoticons and their Codes 92
Table 3: List of Recurring Smileys, their Emoticon counterparts and Codes 93
Table 4: No. of Total Messages Collected 100
Table 5: No. of Total Messages against those Selected Per Genre 100
Table 6: Selected Messages Per Genre 101
Table 7: Distribution of Phonological spelling 104
Table 8: English Nouns with /z/ Suffix 113
Table 9: The /z/ suffix in Kiswahili and Sheng Words in Engsh 113
Table 10: /o/ + /z/ Plural marker 114
Table 11: Compulsory non plural marker /z/ suffix 115
Table 12: /z/ Suffix on Adverbs 115
Table 13: /z/ suffix on English words 116
Table 14: Distribution of Pronounceable Letters among Genres 119
Table 15: Summary of Pronounceable Letters 130
Table 16: Use of Pronounceable Numericals 131
Table 17: Kiswahili and English Homophonic uses for Numerals 134
Table 18: Distribution of Lexical Compressions 150
Table 19: Distribution of Acronyms 162
Table 20: Distribution of the Apostrophe in Contractions 165
Table 21: Distribution of Misspellings and Typos among Genres 169
Table 22: Number of Excessively Capitalised Characters per Genre 178
Table 23: Excessively Capitalised Messages per Genre 179
Table 24: Exclusively Capitalised Messages per Genre 180
Table 25: Message Capitalisation 181
Table 26: Expressively Capitalised Messages per Genre 182
Table 27 Number of Absent Capitalisations per Genre 186
Table 28: Distribution of Punctuations per Genre 192
Table 29: Messages without Termination marks 214
Table 30: Distribution of Smileys and Emoticons 219
Table 31: Frequency use of different Smileys and Emoticons 222
Table 32: Greetings per Genre 233
Table 33: Valedictions per Genre 240
Table 34: Use of Languages in Kenyan CMC Greeting 248
Table 35: Use of Languages in Kenyan CMC Valediction 249
Table 36: Use of Language in Kenyan CMC data 253

Table 37: Distribution of Language use per word across the genres 253
Table 38: No. of Messages with Codeswitching per Genre 263
Table 39: Multilingual Interword Codeswitching 264
Table 40: Trilingual Interword Codeswitching 265
Table 41: Bilingual Interword Codeswitching 265
Table 42: Intraword codeswitching language Combinations 268
Table 43: No. of messages with language Combinations 268
Table 44: Distribution of affixes in Intraword CS 269
Table 45: Engsh Verbs with Engsh Grammar 282
Table 46: Dummy Suffix in Sheng 307
Table 47: Sheng grammatical suffix *-ang-* 308
Table 48: Engsh Vocabulary 309

List of Charts

Chart 1: % Use of Phonological Spelling 105
Chart 2: % per Genre of Pronounceable Letters 119
Chart 3: % Use of each Letter 121
Chart 4: % of Pronounceable Numericals per Genre 132
Chart 5: % Use of each Number 133
Chart 6: % of Lexical Compressions per Genre 151
Chart 7: % Distribution of Spelling Errors in CMC 170
Chart 8: % of Expressively Capitalised Messages 183
Chart 9: % of Absent Capitalisations per genre 186
Chart 10: % of Punctuations per Genre 191
Chart 11: % Use of Punctuation Marks 193
Chart 12: % of Hyphens per Genre 194
Chart 13: % of Exclamation marks per Genre 196
Chart 14: % of Question marks per Genre 198
Chart 15: % of Ellipsis per Genre 201
Chart 16: % of Commas per Genre 205
Chart 17: % of Messages without Termination Marks per Genre 215
Chart 18: % of Full stops per Genre 216
Chart 19: % Use of Smileys and Emoticons 220
Chart 20: % Distribution of Greetings in Messages per CMC genre 234
Chart 21: % Distribution of Valedictions in Messages per CMC genre 241

List of Abbreviations

Adj	Adjective
ASCII	American Standard Code for Information Interchange
ARCC	Regional Centre for Computing
CCK	Communication Commission of Kenya
CMC	Computer Mediated Communication
CMD	Computer Mediated Discourse
CS	Codeswitching
CV	Consonant Vowel
EL	Embedded Language
Email	Electronic mail
Eng	English
FtF	Face to Face
Fut	Future
ICQ	I seek you
ICT	International Communication Technology
Infin	Infinitive
IM	Instant message
Infl	Inflection
ISP	Internet Service Provider
ITU	International Telecommunications Union
KPTC	Kenya Posts & Telecommunications Corporation
Lit	Literal
ML	Matrix Language
MLF	Matrix Language Framework
MSN	The Microsoft Network
N	Noun
Neg	Negation
PDA	Personal Digital Assistants
Pl	Plural
Pt	Past Tense
RO	Rights & Obligations
Sh	Sheng
Sg	Singular
SMS	Short Message Serve
SNS	Social Network Service
Sbj	Subject

Sw	Kiswahili
T9	Text on 9 keys / predictive text input
vrb	Verb
vrn	Vernacular
TraSA	Transcription Statistics tool with automation
UCS	Unified Communications Systems
VNS	Video Network Site
VoIP	Voice over Internet Protocol
VSS	Video Sharing Site
Yuppies	Young Urban Professionals

Chapter 1. Introduction

The focus of this book is on the use of language in Computer Mediated Communication (CMC) in Kenya. The examined CMC genres are Short Messaging Service (SMS), Email, Instant Messages (IM) and Social Network Sites (SNS). In this introductory section of the book, I present a general overview of CMC, its challenges and a presentation of the genres under study. Subsequently, I also give an overview of CMC in relation to language and finally, present the objectives, rationale and hypotheses for the research of which the data and results are discussed at length in the ensuing chapters of the book.

1.1. Computer Mediated Communication (CMC)

Communication is part of human life since time immemorial. Scherba de Valenzuela (1992) describes it as

> "Any act by which one person gives to or receives from another person information about that person's needs, desires, perceptions, knowledge, or affective states. Communication may be intentional or unintentional, may involve conventional or unconventional signals, may take linguistic or nonlinguistic forms, and may occur through spoken or other modes." (Scherba de Valenzuela, 1992:2).

Given the fact that communication is one of the basic necessities to human life, it has been considerably improved and enhanced for ease and expedience in every era right from the earliest known communication. Apart from face to face (FtF) communication, other forms of communication can only be made successful by an intermediary. In fact, Whittaker (2002) captures this very well in his statement that "the natural human communication apparatus is constrained in several ways". There are limits to the distance at which speech is audible, and visible behaviours such as gesture, gaze or facial expressions are perceptible. Furthermore, these natural communication behaviours are transient and do not persist over time. These limitations lead us to rely on some form of mediation if we are to communicate at a distance and across time. People have therefore invented media technologies that attempt

to circumvent these limits to allow remote forms of communication. This is what is meant by Mediated Communication. It is any kind of communication that uses some form of intermediary for it to be accomplished. The mobile phone and the Internet are such mediation technologies that this research focuses on.

The emergence of Internet and cell phone communication in the current age of information has triggered a lot of interest from researchers. In spite of this, most of these studies have focused on the technological aspect and not much has been done using a linguistic approach despite the fact that users keep adapting their languages to fit into the technologies while at the same time manufacturers try to adapt their technologies to fit the users' languages.

CMC is a general acronym for Computer Mediated Communication which refers to the process by which people create, exchange, and perceive information using technologies like networked telecommunications systems that facilitate or mediate encoding, transmitting, and decoding messages. The definition of CMC has undergone a metamorphosis since the term was first coined. Early studies defined CMC as messages exchanged by networked computers. This definition lacked in the aspect of the contribution of the communicators. In the late 1990's studies like December (1997) included the humanity component and defined Computer Mediated Communication as a process of human communication via computers involving people situated in particular contexts engaging in processes to shape media for a variety of purposes. This was maintained over the years until the emergence of the mobile telephone and SMS. CMC researchers then expounded the scope of CMC to include mobile telephony which is operated on digital or analogue networks and thus broadly considered as mediated communication via networks. Arguably in mobile telephony, computer networks are involved at some point in the message transmission process, only that users are not required to interact directly with the computer system via a keyboard or similar computer interface (Lawley 1994). I therefore concur that mobile telephony is indeed computer mediated albeit not as di-

rectly and in such a directly observable way as the others. To take account of all these more specifically, Herring (2007) defines CMC as predominantly text-based human-human interaction mediated by networked computers or mobile telephony. I prefer the term graphic to text in order to also capture the use of Smileys, Emoticons and other graphics in CMC. This encompasses SMS as a text-based format like Instant Messaging (IM), Social Network Service (SNS), and Email thus covering all of them within the remit of CMC. As Eldridge & Grinter (2001:219) aptly sum it up, mobile phones are, in effect, 'mini-terminals' (computers) for text-based communication. This fact then brings me to establish that there are now smartphones which are being manufactured and work just like a mini or pocket computer. They enable one to access and respond to Emails and do instant messaging from anywhere in terms of text-based CMC. The invention and development of these media innovations is very rapid. Höflich & Gebhardt (2005:9-31) explain that media innovations are bringing about a change in the media ecology. A change in the existing cultures of mediation can be seen. New media are added to the previous media repertoire leading to a functional differentiation straight to the point that they overtake functions which earlier media had to take simply for lack of alternatives. Within these processes, the communicative functions of earlier media can change even to the point that they will finally lose their relevancy. A good example of this is the telegram service through the post offices which has been discontinued in many places. Remarkable new forms of "virtual culture" are now developing in this intensely social domain of human interaction (Danet & Herring 2007).

The current most accessible CMC input and output continues to be mainly textual coupled with graphics and pictures and even sound and video clips. More complex communication technologies like oral video conferencing, computer-mediated face to face communication including visual images in real time already exist, but are still expensive and so far not as commonly used although it is envisioned that their usage will spread owing to the speed with which CMC advances.

1.1.1. Challenges of CMC

It is not fair to discuss CMC without exposing its limitations. Just like any other technologies, CMC has its weakness. Bubas (2002) captures these very well and lists the CMC challenges as including:

- Limited social presence
- Anonymity-where the participants are not as clear as in face to face communication
- Reduced/delayed message feedback
- Depersonalised communication

All these in a way make CMC different from face to face communication. It is also worth saying that some of these so called limitations are advantageous to communicators depending on the reasons for communicating. I will discuss these limitations in line with the text CMC that I am interested in.

Short et al. (1976) define social presence as the degree of salience of the other person in a mediated communication and the consequent salience of their interpersonal interactions. It involves the extent to which a medium conveys the actual presence of participants. CMC is limited in this but it is better than the traditional methods of communication. Users of CMC are trying to get around this limitation by inventing Emoticons and Smileys to keep the receiver informed about their feelings. This is a challenge to CMC manufacturers who are designing CMCs like video-conferencing to increase the degree of social presence.

Anonymity is an important feature embedded in CMC. It can be understood as a condition that frees individuals from social evaluation or scrutiny (Pinsonneault & Heppel, 1998). Thus, when the individuals perceive themselves to be anonymous they can contribute without the fear of social repercussions. Visual anonymity enables users to mask their physical or behavioural cues that are undesirable and strategically disclose the ideal self of themselves. They present themselves as perfect according to the communication context. Anonymity also enables users to carefully think and plan on what to say, how to say it and when to say it and

makes receivers to idealise their communication counterpart. Visual anonymity can be advantageous to users who want to make genuine contributions in an anonymous way. But on the other hand, it paves the way for fabrications and deception.

On the feedback issue, CMC (especially text CMC) offers reduced or delayed feedback as compared to FtF communication. This can be nerve-racking for those who need instant feedback. IM is the fastest among the CMC genres in the current study.

CMC can lead to apathetic communication or aloofness and lack of any empathy thus removing any sense of connectedness or intimacy; a situation that makes it easier for people to communicate cruel or inhumane messages.

As already explained, these limitations are advantageous to those who set out to use them for their benefits while at the same time they are limitations and disadvantageous to the receivers.

My conclusion is at par with Riva's (2001) prediction that the technological evolution of the media leads us to believe that CMC could become in the very near future, the predominant *medium*, or rather, it is possible that it will become a general communication interface: an interface used for interpersonal relationship and for the creation and management of information. This prediction is now a reality.

1.2. CMC Genres under Study

In this section, I give a brief introduction of each of the genres that are the focus of this study. The results of this research are from data collected from these genres as used in Kenya. More on the research methodology is presented in chapter 3.

1.2.1. Short Message Service (SMS)

SMS stands for Short Message Service and is also commonly known as text messaging or texting. It began in 1997 when there was a transition from analogue to digital mobile phones. It is a cell phone communication service for sending short text messages to

mobile devices, including cellular phones, smartphones and Personal Digital Assistants (PDA). As early as 1930 Sigmund Freud valued the role of the written word as an alternative to face to face communication. He stated that "*writing was in its origin the voice of an absent person".* Levinson (2004) expounds on this by explaining that a voice can be absent for various reasons; either distance in physical or geographical space, or a time factor whereby the listener arrives before or after the voice has spoken, or even both. The written word or text therefore captures this. This illustrates the importance of SMS which has become a convenient medium that has changed the way we interact and contact each other. No matter where we are, and at any time, we can send these messages to our friends who then have the choice to store, reply, delete or forward them. In message construction, traditional cell phones (*2nd and 3rd generation- 2G and 3G*) restrict the messages to 160 or 224 characters, (approximately 25 words). Newer mobile phones (4G) enable the user to link up to 12 text messages and allow the maximum message length of 1800 characters. Figure 1 further details the transmission of an SMS**.**

SMS messages do not require the mobile phone to be active or within range. They can be held for a number of days until the phone is active and within range. Some networks allow for 'delivery reports' when the message is accessed by the receiver. According to SearchMobileComputing.com[1], SMS can be transmitted in a number of ways, including via:

- One digital phone to another.
- Web-based applications in a Web browser e.g. SMS gateways. These are websites that allow users to send messages to people within the cell served by that gateway. They also serve as international gateways for users with roaming capability.
- Instant messaging clients like *I seek you* (ICQ).
- Voice over Internet Protocol (VoIP) applications like Skype

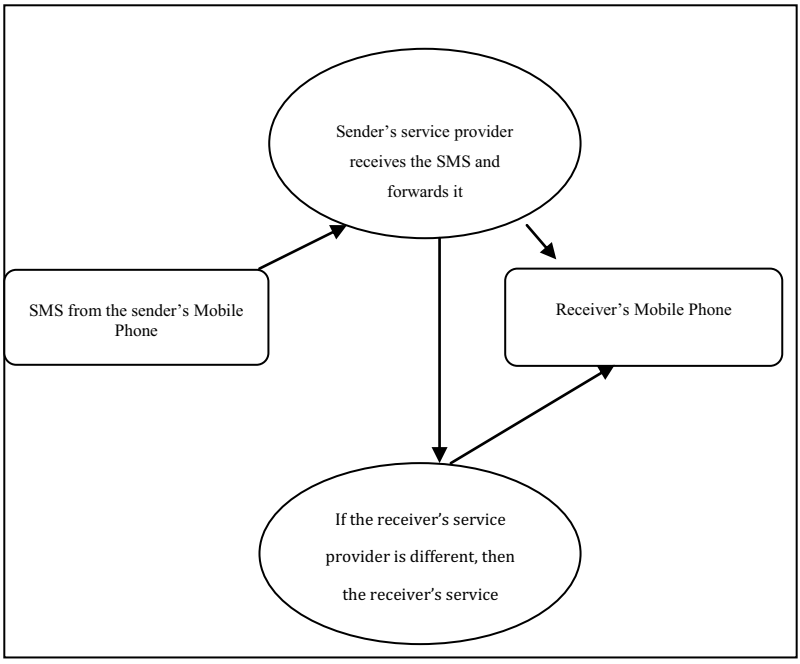
[1] More on this can be found on the searchmobilecomputing.com website at http://searchmobilecomputing.techtarget.com/definition/Short-Message-Service

and Smart VoIP. VoIP technology allows telephone calls to be made over computer networks.

- Unified Communications Systems (UCS). UCS systems enable users to operate all their messages through a single service that can be accessed by several devices.

Currently, over a billion SMS are sent each month worldwide and commercially, the SMS is worth over 130 billion dollars globally per year[2]. In Kenya where the current research is done, use of SMS gained popularity mainly because of its cheaper charges than the actual phone call. This is discussed further on in the following chapter.

Figure 1: An SMS Transmission Flow



The main advantages of using SMS include the following:

- It is cheaper to send/receive compared to the voice call which is expensive.
- It is non-intrusive thus one can write or read an SMS in a

[2] http://communities-dominate.blogs.com/brands/2008/12/trillion-with-a.html

meeting, bus etc and nobody hears you sending the message nor can one decipher what the incoming message is all about.

- It is persistent; it waits until you switch on your phone or until you check your phone.
- It enables direct conveyance of the message without interruption from the recipient. This ensures one way privacy i.e. one has time to compose and sent a message unlike in normal conversations when the recipient interrupts or interferes with your statement.
- It offers a choice whether to reply, forward, or delete. Some phones now have delivery reports such that the sender is notified when the message is read, thus they expect a reply but the good thing is that SMS gives one ample time to figure out the best possible reply.
- It can be saved for future reference unlike the spontaneous spoken word.
- It can be short, casual and precise.

In Kenya some network companies allow free prescribed messages to be sent e.g. Safaricom allows one to send up to five *"please call me*" messages per day to any Safaricom number. This proves handy in emergencies or when one needs to communicate but has no phone airtime credit. The *please call me* initiative on the one hand has helped many but it has been used to different ends too. Many people take advantage of it and send these free messages to each other signalling that they are fine. Some people use the messages to deliberately bother others. This service has been used just like the *beeping* or *flashing* which is done by dialling a person's number and letting the phone ring for a few times and then disconnecting. The receiver will interpret this as a *please call me* message, a greeting message or any other pre-arranged agreement e.g. *please beep me when you arrive home* etc.

Despite these advantages, the SMS has usability issues especially in the message input. The main disadvantages of text messages is that they are cumbersome to type, and one has to think clearly on how to best phrase the message in order to put the point across

with the fewest possible number of words so as not to exceed the character limit. This makes it live up to its name *short message*. In fact in the article on usability issues of sending SMS, Schneider-Hufschmidt (2005) claims that entering the text of a short message on a small device is the hard part of text messaging.

Another general disadvantage of text messaging is that it is only accessible to educated people. This is a reality in the Kenyan community whereby some cannot communicate to their parents in the village via SMS because the parents are uneducated. They therefore have to resort to the much more expensive voice call.

The use of the typical twelve key mobile phone keypad (cf. figure 2) is the common way of entering a text message into the phone. Some mobile phone manufacturers have come up with the touch screen keypad illustrated in figure 3.

This touch screen keypad is considered to be more comfortable than the traditional keypad which requires key presses to type in the characters. The general challenge with these keypads is that some users have far more characters that need to be typed than the phone has keys.

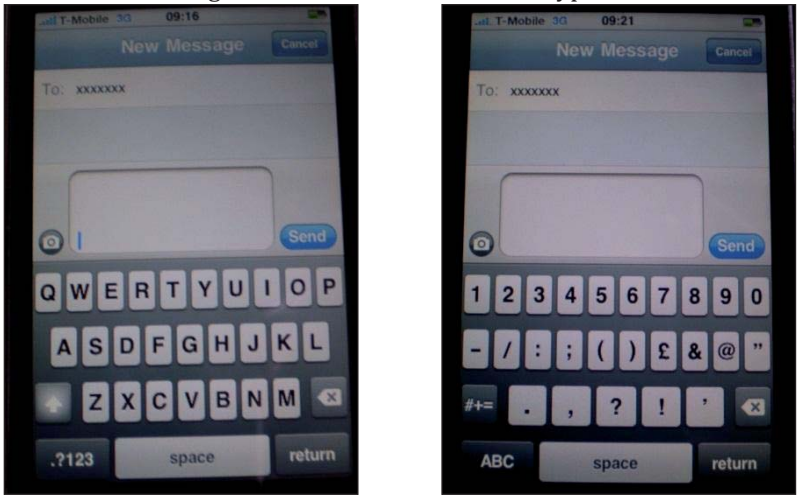
Schneider-Hufschmidt (2005) clearly explains that multiple key presses are common causes for typing errors. He claims that people are not good at counting key presses and are also not good at keeping the time in between key presses such that in an attempt to type a double character, one may press the key in quick succession thus resulting in a different character altogether. Also one can get a double character if one presses the same key a little longer than required.

On the basis of a touch screen one can finally use what is called a virtual keypad, a software program that is able to display a miniaturised keypad on the device.

Figure 2: Traditional mobile phone Keypad

A photo showing a typical mobile phone keypad.

Figure 3: iPhone touch screen keypad



Photos of an iPhone touch screen keypad with a display of alphabetic characters, numerals, symbols & punctuations.

With the use of a pen, or possibly also with the finger, the user types text as on a standard keyboard, one character at a time. In addition to the usually very small size, which makes it hard to touch the right key, there is also the problem of mode switches for the input of special characters or numbers which needs to be understood by the user. Given that the screen of most of today's devices is fairly small, the display of a keypad may obscure most of the text which the user is just then trying to compose.

A number of mobile phone manufacturers have tried to make the process of text input easier for standard situations by providing standard text templates for example; *I will be xx minutes late.* Such templates can easily be modified by users without being forced to type lengthy character sequences. This standard text is usually not modifiable, only the place holders, as the *xx* in the above example, can be replaced by digits or letters. At the end of such a predefined message the user can add personal text. An additional alternative way to reduce the number of keystrokes necessary to type text is the use of a shorthand form *("T42"* being interpreted as *"tea for two"*) to write text. In this research, I describe this phenomenon as a combination of pronounceable letters *(T* for *tea)* and pronounceable numericals (*4* for *for* and *2* for *two*). I discuss this in more detail in the data analysis in chapter 4.

1.2.2. Electronic Mail (Email)

Email is the use of communication of 'letters' mediated by networked computer communication technology. It is a method of transmitting data, text files, digital photos, and audio and video files from one computer to another over the Internet. Each Email user is enabled to compose or write a message for sending. To send the message, the user has to specify the recipient's address. If the user is to send the message to more than one recipient, it is called 'broadcasting'. Similarly, if the user is to send a received message to another person or people, it is called 'forwarding'. Email messages arrive at the mail server from a remote personal computer connected by a modem, or a node on a local-area network. From the server, the messages pass through a router, a special-purpose computer ensuring that each message is sent to its correct destination. A message may pass through several networks to reach its destination. Each network has its own router that determines how best to move the message closer to its destination, taking into account the traffic on the network. A message passes from one network to the next, until it arrives at the destination network, from where it can be sent to the recipient, who has a mailbox on that network. Emails also contain headers and footers above and below the message. They usually state the

sender's name, Email address, and the date that it was sent. A user then can store, delete, reply, or forward the message to others. Most Email programs allow the user to attach files and photos to Emails to send to others. This allows users to append large text- or graphic-based files, including audio and video files and digital photographs, to Email messages. Despite what the World Wide Web offers, Email remains the most important and widely used application of the Internet. For many Internet users, electronic mail has practically replaced the Postal Service for short written transactions.

Statistics in May 2009, by a leading group of messaging analysts called the Radicati Group[3] estimate that there were 1.4 billion Email users in 2009. This is expected to rise to 1.9 billion by 2013. The same source estimates that some 247 billion Emails were sent each day in 2009 and this is expected to double to 507 billion Emails by 2013. In general, 90 trillion Emails were sent in 2009.

1.2.3. Instant Messages (IM)

Instant Messaging (IM) also known popularly as *chat* is a form of synchronous CMC that allows users to exchange typed messages back and forth. IM is mostly referred to as 'chat' but in essence, a chat is a little bit different from IM in that, in Instant messages only two online interlocutors send notes back and forth while in chat, one creates a chat room with a group thus having a sort of group discussion using text and other graphics. In IM, each user defines a list of people that he/she wishes to interact with. IM users can exchange messages with any of the people included in this predefined list (*buddy list* or *contact list*) as long as that person is online; and when one is not online, the instant message is stored and presented as soon as they log in. Typically, the instant messaging system alerts you whenever somebody on your contact list is online. You can then initiate a chat session. Sending a message opens up a small window visible on both screens where the interlocutors can then 'chat'.

[3] Statistics are from: http://www.radicati.com/?p=3237

The IM features enable the users to post an *offline or invisible* sign if they wish to have privacy and not be chatted to. One can also post 'present signs' like; *available, busy,* etc. to announce their presence. They can also put up signs like *stepped out, be right back* etc. These are known as 'away signs/messages'. They announce that a user is still logged on but is temporarily away from their machine in order to alert possible interlocutors not to expect an immediate response to an IM. Baron et al. (2005) assert that as computers are increasingly left on all day, 'away' messages enable IM users to establish a sense of social presence, even when they are not physically at their computers e.g. when having supper, in the bathroom etc. In spite of this, currently, a surprising large number of IM users now post invisible or away messages while sitting at their computers to ward off interruptions.

Unlike Email, IM is faster and allows the user to know if the intended receiver is online at that moment. Also, if one is Emailing back and forth with someone, he/she usually has to click through a few steps. This is why instant messaging (IM) has become so popular.

Besides chatting with text messages, IM users can also

- Share photos and files through IM
- Use voice/video chat
- Make PC phone calls
- Send instant messages to cell phones
- Receive messages while offline
- Personalise IM through use of buddy icons, Smileys and Emoticons
- Join lively discussions online
- Broadcast status e.g. offline, busy etc.

When using IM, the size of an 'utterance' is determined entirely by speaker. Riva (2001:199) writes that in synchronous CMC like IM, 'utterances' are rather short: 5 to 13 words per utterances in 'conversations'. This increases the feeling of interactivity for participants, and lets receivers know that the sender is not idle

and has not finished 'speaking'. In addition, the order of utterances need not be sequentially relevant for meaningful conversation to take place.

IM participants are assumed to be more conscious of the way in which they construct their utterances online. Werry (1996) explains that participants use abbreviations simply to combat the limiting conditions of the medium itself. These, he defines as the pace of channel conversation, channel population and the competition for attention. The use of syntactically-reduced forms, acronyms, symbols, word-clippings is therefore purely for practical reasons. They reduce the time and effort necessary to communicate. Users tend to produce utterances of an average of 6 words. Respect is given to those who can communicate the most information, whether direct or implied in the shortest amount of time. In my view, there is more to it than only the use of language for practical reasons of reducing time and effort. In some cases, it is easier and quicker for the user to type the complete words or phrases than to figure out the most distinct stylish short version or to look up and select a suitable Emoticon. For example, it is easier and quicker to type *hi* than *howdy, how r u,* or by selecting a greeting Emoticon.

Tagliamonte & Denis (2008) did an analysis of English IM by teens to substantiate if it was indeed leading to a breakdown of the English language as had been suggested. They analysed a corpus involving 72 teenagers and over a million words of natural, unmonitored IM. In addition, a corpus of speech from the same teenagers was examined for comparison. They discovered that IM is firmly rooted in the model of the existent language. It reflects the same structured heterogeneity (variation) and the same dynamic, ongoing processes of linguistic change that is currently under way in contemporary varieties of English. At the same time, IM is a unique new hybrid register, exhibiting a fusion of the full range of variants from the speech community that is formal, informal, and highly vernacular. Further, they found out that although IM shared some of the patterns used in speech, its vocabulary and grammar tended to be relatively conservative. For

example, teenagers are more likely to use the phrase *"He was like, 'What's up?'"* than *"He said, 'What's up?'"* when speaking but the opposite is true when they are using IM. According to them, this supports the idea that IM represents a hybrid form of communication. It represents an expansive new linguistic renaissance. They conclude that IM is interactive discourse among friends that is conducive to informal language but at the same time, it is a written interface which tends to be more formal than speech. Therefore, far from ruining teenagers' ability to communicate, IM lets teenagers show off what they can do with language. My view is that the synchronous nature of IM makes it similar to speech and does not employ conservativeness in its grammar and vocabulary. Its unique use of language is based on the need to match the speed of real time speech.

1.2.4. Social Network Sites (SNS)

Social Network sites are defined by Boyd & Ellison (2007) as web-based services that allow individuals to

- Construct a public or semi-public profile within a regulated system
- Articulate a list of other users with whom one shares a connection like common interests, acquaintances and former friends
- View and traverse their list of connections and those made by others within the system.

According to Bodomo (2009:301) the language used in SNS shares many of the characteristics of SMS language. A lot of the SNS language is characterised by what he terms as ghetto-style lexicon, simplified spelling and acronymy. It is a thin, expressive and idiosyncratic language, serving as an identity marker. SNS vary greatly in their features and user base. Some have photo-sharing or video-sharing capabilities; others have built-in blogging and instant messaging technology. There are mobile-specific SNS e.g. Dodgeball, but some web-based SNS also support limited mobile interactions like Facebook, MySpace, and Cyworld (Boyd & Ellison 2007). Some sites are designed with specific ethnic, religious, sex-

ual orientation, political, linguistic groups, geographical regions or other identity-driven categories in mind. Many SNS attract homogeneous populations initially, so it is not uncommon to find groups using sites to segregate themselves by nationality, age, educational level, or other factors that typically segment society even if that was not the intention of the designers (Hargittai 2007). There are even SNS for dogs (Dogster) and cats (Catster), although their owners must manage their profiles.

Users or members of these social network sites are able to connect and network with either strangers or acquaintances with whom they share some common interests. Boyd & Ellison (2007) posit that the primary goal for SNS users is not to meet new people but to connect or network with friends and acquaintances that already exist in their extended social networks. Examples of some SNS include Bebo, Cyworld, Hyves, YouTube, MySpace and Facebook. Some public SNS forums like YouTube mostly connect faceless pen-names of total strangers with similar interests. Apart from the Kenyan daily newspaper comments forum, this study focuses on Facebook, YouTube and Mashada because these are the SNS that are fast gaining a lot of popularity among Kenyan urban youths who have access to Internet. For one to register to these social networks a form with a set of questions about personal details like age, gender, education, location, likes, dislikes, favourite things etc. is provided and the details filled in to create a profile outline for the new member. The member is also encouraged to upload a self-photo to accompany the profile. Some members prefer to upload a fictitious image e.g. a cartoon, flower, car etc. Additionally, some sites allow members to augment their profiles by adding multimedia content. After registration the new member's profile is available for viewing. Sundén (2003) describes the profile as unique pages where one can type oneself into being or existence. The new member then is able to send requests for 'friendship' to other members who share some common interests with the newly registered member. These requests may be send to total strangers or past and recent acquaintances. The requests can be ignored, rejected or accepted thus confirming the friendship. The friends can then access the new member's page. It is worth

mentioning that the new member is to vet and choose the friends who can access his/her full profile and those who are only allowed to access just part of it. The member also has a provision to blacklist some *friends*. The term *friend is* a general term used in social networks to refer to any contact.

In the next section, I focus on the Social Network Sites (SNS) where I collected online comments that I used for the research data.

i) Internet Forums

Internet forums are also known as message boards. Ethan Feerst and Dylan Stewart[4], describe them as an online communication between multiple users. Through such forums, people can share information, experiences, ideas, tips, tricks, etc. The forums mainly use text and are asynchronous in nature. They are based on the idea of a neighbourhood bulletin board, where one posts a message expecting to get reactions. The Internet makes it much easier for people to find specific forums, and board sites to post and respond to information.

A forum consists of a tree like structure beginning with a general heading focused on the forum's content. Next on the structure are different discussion *topics* also known as *threads*. These threads are started by one member who is then referred to as a moderator. The moderator plays the role of a 'chairperson' in the discussion. Finally there are responses and comments to the thread from the members. These responses are called *posts*.

Internet forum participants do not necessarily need to know each other, and neither do they need to share geographical space, time, or language. Their interest in the forum is their common link. Participants can form social bonds and interest groups for any topic of discussion. Such a group can be described as a virtual community. In most cases the participants use a code name. This in a way makes them feel anonymous, free and uninhibited and they can

[4] http://www.videojug.com/expertanswer/internet-communities-and-forums-2/what-is-an-internet-forum

openly comment, discuss and share experiences that they would otherwise be uncomfortable with in 'real life' or face to face.

The current study is confined to forums and message boards that have Kenyans as participants e.g. the Kenyan Newspapers comments section, Kenyan YouTube videos comments, and other online Kenyan discussion forums like Mashada where Kenyans discuss issues related to Kenya. Examples of data from Kenyan SNS sites include;

(1) *obama is a world class leader, our political thugz in the bckgrd have nt even proved 2 b national leaders. where is the comparison pls?????*

Obama is a world class leader, our political thugs in the background have not even proved to be national leaders. Where is the comparison please?????

(2) *Wasee niaje! We maze 2fanye kupressure hao wasee wa EABL, COCACOLA, KWAL NA KBL watunganishie visupa ka SAF COM jo! Drinks za free kuanzia 11 mpaka che! Na kuactivate free drinks lazima uchape kidrink cha guamsa during the dei! Au sio?*

Wasee niaje! We maze 2fanye kupressure hao wasee wa EABL,
Sh Sw Sw+Eng Sw Sh Sw

COCACOLA, KWAL NA KBL watunganishie visupa ka SAF COM jo! Drinks za
Sw Sh Eng Sw
free kuanzia 11 mpaka che! Na kuactivate free drinks lazima uchape
Eng Sw Sw Sw+Eng Eng Sw Sh

kidrink cha guamsa during the dei! Au sio?
Sw+Eng Sw Sh Eng Sw

Hi guys, lets pressurise EABL, COCACOLA KWAL and KBL (drinks companies) so that like Safaricom (a phone network) they begin a plan of flat rates of drinks coupled with free drinks from 23.00 till morning, and activation of the free drinks voucher should be preceded by a huge drink during the day, or not?

Note that Sw indicates Kiswahili, Eng- English, Sh- Sheng and vrn- the Kenyan indigenous vernacular languages.

The provided link [5] leads to a sample of a reaction comment from the Daily Nation forum which is one of the major Kenyan dailies.

(ii) Facebook

Facebook[6] statistics indicate that by March 2010, there were more than 400 million people on Facebook and that half of the users log in every day.

According to its general homepage, Facebook serves to

- Keep up with friends and family
- Share photos and videos
- Control privacy online
- (Find)[7] and reconnect with old classmates (friends, family, acquaintances)
- Discuss interests and hobbies
- Plan parties and other events

Facebook was founded by Mark Zuckerberg who was by then a student of psychology at Harvard University. He launched it in February 2004 as "The Facebook", which was the name taken from the sheets of paper distributed to freshmen, profiling students and staff. Within 24 hours, 1,200 Harvard students had signed up, and after one month, over half of the undergraduate population had a Facebook profile (Hanson 2007:86). It became Facebook.com in August 2005 and had by then spread to all American and UK universities. As of September 2006, the network was extended beyond educational institutions to anyone with a registered Email address. The site remains free to join, and makes a profit through advertising revenue. All that is required to join is to fill a form that is made to constitute one's personal homepage. This personal homepage allows links to one's profile, edits, friends and inbox. The profile displays personal information (one is free to leave out the private information) like the name which can be

[5] http://www.nation.co.ke/News/-/1056/505894/-/u0nx0f/-/index.html

[6] Statistics are from: http://www.facebook.com/press/info.php?statistics

[7] Brackets signify my own additions.

real, fictitious or a made up nick name or code name, gender, date of birth, marital status, political and religious views etc. The edit link allows one to edit any entries, the friends link shows the 'friends' that one has made on Facebook. In most cases these are associates and acquaintances in real life or to a little extent, they may be strangers but with a shared passion (similar to web chats). The personal Facebook homepage also allows visitors to view one's photos or photos of their list of 'friends'. In summary, like any other social network site, each member of Facebook is allocated a page where other members 'friends' can view his/her

- Profile
- Photos
- List of friends, their representative photos and links to their pages.
- Comments and messages left by friends on the 'wall'. The Facebook superwall is a kind of virtual board where a friend or visitor can leave a message.
- Any updates on the profile or the page in general.

Facebook is mainly used to maintain informal contacts, sharing photos, links and videos. Members of Facebook can join activities concerning many themes. Examples of these themes are study related themes, family, entertainment, science or policy. Furthermore, Facebook members can keep their network up to date about events and there is a button 'marketplace' on which products and services are sold. It is worth noting here that there are similar networks to Facebook in different countries, for example Hyves in the Netherlands. Nevertheless, Facebook is more internationally widespread such that members of these other networks sometimes also have Facebook accounts. Facebook as a form of SNS is quickly gaining popularity in Africa among young educated urban professionals (yuppies) who have free access to Internet mainly in their work places or in their homes late at night when the connections are stable. Most recently, it can be also accessed via mobile phones at affordable prices. Facebook in Kenya is mainly used to upload photos and to connect with friends and acquaintances especially those who had lost touch with each other after school,

college, university etc. Most comments are posted on the 'super wall' which is a sort of message board and other members who belong to the member's network or list of friends can view the comments. Most of these comments are casual, jocular or sometimes just plain comments about a photo posted by the member, or any other general situation.

The language used on Facebook is mainly casual and brief. Kenyans use English, Kiswahili Sheng and the vernacular languages depending on the relationship between the members. For example members who belong to the same vernacular language group may use their vernacular language to lock out others. Those who do not share a vernacular language background may use English, Kiswahili, and Sheng. At this juncture I must point out that on 15th June 2009 Facebook was launched in Kiswahili with the name *Sura Kitabu*. This makes it possible for a user to set Kiswahili as the language of his or her page. The initiators of this move claim that it will help preserve the language by popularizing it amongst the youth who are the majority users of Facebook. The projection is that it will target more than 110 million speakers of the language. Examples of messages from Facebook include:

(3) *Supuu!! sasa,*
op u r doin grt!! else MERRY X-MASS en Happy new Year! 2009 jazwad wit muenjoyoz
Hve blessed new bgnin.... oriti.

Supuu!! sasa, op u r doin grt!! else MERRY X-MASS en Happy new Year!
Sh Eng
2009 jazwad wit muenjoyoz Hve blessed new bgnin.... oriti.
Sw+Eng Eng Sw+Eng+Sw+Eng Eng vrn(Luo)

Hi pretty, how are you, I hope you are doing great! Have a Merry Christmas and a Happy new Year 2009! full of happiness and enjoyment. May you have a blessed new beginning? Be blessed, bye.

This study accessed Facebook and collected data from pages belonging to Kenyans. The main area of interest was in the use of language. I first identified the languages used, and then gave a de-

tailed linguistic description of how they were used. Link[8] is a sample of an English homepage from Facebook and link[9] is a sample of the Kiswahili Facebook version which was launched on 15th June 2009.

(iii) YouTube

YouTube is also classified as Social Network Site (SNS) (Boyd & Ellison 2007). It is an international site where members share videos that can be viewed online. According to Cloud (2006), YouTube was founded by Chad Hurley, Steve Chen and Jawed Karim who were all employees of PayPal (PayPal is a payment method that enables any individual or business with an Email address to securely, easily and quickly send and receive payments). The domain name "YouTube.com" was activated on February 15, 2005 and the website was developed over the subsequent months. The World History site 2008[10] reports that YouTube turned out to be something quite different from what its creators had earlier imagined. Their initial idea was to create a dating website based on videos. This was launched in April 2005 but it didn't take off. People did not use it very much. The founders revisited their idea and decided to make YouTube accessible to everyone. They also opened up the interface so that users could choose what they wanted to watch, search for videos, and link to related videos. Additionally, they enabled the ability to 'tag' videos so that others could locate the videos with the use of a keyword. They generally made it easier to post and locate videos, regardless of use or intent. The decision to allow users to tag their videos for better identification and retrieval was an important factor in the site's success. In this way, YouTube progressed explosively. The success of YouTube was coupled with the fact that young people wanted to express themselves on the Internet but they did not have a good way to post videos. There was a video-sharing craze in 2004, spurred by the greater availability of video cameras and cell phones. The tsunami in the Indian Ocean showed the importance

[8] http://www.facebook.com/groups.php

[9] http://news.bbc.co.uk/2/hi/africa/8100295.stm

[10]The World History site can be accessed at http://www.worldhistorysite.com/cthistory.html

of videos taken on cell phones when video-equipped media could not be on the scene. All this amateur-made video created a backlog of materials waiting to be used in personally expressive ways on the Internet.

It is worth noting that before 2004, one could usually not send videos by Email because this required too much bandwidth. The broadband capacity increased for home users in the period between 2004 and 2006. Furthermore, the hosting costs for dedicated servers came down. More people were able to send and post videos, fueling a demand for YouTube. The site also benefited greatly from "viral communication", which means that users could tell and link their friends to YouTube. For example, when a user sent a video via YouTube, the receiver would be informed that they had received a video but to access it, they first had to register and establish a YouTube account free of charge. After gaining popularity, the rules of viewing were made more flexible. Currently, YouTube videos are accessible to both members and non-members.

While it is generally agreed that YouTube as a social network enables people to maintain both informal and formal contacts, its main aim is to work as a Video Sharing Site (VSS). Video sharing sites are sites where members post short videos that can be viewed by others. As already explained, although YouTube is free, one needs to register in order to be a member. The registration is mainly to confer one a personal page with an inbox, account, contacts, playlists, favourites and the sharing of videos option. It also enables one to post or upload videos, send invitations, make friends or become a member of somebody's personal profile. In addition, members are able to make their own playlist, to share videos and categorise their videos. YouTube makes it easy to search for videos of one's interest because it requires video posts to have as many tags as possible to make them easily identifiable in quick searches. When one does not know the exact title of the video, all they have to do is to get to the search box and type any tag describing the video and chances are that the required video will be among the options that will be provided to choose from.

YouTube is used for informal as well as business goals. Furthermore, on YouTube people can share their interests concerning social issues. When uploading a video, the member needs to give a full overview of the video, e.g. the title, description, video category, tags etc. The uploader is also given options for example:

- Whether to allow public or only private viewing (Private viewing allows for the uploader and up to 25 other people to view the video).
- Whether to show the date and map of the location that the video was taken.
- Whether to allow comments automatically by the public or only comments from friends.
- Whether to allow these comments after approving them or not to allow any comments at all.
- Whether to allow viewers to vote on comments or not.
- Whether to allow video responses automatically, to allow video responses after approval or not to allow them at all because after an upload, another member can make a response by uploading their version or another video related to the one initially posted.
- Whether to allow the video to be rated by viewers, or not (Users can be allowed to give points to the video).
- Whether to allow external sites to hyperlink and play the video or not.
- Whether to allow the video to be available on cellphones and televisions or not.

The short videos with each limited to around 7 minutes on YouTube can be of anything ranging from media (music, movies, TV programs) to personal things like showing the best way to shave, swim or the easiest way to put graffiti on public space etc. Nonetheless, pornography is prohibited although users still post it albeit with a lot of editing coupled with other tricks to disguise it. Below each video, there is a provision for viewers to post their comments about the video, and also respond to each other's comments. This is the ideal because the function is overwhelmingly being misused by advertisers of materials unrelated to the

video. Each comment is limited to 500 characters and can be in any language whatsoever. My main interest was the language used in these comments. I visited YouTube video posts related to Kenya like Kenyan music, features, news, culture, shows, politics and Kenyan life in general. I collected data from the viewers' comments beneath the video. Most people who visited a YouTube video site must have a particular interest. For example if one wants to view Kenyan gospel, all they need is to type the phrase "Kenyan Gospel" in the search bar and they will access a long list of all the videos that have been tagged (labelled) as such. After viewing, they have the chance to leave a comment for other viewers. For example,

(4) *mungu nakupenda, hi wimbo ni poa sana. mungu na omba u help me through my problems i believe in u na jua unanipenda MUNGU ASIFIWE!!!!!!!!!!!!!!!*

mungu nakupenda, hi wimbo ni poa sana. mungu na omba u help me
Sw Sh Sw Sw Eng
through my problems i believe in u na jua unanipenda MUNGU
Sw
ASIFIWE!!!!!!!!!!!!!!!
God I love you. This song is very nice. God I pray that you help me through my problems. I believe in you and know that you love me. Praise God!

(5) *banjukeni man life ni fupi we should take it easy while banjukaring*
Sh Eng Sw Eng Sh+Eng
Dance man, life is short and we should take it easy by dancing

The collected corpus data served its purpose by presenting data that could give an idea of how language is used in SNS. In order to have a wider sample, the data is used in combination with similar data from other SNS sites like the Kenyan daily newspapers, Facebook, and other Kenyan SNS forums. The footnote below has a sample of a Kenyan YouTube page[11]. The page clearly portrays not

[11] http://www.YouTube.com/watch?v=2ijbCu4acIg&feature=related

only viewers comments based on the video, but it goes further to show what Bodomo (2009) describes as viewers engaging in discussion based on the video.

1.3. Computer Mediated Communication and Language

In this section, I discuss studies that have tackled CMC in relation to several aspects of language. Hinrichs (2006:21) begins by explaining that in CMC we are not witnessing language change but more of innovative types of language use. As previously explained, this research considers CMC to include both Computer Mediated Communication (CMC) forms like Email, IM, SNS forums and short message texting via the cell phone popularly known as SMS. Computer Mediated Communication (CMC) has become an important alternative to conventional means of communication in an age of rapidly developing electronic communication technology. The evident fact is that language and communication are interdependent on each other. That is, in order to succeed, communication needs some form of language system while on the other hand language exists to enhance communication. Hence CMC cannot survive without the use of some form of language system. As communication crosses the borders of languages and cultures, CMC has become an instrument of international and intercultural communication (Mi-Kyung 2005). All the same, Danet & Herring (2003) report that the early planners of the Internet were generally American, and were implicitly thinking only about how to facilitate communication in English. They did not anticipate the challenges that might arise when other languages were introduced to communicate online. The text-transmission protocol on the Internet is based on the American Standard Code for Information Interchange (ASCII) character set. ASCII, was established in the 1960s, and contains 128 seven-bit codes (unique combinations of 1's and 0's), 95 of which are available for use. This character set is based on the Roman alphabet. The expression "plain text," as in Email and IM, refers to a format that contains only basic ASCII characters, whether written in English or in some other language.

The cyber Atlas (2003) states that already by 2003 roughly two-thirds of all Internet users were non-native speakers of English. This is unlike earlier when native speakers of English dominated the Internet for many years. Danet & Herring (2003) claim that in 2003 only in four of the fifteen top countries online (US, UK, Canada, Australia) was English the official or dominant language. China and Japan together accounted for nearly another fifth of the total. This clearly shows that hundreds of millions of people are participating online and with CMC generally in languages other than English or in some form of non-native English. This scenario is even more pronounced in the use of cell phones. Cell phones are much more widespread than computers and the Internet. The Sub-Saharan Africa experience has indicated that people residing in villages without electricity or Internet networks can still utilise cell phones maximally. These cell phones are designed for the western market and set with mainly English as the base language before finding their way into the African market.

The relationship of language and CMC is now dawning on researchers. Most of these studies have contemplated whether to categorise CMC as written or spoken language. The majority have settled on the fact that it shares qualities from both forms.

The following is a discussion of some studies that have been carried out and brought about new dimensions on the relationship between CMC and linguistics.

It could be claimed that Naomi Baron (1984) was the first to conduct a study on CMC in relationship to language by publishing an article focusing on computer conferencing which speculated on the effects of computer mediated communication as a force in language change. Her hypothesis was that the lack of physical presence had some effects on the language used by the participants. She notes that the lack of physical presence makes participants in CMC aggressive and rowdy because they use all possible means to communicate their point.

1.3.1. CMC and Discourse

Herring (2001) devised the term Computer Mediated Discourse Analysis (CMDA) after claiming that online interaction overwhelmingly takes place by means of discourse. That is, participants interact by means of verbal language, usually typed on a keyboard and read as text on a computer screen. Hinrichs (2006:19) claims that early CMC studies described CMC language as of hybrid nature, displaying a mix of features between spoken and written. An example is Baron (1998:164) who suggests that electronic language is a new phenomenon resulting from contact between the modalities of speech and writing. Some labelled it as conceptual orality in order to explain its informal style. Although text CMC shares some of the spoken word's transience, it offers other traits including simultaneity, which is not possible in spoken language. For instance, one can have conversations with over 20 people in a computer chat room, something not even the most extrovert person can accomplish at a party. Additionally, it is also possible for both or all participants to type at the same time in text CMC a trait that is not possible in discourse, where the participants cannot all speak at the same time. Further on, CMC always leaves some form of trace and for this reason, it is also referred to as persistent conversation (Erickson & Herring 2001). It is possible to lose sight of this fundamental fact at times, given the complex behaviours people engage in on the Internet, from forming interpersonal relationships (Baker, 1998) to implementing systems of group governance (Dibbell, 1993; Kolko & Reid, 1998). Yet these behaviours are constituted through and by means of discourse: language is doing, in the truest performative sense, on the Internet, where physical bodies (and their actions) are technically lacking (Kolko, 1995).

This study concurs with this but generalises online communication to mediated communication in order to accommodate mobile telephony. The current study also proposes that participants of CMC interact by means of a medium which is not purely verbal and only typed. The supposed medium has a close resemblance to verbal communication like in salutation but it differs on

some aspects for example lexical compression which will be discussed in the variables.

1.3.2. CMC and Sociolinguistics

According to Androutsopoulos (2006:419) CMC provides a new empirical arena for various research traditions in sociolinguistics. He also posits that conversely, sociolinguistics can contribute to the interdisciplinary theorizing of CMC by demonstrating the role of language use and linguistic variability in the construction of interpersonal relationships and social identities.

Schler et al. (2006) conducted an interesting sociolinguistic research on the effects of age and gender on blogging. The results were that despite the strong stereotypical differences in content between male and female bloggers, stylistic differences remain more telling than content differences. They concluded that teenage bloggers are predominantly female, while older bloggers are predominantly male. Moreover, within each age group, male and female bloggers discuss different things and use different styles. Male bloggers of all ages write more about politics, technology and money than do their female cohorts. On the other hand female bloggers discuss their personal lives and use personal writing style much more than males do. Furthermore, for bloggers of each gender, a clear pattern of differences in content and style over age is apparent. They also noted that regardless of gender, the general writing style grows increasingly "male" with age: pronouns and assent/negation become scarcer, while prepositions and determiners become more frequent.

Baron (2004) did a study on gender issues in IM. She analysed a corpus of IM conversations by American college students. She found various aspects which reflected differences between genders in the IM chats. For example, female chats were longer and used more Emoticons than male ones. On the other hand, males were reported to use more contractions than females. She concluded that although females were more 'chatty', they used the language more formally than their male counterparts in their chats.

For this research, most of the informants preferred anonymity, a request that I honoured as required by research ethics. I therefore did not collect much data on personal information for analysis. More on this is explained in section 3.1. The closest link to personal information that the present work comes to is in the employment of a uniform group of informants who belong to the same social class as university/college students and young urban professionals and that their age bracket is post youth.

Sociolinguists like Romaine (1984), Eckert (1997) and Androutsopoulos & Georgakopoulou (2003) have carried out extensive studies on youth discourse and concur that the typical sociolinguistic features of adolescent's or youth language include *heavy vernacular use, preference for local varieties, fondness of slang, heavy use of taboo words and modifications e.g. clipping and syllable re-ordering.* Although the current CMC research group age bracket is post youth, I still expect to record some of these sociolinguistic features for example fondness of slang and modifications.

1.3.3. CMC Register

CMC register has for a long time been considered synonymous to a 'non standard register' of language (Thurlow 2007). It has been labeled as *Internet slang, webslang, chattisch, netspeak, netlingua, digital English, textese* and so on. These labels have implied that CMC register is a form of degenerated language. Instant messaging, Emailing and especially text messaging have been, for example, described throughout as destroying, impacting, harming, limiting, damaging, ruining, threatening, massacring, corrupting or eroding standard English and received standards of literacy (Cameron 1995). The current study views language use in CMC as pure innovation and creativity by participants to communicate with each other economically while saving time, space, effort and at times to show off or feel smart.

Crystal (2001) coined the term 'netspeak' in his pioneering work that approached CMC from a linguistic point of view. He defines

netspeak as a type of language displaying features that are unique to the Internet. He refers to it as both a language variety and a new linguistic medium. He further divides 'netspeak' into sub-varieties that are related to different communication genres like Email, chats, and instant messaging. He gives a detailed description of each medium and the specific features of the genres. He curtailed many arguments on whether CMC was written or spoken language by his proposition that CMC is neither written nor spoken language; for him, CMC was a 'third medium' that is in process of evolving its own systematic rules to suit new circumstances. He argued that CMC is developing into a new medium that shows language users at their most inventive, adapting a variety of styles for a variety of purposes of which some are formal and some are highly informal. CMC is fundamentally different from speaking and writing; it shares in their properties, but goes further and does something neither could possibly do.

In addition to this, Crystal is also the first to make a distinction between synchronous and asynchronous CMC. In synchronous CMC, the communicators are required to be available at the same moment (Real time communication) in order for the communication to take place successfully. Examples of these include instant messaging and Chat forums. Asynchronous CMC like Emails and SMS on the other hand does not necessarily require the communicators to be present at the same time for the communication to succeed. Notably Social Network Sites (SNS) can fall into either category. They can be synchronous if the users are logged in at the same time and exchange messages, yet they can be asynchronous in that messages are left for reaction when the other users log in. Both these forms have their own strengths and weaknesses. For example, although the synchronous communication is instant, it does not leave the user with ample time to reflect. It needs immediate reaction and feedback. In addition to this, synchronous communication requires the presence of both users at the same time which may be a disadvantage in cases where one of the interlocutors is indisposed. On the other hand asynchronous communication is disadvantageous in cases where a quick reaction is required.

Bodomo (2009:17) regards this distinction as archaic. He cites cases where people in different continents can synchronously exchange Emails with immediacy almost similar to real time chats owing to the developments in broadband and Internet connectivity. He also argues that when one is offline, IM messages will pend awaiting sign in just like in asynchronous CMC. In this research, I still hold the synchronous and asynchronous distinction because it is closely linked to the principle of rapid communication. However fast Email or SMS exchanges can be, the immediacy does not compare to that of Instant Messages and Chats when all the involved people are present.

All in all, Crystal's work made a large contribution to linguistic CMC studies in terms of language description, non verbal cues like Emoticons, the hybrid combination of written and spoken features and the principal differences between synchronous and asynchronous CMC. Crystal's main omission that the current study fulfils is to embed all the linguistic features of the different genres into their socially situated discourses.

1.3.4. Language Change and Variation in CMC

CMC researchers always acknowledge some form of language change in relation to CMC. The point of contention nonetheless is whether this change is directly attributed to the emergence of the technology. Some researchers like Kress (1998:53) and Luke (2000:83) think that there is more to the changes and find it erroneous to simply attribute them to a technological innovation. Other researchers like Adams (1996:73), Baron (1984:139) and Crystal (2001) closely link the language changes to the technology involved (Bodomo 2009:8). Indeed, new practices of language and literacy may be attributed to a set of unique properties in new communications technology claims Bodomo (2009:23) who holds a broad view on language change in CMC. He describes the change as involving

- change in linguistic forms
- change in the use of language

- modification of existing forms and uses
- emergence of novel, original creations of language.

Relatively few studies in CMC are based on quantitative methodologies. Even fewer make an explicit connection to variationism (Paolillo 2001). Androutsopoulos (2006) partly attributes this to the fact that anonymity in CMC raises problems for traditional variationist methods which assume that reliable information about participant gender, age, social class, race and geographical location is available to the researcher. It can also be argued that this is also due to the absence of phonetics and phonology which is the main type of linguistic variable in the correlative paradigm.

Notwithstanding, some studies have quantified linguistic features like Emoticons, unconventional spellings, representation of spoken language features, regional dialects features, obscenities and codeswitching. (Androutsopoulos & Ziegler 2004; Herring 2003; Huffaker & Calvert 2005; Paolillo 2001; Siebenhaar 2005; Witmer & Katzman 1997). Analyses based on these features demonstrate that language variation online is patterned by age, gender and region. Witmer & Katzman (1997) correlate the frequency of Emoticons to gender. They concluded that Emoticons are used more frequently by females. Conversely Huffaker & Calvert (2005) found that teenage males used Emoticons more frequently.

Crystal (2001) also attempts to point out the existence of variation in CMC communication. He classifies the different variants according to the different genres of CMC and discusses the linguistic features and structure of what he calls *Email language* and *chat language*. He initially dismissed SMS language as a pastime for idle teenagers who now have something to occupy them. Contrary to this, in his latter 2008 book, he acknowledges that both teenagers and adults use SMS language with the former having a higher tendency (Crystal 2008).
The current study tackles variation in relation to the different CMC genres under investigation.

1.3.5. CMC and Multilingualism

According to Diki-Kidiri (2002), languages and cultures are treated differently on the Internet. First, there are "working languages", which people use to communicate on the web. This group is dominated by the worldwide lingua franca like English and French. Other languages of this category are those supported by speaker communities and aspire to play their full part in Internet-based exchanges. An example of this is Kiswahili. Next there are languages which are "spoken about". Among these are a number of languages which have been the subject of linguistic research, and which are only mentioned on the Internet as the subjects of study. Some of these languages are presented quite comprehensively, and indeed lessons in them may even be offered on the Internet, while others are known only by mention of their names. Finally, there is a group of languages which are absent from the Internet. True linguistic and cultural diversity on the Internet should reflect the diversity of the logosphere. Every language ought to have the opportunity to serve as a vehicle for culture and communication on the Internet, which implies the existence of an active Internet community using each language. The current situation is different. Very many languages are represented online. Even many of the minority languages that he calls 'absent' are either currently thriving online through blogs or at least they have risen to the spoken about category. There is also a lot of informal use of language like codeswitching and slang being used online currently.

Research on multilingualism in CMC is in twofold; on one hand the research is on the dominance of English as a lingua franca of transnational communication and on the other hand, the representation of linguistic diversity online (Danet & Herring 2003). Hinrichs (2006) has given this a new angle by focusing on codeswitching on the web. Crystal (2008) also touches on codeswitching in SMS abbreviations. Bodomo (2009:24) explains the dominance of English in CMC as being caused by the facts that

- the Internet begun in the US and naturally adapted English as its language

- English is still regarded as an international language for global information exchange
- English has a fair amount of native speakers and many second language speakers worldwide, such that even native speakers of other languages still have the option to use English as well as their language on the Internet
- there are quite substantial character inputting and encoding problems, which leads to difficulties in using other languages, especially those that do not have an alphabetical writing system

Bodomo identifies some attempts made to make the Internet more multilingual and user friendly for other languages for example by developing codes compatible for most languages, known as the Unicode system in order to solve character inputting and decoding problems. In fact, some inputting systems have been developed for converting alphabetical writing into other orthographical systems such as Chinese and Japanese characters. Additionally, multilingual websites in which the same information is written in several languages on the same website are being popularised. Furthermore, instant real-time multilingual translators for websites, like Systran[12] and AltaVista Babel Fish[13] have been developed.

UNESCO Institute for Statistics in 2005 observes that unlike before, there is now a slightly weakening dominance of English on the Internet. This dominance could have been weakened more were it not for the fact that the world's richest multilingual areas are on the 'wrong' side of the digital divide. The presence of lesser-used languages on the Internet crucially depends on localised software and computer fonts but their availability in tandem depends on the market volume of the respective populations (Maurais 2003; Ouakrime 2001).

[12] http://www.systransoft.com/Papers/ppr_alta.htm

[13] http://world.altavista.com/.

The impact of technology is particularly manifested in the romanised transliteration of native scripts that is reported for, among others, Greek, Arabic and Persian (Palfreyman & Khalid 2003). Often diverging from official transliteration systems with innovative correspondences between native and Roman graphs, these vernacular language transliterations seem to persist, despite the development of Unicode especially in settings of transnational and diasporic contact (Maurais 2003).

On the representation of linguistic diversity online, Debski (2004); Ouakrime (2001); Sperlich (2005) and Warschauer (2000) concur that the Internet may contribute to the maintenance of endangered and minority languages by providing a space for their documentation and literacy promotion. According to their studies, the Internet affords small languages an increase in written language domains and endows them with prestige by demonstrating their compatibility with technology and modern communications media. Although in any case, the success of these initiatives ultimately depends on the active participation of the population concerned which often lacks the required technology and computer literacy. In some cases even if these requirements are met, the use of small endangered languages does not come automatically or is very limited for example only in phatic communication.

On language choice, Wright (2004) performed a comparative study by investigating language use online by educated speakers in various countries including Indonesia, Italy, Japan and Ukraine. He found out that the use of English decreases when CMC resources become available in users' own languages. Notwithstanding, reported language choices also varied according to the communication mode and the web content. Other studies like Durham (2003) suggest that English is favoured as a lingua franca of professional communication in multilingual networks. It is also worth noting that seemingly, the lack of institutional constraints and the 'triumph of informality' in vernacular languages in CMC encourage the 'literalization' of varieties that were traditionally confined to spoken disc.

1.3.6. Other Languages in CMC

Three main comprehensive studies similar to this investigating and comparing the use of text language in the different CMC genres have been carried out on Swedish by Hård af Segerstad (2002), English, Swiss and German by Frehner (2008) and English, Chinese and some French by Bodomo (2009). So far none has been carried out in a multilingual context in Africa or more specifically Kenya. This fact was one of the strongest motivations for this study. The multilingual nature of Kenya with over 42 indigenous vernacular languages coupled with English and Kiswahili makes the use of language in Kenyan CMC very interesting in terms of the language choices coupled with the linguistic creativity and manipulation of language to comply with the limitations that these CMCs offer. This study endeavours to give a detailed description of these practices.

Hård af Segerstad's (2002) investigation was to find out how Swedish written language is used and adapted to suit the conditions of CMC. Similar to this research, the study dealt with texts from four genres of CMC that is; SMS, Email, Web chats and Instant Messaging. The study also incorporated a corpus of traditional handwritten letters. The main difference with this study besides the languages of focus is in the approach of collecting the Email and traditional handwritten letter corpus from private individuals to an anonymous authority of a city council. This may not have been quite representative data because the sender and receivers were unknown to each other and most of these Emails and letters were official and mainly bore complaints and requests unlike mail between peers, friends, family (people who knew each other well) which I used for the corpus. On data analysis, Hård af Segerstad mainly dealt with quantitative data by using the TraSA (Transcription Statistics tool with automation) software. In the findings, Hård af Segerstad proposes that three interdependent factors influence language use: synchronicity (instant/delayed), means of expression (communication mode) and situation (context). She claims:

> "Production and perception conditions such as text input technique, limited message size, as well as situation parameters such as relationship between communicators, goal of interaction are found to influence message composition." (Hård af Segerstad, 2002:93)

In line with my study, Hård af Segerstad also contests the popular assumption that language is deteriorating because of increased use in CMC. The results show that language is used creatively and adapted to suit different CMC genres.

Frehner (2008) did a comprehensive study on English, Swiss and German in order to ascertain whether Emails and SMS were speech or writing. It was discovered that they are hybrid because they make use of both literacy and orality. This study went on to compare Emails and SMS to phone calls and telegrams. The result was that in a way, text messages can be regarded as the renaissance of telegrams. On language use, they concluded that although CMC language use showed new trends, it actually used old features. It was claimed that many features of CMC and particularly text speak are not actually new. This concurs with Bergs and Kesseler (2003) who concluded that SMS display similar characteristics as common letter writing in the 19th century mainly in terms of lexical reductions and lack of spaces between words. In addition to the different languages involved, Frehner's main deviation of her study from this is in her data collection. The data was collected by means of an electronic and paper based survey respectively. The SMS corpus was collected from questionnaires where informants were asked to copy SMSs that they had received. The main weakness of this approach was that it could not be ascertained whether the informants gave accurate copies of the SMS. Writing down SMS using a pen and paper is more conscious and thus different from typing it on the phone. The informants could have easily corrected what they considered as errors or edited the messages. In addition, this approach could have been an infringement of ethics because the messages used were submitted for research by the receivers without the sender's consent. I required informants to forward their own messages to me. They

are not informed about what I am looking for in order to discourage them from editing the messages.

Another interesting study is Ling (2004) who explored the intersection of linguistic and social aspects of SMS. Ling's main interest was in the SMS texting culture among Norwegian teens, particularly females. Ling reports that teenage girls sent more text messages, used more complex syntax, included more salutations and closings and even employed better punctuation than their male or their older counterparts. Ling's conclusion was that females are more sophisticated users of the medium. This is consonant with other research findings that female writing (and also speech) tends to approach normative standards more than that of men (Labov, 1994; National Centre for Educational Statistics, 2002; Baron, 2004). Unlike Ling's study, my study does not approach CMC based on differences of language use through gender. My focus is on the general language use in Kenyan CMC.

In conclusion, CMC involves language. The emergence of CMC has had a major influence on adapting language to these media and at the same time, language has been a major influence on CMC with manufacturers trying to adapt CMC to languages.

1.4. The Current Research

The current research focuses on issues concerned with CMC language use in Email, SMS, IM and SNS in Kenya. Its objectives, rationale, hypotheses, and methodology are elaborated hereafter.

1.4.1. Research Objectives

This research has two main objectives:
1. The first is to analyse the use of language in CMC texts as compared to the everyday formal and informal *standard* language use. This will in turn reveal

a) in what ways it deviates from the standard language use,
b) what motivates these deviations, and
c) what arises as norm in CMC language.

2. The second objective is to discuss the similarities and differences among the individual CMC genres in terms of their specific registers. Register in this case is used to refer to the unique language use for each genre.

1.4.2. Rationale of the Research

The few linguistic researches on CMC have mostly focused on English, European and Asian languages probably because CMC evolved from these regions. CMC in Sub-Saharan Africa in relation to language has hardly been researched into despite the region being a host to diverse linguistic groups. Danet and Herring (2003:3) aptly claim that

> "To date however, the research literature in English on Computer - mediated communication has focused mainly exclusively on emergent practices in English, neglecting developments within populations communicating (online) in other languages." (Danet & Herring, 2003:3).

The current research investigates the use and characteristics of Kenyan languages in text-based CMC. The focus on Kenya as an African country is particularly interesting because of

i. the rapid growth of CMC in the region.
ii. the multilingual context in the region.
iii. the emerging literacy especially in vernacular languages.

As already mentioned, Kenya is a highly multilingual country with over 42 indigenous vernacular languages spoken besides English and Kiswahili which are lingua francae for national and official communication. This research confines itself to the use of text-based language in the following four genres of CMC:

- Short text messaging (SMS)
- Electronic mail (Email)
- Instant Messaging (IM)
- Social and Video Network sites (SNS e.g. Facebook and Mashada forums and VNS e.g. YouTube).

Besides being all text-based, these four CMC genres have been selected because they are relatively easily accessible to the population group (see section 3.1 for further discussion on the population group).

The four CMC genres fairly represent both the synchronous and asynchronous forms of CMC. The selection of YouTube and Facebook as part of the SNS genre further represents what Bodomo (2009:10) describes as video-based CMC. He claims that there is an emerging paradigm shift from purely text-based CMC to video-based. He says that

> "[...] a new theme has emerged from text-based CMC to video-based CMC. Video-based Computer-Mediated Communication may be defined as interaction and transfer of information through the medium of the computer and related digital devices mainly in the form of dynamic image streams. Most contemporary social networking tools like Facebook and YouTube are implemented with video-based CMC. Of course, Video CMC still contains the written word, but the written word is mainly meant to just express talk around the main issue, the Video event. Young users of the Internet have radically moved away from communication through the plain written word to communication in the medium of video clips and voice-image interactions through video-based media such as Facebook, YouTube, video games, and Skype." (Bodomo, 2009:10).

Despite my main focus being on the use of language in the CMC text discussions, I have deliberately included YouTube which is a Video Network Site (VNS) as a form of Social Network Site (SNS) in order to have a balanced representative sample of the SNS genre composed of both Video Network Sites (VNS) and other text-based SNS forums. My focus on text-based CMC is motivated by various reasons. First of all, for practical reasons, text-based CMC is selected because given its relatively low costs, it is especially endeared to the youth who make use of it, which

consequently makes it easy to set up a corpus. Secondly, text-based CMC is a new invention of presenting language in a visual way and needs to be researched into. Thirdly, text-based CMC always leaves a trace, a fact that has proved beneficial in the process of data collection. Finally, as already mentioned, not much research has been carried out on sub Saharan African CMC in general despite Africa's growing CMC market coupled with the multilingual nature of the continent.

1.5. Hypotheses

Several principles are tested as hypotheses in this research. These are the principles of rapid communication, least effort, mode limitation, and informal communication that includes codeswitching and peer communication. The principles of rapid communication, least effort and mode limitation have taken the lead in explicating CMC practices, especially in the SMS. They have been used by CMC researchers like Schlobinski et. al. (2001), Döring (2002), Hård af Segerstad (2002), Frehner (2008) and Bodomo (2009). I adapt them in order to test their applicability to my data. These three were my initially set hypotheses, but after going through the data, I observed additional crucial dimensions of codeswitching and peer communication that necessitated me to incorporate the principle of informal communication. There was a general recurrence of codeswitching between vernacular languages, Kiswahili and English in all the CMC genres, a fact which required a detailed discussion. Additionally, there was the use of Sheng and Engsh (cf. 2.1.4) and general creativity by the population group who comprised college/university students and young urban professionals. This formed a peer group of the educated and the technologically savvy (cf. 3.1).

1.5.1 Principle of Rapid Communication

Most features of CMC are triggered by the need for rapid communication. This in turn suggests that synchronous CMC like IM displays more features similar to speech than asynchronous CMC (SMS, Email and SNS). This is because synchronous CMC is hurried in order to make it flow with turn taking and acceptable time frames like face-to-face communication, while asynchronous CMC

is more consciously composed. Therefore IM leads in rapidity, followed by SMS, then finally Email and SNS. This principle has been confirmed for synchronous CMC by Hård af Segerstad (2002), Rheingold (2008), and Bodomo (2009). Examples of CMC messages composed rapidly include;

(6) *sori im hurryn nt my ofis bt m waitin. Mis u 2*
Sorry I'm in a hurry. I'm not in my office but I'm waiting. I miss you too.

(7) *g2g now…cach u 2moro*
I've got to go now. I'll catch you tomorrow.

Message (6) displays an omission of grammatical words for example the 1st person pronoun *I,* the auxiliary *am (I'm)* and the preposition *in*. Message (7) uses a phrase acronym *g2g* for *got to go* and also omits the 1st person pronoun *I* and the modal verb *have (I've).* All these features of missing forms and use of acronyms arise from the rapidness involved in composing the message.

1.5.2 Principle of Least Effort

When confronted with text input in CMC, users will often choose the most convenient input that requires the least effort to avoid strain. A good example of this is the neglect of capital letters at the beginning of sentences/messages and proper nouns. This principle was put forth by Zipf (1940) and has been used for CMC researches by Schlobinski et.al. (2001), Hård af Segerstad (2002), Frehner (2008) and Bodomo (2009). They all affirm the employment of least effort or economy in CMC texts. An example where the initial capital letter of the message and that of a proper noun is ignored is:

(8) ***t****utaenda* ***n****ai kesho*
T*utaenda* ***N****airobi kesho.*
We will go to Nairobi tomorrow.

Besides the omission of the capital letters, the proper noun *Nairobi* is also shortened to *nai* for economy and least effort.

1.5.3 Principle of Mode Limitation

This is also known as the principle of least space. Limited buffer size results in conscious, carefully edited input with a lot of codeswitching and clippings for economy. Users would rather resort to apply these strategies in order to save space e.g. the following example is a message from the data. This message is written using codeswitching between English and Kiswahili. It appears as:

(9) *dei fungad our uni juzi*
Eng Sw+Eng Eng Sw
They closed our university the day before yesterday.

Equivalents of the message in English or Kiswahili show that the initial codeswitched version is relatively shorter and uses the least space thus saving the user a lot of space.

> Initial message = *dei fungad our uni juzi* (23 characters)
> English equivalent = *they closed our university the day before yesterday* (Eng=51 characters)
> Kiswahili equivalent = *walifunga chuo kikuu chetu juzi* (Sw=31 characters).

Apart from clipping, in this case, codeswitching has also been used to shorten the message. In the Kenyan scenario, this practice is expected mainly in SMS and SNS. In SMS, a shorter message within the limit of 161 characters is cheaper, otherwise it would be split and charged as two messages if it surpasses 161 characters. Similarly, many SNS sites limit the number of characters to 500. This limitation has an influence on the language used in the message. IM and Email do not have mode limitation.

1.5.4 Principle of Informal Communication

This principle encompasses both codeswitching and peer communication as practices of informal use of language. Informal communication is frequent, interactive, and expressive. It is traditionally mediated by physical proximity. Indeed the occurrence of informal communication features in the CMC data shows that indeed CMC is shrinking distances and recreating settings similar to close physical proximity. Generally, SMS and IM are the most in-

formal genres owing to the hypotheses discussed. Email is the most formal although its formality is not close to formal letters. The level of formality in SNS is variable depending on the network site. Entertainment and discussion sites like YouTube and Mashada are likely to be more informal than commentary sites like on the online Daily Nation comments section. Codeswitching and peer communication and identity will be discussed under this principle of informal communication.

Codeswitching

Owing to the multilingual nature of Kenyans, CMC texts are likely to display frequent instances of codeswitching for reasons like identity, economy, accuracy and even show off or at times just for fun. Another interesting observation in codeswitching is the use of a different language for salutations. Codeswitching between English, Kiswahili and Sheng is very common in the Kenyan CMC and may act as an identity marker for students and yuppies. An example of CMC codeswitching includes,

(10) *thengiu!!! some few years ago it hit the headline gutuikite ati funda icio ciohwo nappy citige guthukia mazingira.*

thengiu!!! some few years ago it hit the headline gutuikite ati funda icio ciohwo
Eng Kikuyu
nappy citige guthukia mazingira.
Eng Sw
Thank you! Some few years ago, it hit the headlines that those donkeys should be tied a nappy so that they stop making the environment dirty.

The observed codeswitching is,

thengiu – It is the English *thank you* but has been written the way it would be pronounced in Kikuyu.
some few years ago it hit the headline - English
gutuikite ati funda icio ciohwo-Kikuyu (vernacular language)
nappy- English
citige guthukia-Kikuyu
mazingira- Kiswahili

Peer Communication and Identity

The population group of this research forms a kind of peer group. They are all educated and well-placed academically/ profession-

ally in society (cf. section 3.1.). Peer communication on the whole uses peer language. Youth peer communication in CMC also follows suit. The language use is free and natural. Since the participants belong to this kind of peer group, I assume that their language usage in CMC employs spontaneous use of natural language unlike in communication between non peers. For example, in order to be as close as possible to natural language, the peers use Sheng which their peer group understands in order to express their identity as a separate group. It is notable though that the peer language here is not only for its own sake but also serves a purpose of show off. This show off in most cases occurs in the modification of the language of communication and even the use of 'new' vocabulary. This is captured in the use of Sheng in the messages.

Peer communication and identity has been set to describe the use of language in terms of style associated with the peer group. This style includes the purposeful deviations from the official norm and the use of different forms of creativity. Peer communication in CMC is expected to yield a high frequency of free and natural language.

All the four hypotheses, that is: the principles of rapid communication, least effort, mode limitation and informal communication which involves peer communication and identity, are tested in the corpus database set up.

Chapter 2. Language and Communication in Kenya

This study aims to shed some light on the current practices and use of text language in Computer Mediated Communication (CMC) in Kenya. In a bid to do this, it is important to present an impression of the language and CMC situation in Kenya. This chapter presents the necessary background to both the language and CMC situation of Kenya.

2.1. Language in Kenya

The African continent generally constitutes a highly complex multilingual area. The complexity results from the high number of languages, the distribution of these languages, the relatively low numbers of speakers per language, the intensive language contact in many areas of the continent and the widespread multilingualism in the continent. As Li-Wei (2000) aptly puts it in his introduction as an editor, "Africa's heterogeneity is reflected in language. Per capita there is a wider range of languages in Africa than in any other continent in the world". This view is also shared by other researchers like Makoni & Kamwangamalu (2000), Prah (1998) and Diki- Kidiri (2001).

The introduction of new communication technologies and in particular CMC into the African community has posed a challenge in terms of how to best utilise languages in order to communicate effectively using them. It is given that these technologies have not been manufactured for the African community but are mass-produced elsewhere and exported into Africa with their original specifications adapted to foreign languages.

Most people in Africa speak one or more indigenous vernacular languages, as well as an indigenous lingua franca, e.g. Kiswahili in Kenya, which has become the medium of communication between different ethnic groups or speech communities. Such individuals may also speak a foreign language such as English or French introduced to the communities as a consequence of colonization or during the process of international communication. The latter language is often the official language of education, bureaucracy

and privilege. The Kenyan scenario is no different. Besides having English and Kiswahili as the general lingua francae, the country comprises approximately 42 indigenous vernacular languages which are distributed across 7 provinces, excluding the plurilingual Nairobi province/area that encompasses the capital city and does not have a distinctive vernacular language. The 7/8 provinces that constitute the republic of Kenya are, in effect, linguistic units[14].

The Kenyan languages fall into 3 linguistic families of Bantu, Nilotic and Cushitic groups. About 65% of Kenyans speak a Bantu language, for example; *Luhyia*, *Kamba*, *Kikuyu, Kisii and Mijikenda*. 30% are Nilotic including *Kalenjin*, *Luo,* M*aasai*, *Samburu* and *Turkana*. The Cushitic family, mostly composed of *Somali and Rendile* speakers, represents about 3% of the population. The remaining 2% are speakers of European, Indian or other languages (Marhoum & Samper 2003). The number of speakers of these vernacular languages varies. For example according to the World Ethnologue report[15], a Cushitic language like Yaaku also referred to as Ndorobo, which originates from the Laikipia district in the Rift valley province only has a handful of speakers while Kikuyu, a Bantu language from the central province, has approximately 7,180,000, which constitutes 15% of the total country population. The Maasai, who are Nilotes from the Rift valley province, are approximately 590,000 in Kenya (2009 Ethnologue report). Speakers of the Kenyan vernacular languages interact freely, leading to language contact and multilingualism.

I will now give an overview of the Kenyan languages that this study deals with. These are Kenyan indigenous vernacular languages, Kiswahili, English and Sheng.

2.1.1. Indigenous Vernacular Languages

Kenya's population was given as 36,913,721 in 2007[16]. As noted

[14] See inset see maps of Kenya showing the 8 provinces and the languages spoken on http://lcweb2.loc.gov/frd/cs/profiles/Kenya.pdf

[15] World Ethnologue statistics at: http://www.ethnologue.com/

[16] http://lcweb2.loc.gov/frd/cs/profiles/Kenya.pdf

earlier, Kenya is a highly multilingual area but the exact number of its languages has not been established. A reason for this is that it is generally difficult to establish 'the number of languages' because there is no universally accepted notion of 'a language'. Additionally, in Kenya, the census reports are manipulated to suit the interests of the government of the day. Suffice it to say that the total number is approximated at 42 indigenous vernacular languages which are actively spoken (Mbaabu 1996:147). This study refers to these indigenous languages generally as vernacular languages. Some of these vernacular languages form dialect clusters. For example the Luhyia group (Bantu) has 19 dialects of which some are mutually intelligible while others are not. To expound more, each of the Kenyan vernacular languages has its *homeland* so that linguistic differences largely coincide with regional differences as shown in figure 4.

For research purposes, I consider all indigenous vernacular languages to function at the same sociolinguistic level. Kenyan vernacular languages are used in similar ways in relationships existing between people from the same ethnic group. They are typically used in the daily lives of members in the speech area, while they are mainly used at home (if at all) outside the speech area. The use of these vernacular languages generally may be triggered by various contexts, for example

- for communication amongst friends and relations sharing a similar vernacular language group;
- to exclude others e.g. in secret agendas;
- to show allegiance to one's ethnicity e.g. in politics;
- to avoid inaccurate translations by using vernacular language words or phrases that capture the exact meaning intended.

In contrast to this, I note that Kenyan vernacular languages are hardly used in writing or in official communication but this may be changing with the emergence of CMC. Generally, these vernacular languages are all comparable in their usage and therefore this study treats them equally under the tag 'vernacular' (vrn).

In relation to the vernacular languages, this research intends to find out and describe the general degree and style of vernacular language use in text CMC in Kenya.

2.1.2. Kiswahili

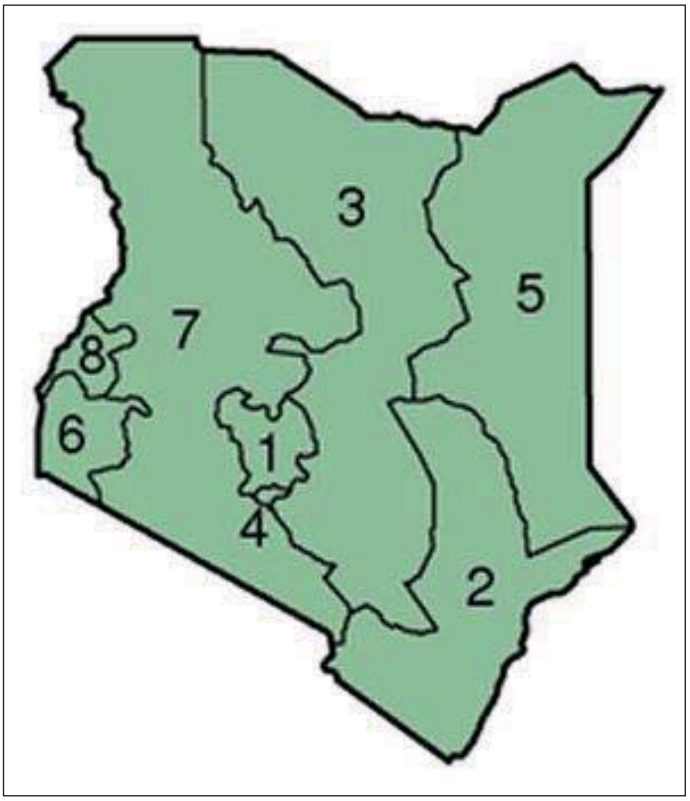
Besides the indigenous vernacular languages in Kenya, there is also Kiswahili which functions as a lingua franca both nationally and in parts of East and Central Africa. Kiswahili is accorded the status of a national language and is spoken by roughly 95% of the population. It is mostly used as a lingua franca for inter-ethnic communication.

According to Nurse & Spear (1985:1), Kiswahili is a language that originated at the East African coast. Long-time interactions with other people bordering the Indian Ocean spread the Kiswahili language to distant places such as on the islands of Comoro and Madagascar and even far beyond to Oman and United Arab Emirates. Trade and migration from the Kiswahili coast during the nineteenth-century helped spread the language to the interior of particularly Tanzania. It also reached Uganda, Rwanda, Burundi, Congo, Central African Republic, and Mozambique.

Currently Kiswahili is spoken in many countries of Eastern Africa. For Tanzania, deliberate efforts were made by the independent nation to promote the language, thanks to the efforts of the former head of state, Julius K. Nyerere.

Kiswahili is the national as well as the official language in Tanzania - almost all Tanzanians speak Kiswahili proficiently and are unified by it. Tanzania's special relation with countries in southern Africa was the chief reason behind the spread of Kiswahili to Zambia, Malawi, and other neighbouring countries.

Figure 4: Map showing the 8 provinces in Kenya



Key: Province and major languages spoken[17]

(1) Central province – Kikuyu

(2) Coast Province – Mijikenda (Taita, Digo, Giriama, etc)

(3) Eastern Province – Kamba, Meru, Embu

(4) Nairobi Area – Plurilingual capital city

(5) North Eastern Province- Somali, Borana, Rendile

(6) Nyanza Province – Luo, Suba, Kisii, Kuria

(7) Rift Valley Province – Kalenjin (Nandi, Kipsigis, Keiyo, Marakwet etc.), Maasai

(8) Western Province – Luhyia (Bukusu, Maragoli, Isukha etc.), Sabaot, Teso

[17] For a complete list, see http://www.kenya-advisor.com/kenya-map.html

In Kenya, Kiswahili is the national language, but official correspondence is still conducted in English. In Uganda, the national language is English but Kiswahili enjoys a large number of speakers especially in the military. As a matter of fact, during the Idi Amin rule from 1971-1979 Kiswahili was declared the national language of Uganda. However, the declaration has never been seriously observed nor repealed by the successive governments.

Thus, Kiswahili is the most widely spoken language of Eastern Africa and many world institutions have responded to its spread. It is one of the languages that feature in some world radio stations such as the BBC, Radio Cairo (Egypt), the Voice of America (U.S.A.), Radio Deutsche Welle (Germany), Radio Moscow International (Russia), Radio Japan International, Radio China International, Radio Sudan, and Radio South Africa. The Kiswahili language has also made its presence in the art world - in songs, theatres, movies and television programs. For example, the lyrics for the song titled *Liberian girl* by the late Michael Jackson have Kiswahili phrases: *Nakupenda pia, nakutaka pia, mpenzi we!* (I love you, and I want you, my dear!). The well-celebrated Disney movie, *The Lion King* also features several Kiswahili words, for example *Simba* (lion), *rafiki* (friend), as the names of the characters. The Kiswahili phrase *hakuna matata* (No troubles or no problems) was also used in that movie. The promotion of Kiswahili language is not only in its use but also deliberate efforts have been made throughout the world to include it in the education curriculum for higher institutions of learning in many parts of the world including the Netherlands, Germany, Finland, U.S.A, Japan, and China. Additionally, it is now used on websites such as Facebook and Wikipedia. With the expanding growth of Kiswahili both nationally and internationally, it is my hope that this study will provide insight into the frequency and style of its usage in Kenyan CMC.

2.1.3. English

English was brought to Kenya by the missionaries and was spread further by the British colonisers. After these groups left, it was cheaper and more practical to endorse the use of English for the

running of all the concerned activities. Thus currently, English is the main official language. It is the language of education, commerce and communication, judicial, parliamentary debating and administration in Kenya.

It is spoken by roughly 75% of the population and is practically the main official language. As Kembo-Sure (1994) aptly puts it,

> "In Kenya, English has acquired such functional prominence that the question now is whether to consider it as "another people's language," or to regard it as just another Kenyan language." (Kembo-Sure, 1994:69).

All school-going Kenyans are required to learn English as a compulsory subject. After the first four years of school, English becomes the language of instruction in the schools and institutions of higher learning. Therefore all Kenyans who have been to school have acquired some proficiency in English.

According to Muthwii & Kioko (2004),

> "English (in Kenya) is mainly learned in formal educational settings. Because it is an important language for participation in the public domain almost all Kenyans with some education have acquired English, albeit with certain variations." (Muthwii & Kioko 2004:36).

Sociolinguists' analysis of discourse based on speech of multilingual speakers (cf. Muthwii, 1986; Myers-Scotton, 1993b) demonstrate that intelligence, ambition, expertness and confidence in many formerly colonised parts of the world are attributes that have been associated with the use of the English language vis-à-vis the use of ethnic languages. Over the years, as English was used as a language of power in these regions, these attributes became part of the social meaning of the English language.

This is the position of English in Kenya. Many use it as a language of prestige and power. It is expected that this study will shed more light on the usage of English in CMC in Kenya.

2.1.4. Sheng and Engsh

Sheng is one of the world's most remarkable and intriguing codes. King'ei (1987) and Mazrui (1995) describe it as a hybrid linguistic code that defies the classification categories of pidgin, creole, slang, or jargon. It is a language variety that emerged from the youth in the poor neighbourhood of Nairobi Eastlands in the 1960s and 1970s (Mazrui 1995). Its evolution and use has been attributed to a variety of factors ranging from non-linguistic factors, to language contact and amalgamation of various languages, inadequate knowledge of standard Kiswahili and English and so on (Osinde 1986, Mazrui 1995, Samper 2002 among others).

Ogechi (2005) points out that word/morpheme order in Sheng largely conforms to the Kiswahili word/morpheme order. The positions of grammatical morphemes on inflected Sheng words also appear to be identical to those in Kiswahili words. In addition, there appears to be a shared surface form of some of the inflectional affixes on both Sheng and Kiswahili words. In spite of this, these affixes do not always follow the type of concord required by Kiswahili syntax when they are used in Sheng constructions. Furthermore, the affixes are used on Sheng lexemes regardless of whether they are sourced from Kiswahili, English, other Kenyan indigenous languages or the coined ones. Ogechi continues that this implies that it is possible to identify Sheng lexemes but it is difficult to posit a Sheng morphosyntax and as such Sheng participates in codeswitching as a code largely identifiable through its lexemes.

A mirror image of Sheng is Engsh which I describe as a code within Sheng. According to Kießling & Mous (2004), Engsh developed as an anti-language in the richer neighbourhoods of Nairobi (Westlands) in reaction to Sheng. It consists of an English base with insertions from Kiswahili and other languages. It is a kind of Sheng used in the more affluent neighbourhoods by the elite

youths. Engsh uses a lot of words in the English forms coupled with Kiswahili. On the syntactic level, Mazrui, (1995) explains that Sheng exhibits a Kiswahili syntactic base form while Engsh takes on an English base. For Sheng, the Kiswahili word order is prevalent while the English word order is prevalent for Engsh.

The following are examples of Sheng and Engsh. Examples (11), (12) and (13) are Sheng while (14) and (15) are Engsh. These examples give a more or less clear differentiation between Sheng and Engsh, but it is not always easy to identify which is which in the data since there is a lot of mixing leading to unclear cases. The underlined words in (11) indicate the English vocabulary additions to the Kiswahili base form while *alihora* in (13) shows a combination of a Kiswahili base form a-li- (third person, past tense) and vernacular (Kikuyu) vocabulary *hora* (beat). In Sheng, this word *hora* is used with the slang meaning of the phrase beat it. The underlining in (14) indicates Kiswahili vocabulary additions to the English base. The bold word in (15) *bwog* is from *bwogo* (fall/defeat) which is in fact a word from Luo; one of Kenya's vernacular languages. This example shows a mix of English base forms and vernacular language vocabulary which has been adopted into Sheng.

(11) *Hope umeklia jobo coz tonite tutahave bash moja noma usihate coz ni ya kukata na axe.*
Hope ume-clear jobo coz tonite tuta-have bash moja noma usi-hate coz ni ya kukata na axe.
Eng Sw+Eng Eng Sw+Eng Sh Sw Sh Sw+Sh Eng Sw Eng
I hope you have finished your work because we will have a very exciting party tonight. Please don't miss it.

(12) *nitatry kucome jioni but msiblemiane haki*
nita-try ku-come jioni but msi-blem-iane haki
Sw+Eng Sw+Eng Sw Eng Sw+Eng +Sw Sw
I will try to come in the evening. but please don't blame each other.

(13) *imagin alihora na hummer yangu! wacha nitamuunleashia stress eish!*
imagin ali-hora na hummer yangu! wacha nitamu-unleash-ia stress eish[18]!
Eng Sw+vrn Sw Sh Sw Sw Sw+Eng+Sw Eng
Imagine he disappeared with my car! Just wait, I'll (stress him) give him a hard time esh!

[18] '*eish*' is an irritation marker.

Although the messages above have a mix of some English, Kiswahili, Sheng and vernacular languages vocabulary, it is notable that the base morphology is Kiswahili. In contrast, the following messages have vocabulary from the different languages but the morphology is English.

(14) They changad some becks for him to use in liparing the loan before he uzas the hao

They changa-d some becks for him to use in lipa-r-ing the loan before he uza-s the ha-o
Eng Sw+Eng Eng Sh Eng Sw + Eng Eng Sw+Eng Sh
They contributed some money to help him service the loan before he sells the house.

(15) Raila's team is unbwogable, kwanza their hummers are hot!

Railas team is un-**bwog**-able, kwanza their hummer-s are hot!
Eng Eng+vrn+Eng Sw Eng Sh+Eng Eng
Raila's team is unbeatable, in fact their cars are really good.

This study views Engsh as a code within the code Sheng. Sheng is the source of Engsh. The target population who are mainly university students and young professionals use the two language varieties not only interchangeably but also in the same constructions such that they overlap. This makes it difficult to distinguish between what is Sheng and what is Engsh. For example when the students and young professionals use text CMC to communicate with their peers outside the university, they may use Sheng with Engsh insertions to show off their high level of education. On the other hand when they communicate with their fellow university peers, they may use Engsh with Sheng insertions to flaunt their prowess and toughness that is associated with Sheng.

At the quantitative level, this study will be more concerned with identifying Sheng/Engsh vocabulary in general as compared to the other languages which are English, Kiswahili and vernacular languages. It is not surprising for this study to observe that Sheng and Engsh provides some complexities in relation to codeswitching similar to what Kießling & Mous (2004) observe in their claim that codeswitching is extensively used in urban youth languages

and plays a significant role but it is insufficient in itself to describe and explain the phenomenon of urban youth languages. Myers-Scotton (1993a:213) also captures this in the claim that urban youth languages are similar to codeswitching since they often arise from situations where codeswitching is the unmarked choice.

In relation to this, Mazrui (1995) characterises Sheng as codified Swa-Eng codeswitching. Thus two main properties distinguish Sheng from mere codeswitching. First, there is a strong norm imposed on certain *switches*, in that certain lexemes will be conventionalised as switches and there will be no choice as to switch or not to switch. For example in Sheng, some vocabulary is conventionalised e.g.

- *keja: house (from cage)*
- *bash: party*
- *matha: woman*

Thus by discerning their meaning, it becomes easy to classify them as standard Sheng vocabulary and not as codeswitching. Secondly, there are lexical items that are peculiar to the urban youth language and could not be ascribed to either of the basic codes involved in the switching or to the processes of lexical manipulation e.g.

- *mafonyifo wananyarwa*: *Prostitutes are being arrested*
- *mboch: house girl (maid)*
- *fala: idiot*
- *karao: Police*
- *mbota: watch*
- *mbwenya: overcoat or trench coat.*
- *mdosi/sonko: boss, rich man*
- *mdosi/mbuyu: also means dad*
- *munde/chapa/niado/ganji/dough: money*
- *nare: fire/matches*
- *murenga/dinga: car*
- *njumu/njuti: shoes*

- *maunenge: hunger*
- *veve/mbachu: khat*

This vocabulary is not simply taken from any known languages surrounding Sheng. They must be described as Sheng vocabulary. Table 1 gives a summary of some Sheng words with their origin and meaning. Such is the kind of lexicon that I intend to label as Sheng.

Table 1: Sheng Lexemes with their origin

Word	**Language of Origin**	**Original Meaning**	**Sheng Meaning**
a) sanya b) ishia c) pewa d) mbuyu e) chorea	Kiswahili	gather finish (isha) be given a type of old tree draw (chora)	steal go away be made drunk father to plan
a) msupa/mpasu b) hepi c) vibe d) blast e) winch	English	super happy vibrations blast wench	Pretty girl fun talk Reprimand (tell off) coins
a) noma b) genya c) banjuka d) sota	Sheng		trouble die dance financially broke
a) nyita b) unbwogable c) ndai d) kuthela	vernacular Kikuyu Luo (buogo) Luhyia Kamba	catch be shaken good	understand cannot be scared car to finish/complete

In order to categorise such data, this study deals with Sheng/Engsh at the vocabulary level by identifying only the conventionalised Sheng/Engsh vocabulary e.g. *masaa* 'with speed/hasty', becks 'money', *niaje* 'how are you?', etc.

In terms of language identification, codeswitched vocabularies do not pose difficulties for the current study because they will be simply classified as that under codeswitching. For example, *nilikushow*,

(16)	*ni*	*li*	*ku*	*show*
	1SG.SBJ(Sw)	-PST-(Sw)	1SN.OBJ (Sw)	told you (Sh)

I brought to your attention.

This word would have been difficult to categorise. It appears as both a codeswitch of Kiswahili prefixes (ni-li-ku) and English verb (show). It also appears as a Sheng or Engsh vocabulary in terms of meaning. The root show is evidently English but it has another meaning of 'told' in Sheng/Engsh. I classify this generally in the category of codeswitched vocabulary. The main difficulty arises at the syntactic level. It is difficult to identify what is codeswitching and what is typically Sheng. Compare a message like:

(17) *niaje bro, sina credo nitatry kuget tubonge moro*

ni	aje	bro	si	na	credo
it is(Sw)	how(Sh)	brother(Eng)	I don't(Sw)	have(Sw)	credit(Eng/Sh)

ni	ta	try	ku	get	tu
1SG.SBJ(Sw)	FT(Sw)	try(Eng)	prep(Sw)	get(Eng)	2PL.SBJ(Sw)

bonge	moro
talk(Sh)	tomorrow(Eng/Sh)

Hi brother, I don't have airtime now but I will try to get so that we talk tomorrow.

This message can be wholly considered as Sheng/Engsh, yet at the same time it can be argued that it is a mere codeswitch between Sheng, English and Kiswahili. At the sentence level, I will categorise it as interword codeswitching involving codeswitching between English, Kiswahili and Sheng/Engsh basing on the language and meaning of each word used. I do not categorise Sheng/Engsh at the sentence or message level due to this ambiguity.

2.1.5. Conclusion

In conclusion, this study confines itself to indigenous vernacular languages as a block, in addition to Kiswahili, English and Sheng. They are the main languages used by my CMC target group. Com-

municators who share a common vernacular language may easily use their language to communicate to each other whereas those that belong to different vernacular language groups may use Kiswahili, English or Sheng to communicate. It is interesting that the functions of English and Kiswahili overlap. Both Webb & Kembo-Sure (2000) and Ferrari (2005) aver that English and Kiswahili enjoy a diglossic relationship with English being used as the language of government business (parliamentary debating and administration), judicial system, commerce and communication, instruction in schools and popular mass media. Kiswahili on the other hand is consigned for more social functions and for communication across ethnic communities and is used as a carrier language in the region. Although this is mainly the case, it is important to point out that the elite use English to communicate to each other in social gatherings. Consequently, my general view is that these languages play similar roles depending on the social class. Sheng on the other hand is mainly used by the youth and young adults. It is very popular in social gatherings and communication among young age-mates. As earlier mentioned Sheng changes very fast and involves a lot of linguistic creativity. This is the reason why I deem it necessary to investigate how it is being utilised in CMC.

To summarise this section, I conclude that in this complex multilingual environment, the average Kenyan has at least 3 languages, that is, a vernacular language, Kiswahili and English. Besides the three languages, typical Kenyan youths and young adults additionally have Sheng in their linguistic repertoire. On this note, it is worthwhile to point out that owing to various reasons like urbanization and intermarriages, it is now becoming common for people to possess a combination of Kiswahili, English and Sheng with no vernacular language.

As noted previously, Sheng changes at a very fast rate, therefore unless one keeps updating it, one's knowledge of Sheng may be out-dated or different from the current Sheng. Generally, Sheng is associated with modernity and urbanity among the youth and young adults who want to 'belong' to the current 'in' group.

Vernacular languages on the other hand are associated with *ushamba* -traditional values, and lack of modernization and/or education. Kiswahili is associated with African urbanism, trade and blue-collar jobs. It dominates social interaction and is the language of national unity. But in comparison to English speakers, Kiswahili speakers are perceived as disadvantaged. English is associated with government service, the professions and high status jobs. It is the language of prestige and upward mobility. English speakers are considered as the young, modern Kenyans, the educated, clerical workers (Parkin 1977, Samper 2002, Whiteley 1974, Abdulaziz & Osinde 1997). Basing on these associations, the trilingual average Kenyan uses all these languages in desirable contexts, for example the mother tongue is used in more personal settings and topics at home with family, relatives etc, then Kiswahili is used in the general public sphere e.g. on the bus, in the market etc. Then English is used for formal transactions e.g. at the office. The youth and young adults additionally use Sheng when interacting with peers, e.g. during lunch break or after work etc.

The data for this study is collected from youths and yuppies. It is believed that this is the population that has embraced the use of new forms of Computer Mediated Communication with zeal. These groups also have two additional advantages for this study because of their possession of Sheng besides English, Kiswahili and a vernacular language, and secondly, their literacy because the data is text-based. They typically use these languages daily in different contexts and for different reasons as described earlier e.g. use of a vernacular language to communicate with the family, Kiswahili for the public and some intimate social settings, English for academic and professional services, and finally Sheng in informal settings with peers.

On the literacy aspect, English is mostly associated with literacy, followed by Kiswahili, a vernacular language and Sheng. However, with the emergence of CMCs whose language use is described as a hybrid between verbal and written language, more people are now using their languages in text CMC just as they would use them

verbally. Similarly, multilingual people are now creatively engaging their different languages to enable them get past some CMC barriers like the limitation of the number of characters that form messages.

2.2. Computer Mediated Communication in Kenya

Notably, although CMC in Kenya is still a rather recent mode of interaction, it has become increasingly popular for social purposes. Communication via the Internet and cell phone in both speech and text forms is now used more and more in people's daily lives for personal and professional purposes. Just like in the rest of Africa, SMS in Kenya is more important and heavily utilised than communication via the Internet and voice calls, due to its affordability and expediency. Limo (2008) captures this by claiming that more and more Kenyans use the short message service (SMS) to communicate more than they use voice calls. Besides SMS, the other most common genres of Computer mediated text communication in the African context are the Electronic mail (Email), Instant messages and SNS, like Facebook and Kenyan daily newspaper comments forum. Video Network Sites like YouTube are also on the rise despite the low and unstable connection speed (Barasa 2007).

It is worth noting that the use of computers is generally a reserve of the 'young'. Most people in the older generation are computer illiterate and lack typing skills. This includes professionals who can easily access computers but have secretaries and typists to do their computer tasks. The youth and yuppies use computers more commonly. Even the ones who cannot access a computer at work will do so at a cyber cafe. It is also interesting that in some cyber cafes, the computer keyboards are in Chinese characters or are generally eroded but this does not deter users from typing.

Rankings compiled by Alexa Web Information Company[19], indicate that the 10 most popular Internet sites visited by Kenyans are ranked as follows:

[19] http://www.alexa.com/topsites/countries/KE

1. Google.co.ke
2. Google.com
3. Yahoo.com
4. Facebook.com
5. Youtube.com
6. Wikipedia.org
7. Blogger.com
8. Live.com
9. Msn.com
10. Nation.co.ke

Instant messaging (IM) which is a synchronous genre is also picking up popularity especially amongst young upcoming middle and upper class professionals in urban centres with unlimited and 'free' access to Internet connection at their work stations. To make these so called genres more entwined, in November 2007, one of the network provider companies (Celtel now renamed Zain) pioneered a new service where a chat message typed on a computer can be received on the cell phone like an SMS, and vice versa just like in an instant message (IM) genre albeit slower. This communication is also swiftly picking up popularity because of its convenience such that one can easily chat online and access Emails without necessarily having a computer in the vicinity.

In addition, as mentioned, there are now smartphones which are being manufactured and work just like a mini or pocket computer. They enable one to surf the Internet, access and respond to Emails and do Instant Messaging from anywhere. Mobile phones have become extremely useful in Kenya and are even used to transact money (M-Pesa). Ingenuity is making people use the phone in many ways that they probably did not buy it for in the first place.

The ensuing section gives a brief description of each of the CMC providers in Kenya.

2.2.1. Mobile Phones and Networks

Mobile phones are designed to meet Western needs. Yet, subscribers in developing countries now represent the majority of mobile phone users worldwide. Africa is currently the fastest growing mobile phone market in the world with approximately 300 million mobile phone subscribers.

The projection by statistics from the African Mobile Factobook[20] in graph 1 indicates that mobile phone subscription in Africa will increase up to close to 600 million by the year 2012. It is also projected that 20% of the subscribers will move on to use 3G[21] mobile phones. At least 15 service provider companies have announced plans of introducing 3G voice and data services (including among others, Kenya, Nigeria and Tanzania).

According to the Information Society Statistical profiles 2009, the distribution of the leading mobile phone subscription in Africa in the year 2008 is shown in graph 2 Nigeria and South Africa have the most subscribers. Next is Kenya despite the fact that East Africans pay taxes of between 25% and 30% on mobile phone services, compared with an average of 17% across Africa (ITU 2010). The majority of this subscription is on pre-paid terms. The growth of Kenya's mobile phone industry has grown rapidly from the year 1999 to 2009. In June of 1999, Kenya had 15,000 mobile phone subscribers. By the end of 2004 the country had 3.4 million subscribers and by 2006, the number had grown to over 5.6 million. This tripled to 17.6 million by mid 2009. This growth has been encouraged by the introduction of cheaper mobile phones coupled with the reduction of calling rates, the spreading of network service to rural areas, and the general competition by service providers.

It is projected that by the end of 2010 there will be around 25 million subscribers in Kenya as shown in
graph 3. This will be driven by the new services introduced by the service providers for example, access to Internet and money transfer services.

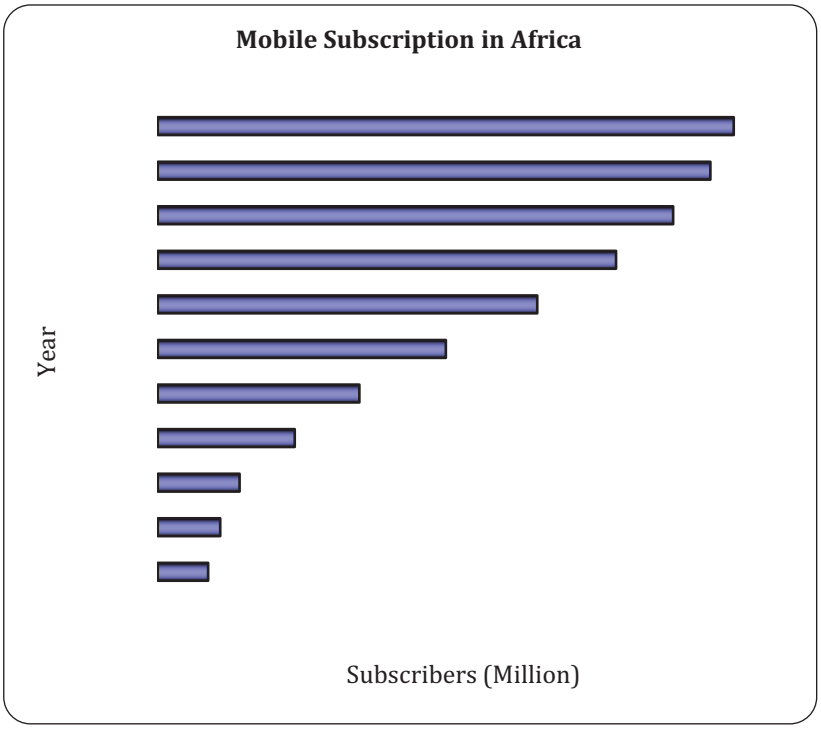
It is claimed in the African Mobile Factobook that across most of Africa, SMS is likely to be the only non-voice value-added service to gain mass-market popularity in the immediate future owing to its affordable price.

---

[20] Source: www.africantelecomsnews.com

[21] 3rd Generation Phones.

Graph 1: Mobile Subscription in Africa (2002-2012)[22]



Currently, in addition to communicating with friends and family, SMS is being used in Kenya in innovative ways such as pricing information for agricultural products, mobile banking and human rights abuse notifications. This clearly shows how popular this genre of communication is.

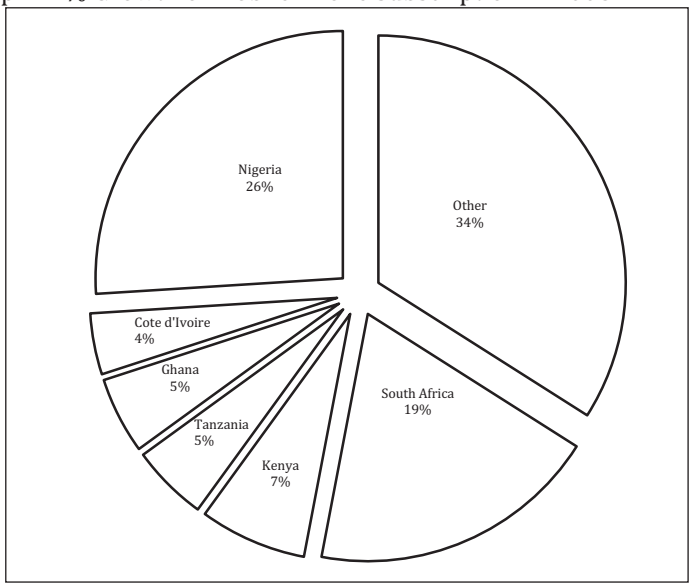
With the relatively high number of subscribers, it is then important to research into the language use in order to come up with a detailed description and informed conclusions.

Despite this need, to date, most research literature on CMC has focused almost exclusively on emergent practices in English, European and Asian countries neglecting developments within

[22] Data compiled from the Africa Mobile Factobook

other populations.

Graph 2: % Growth of Mobile Phone Subscription in 2008 in Africa[23]



The current mobile service providers in Kenya are Safaricom (backed by Vodafone), Zain Kenya, Orange/Telkom Kenya (backed by France Telecom) and the most recent Yu Econet wireless (backed by Essar global). I will present each of the service providers hereafter.

2.2.2. Safaricom

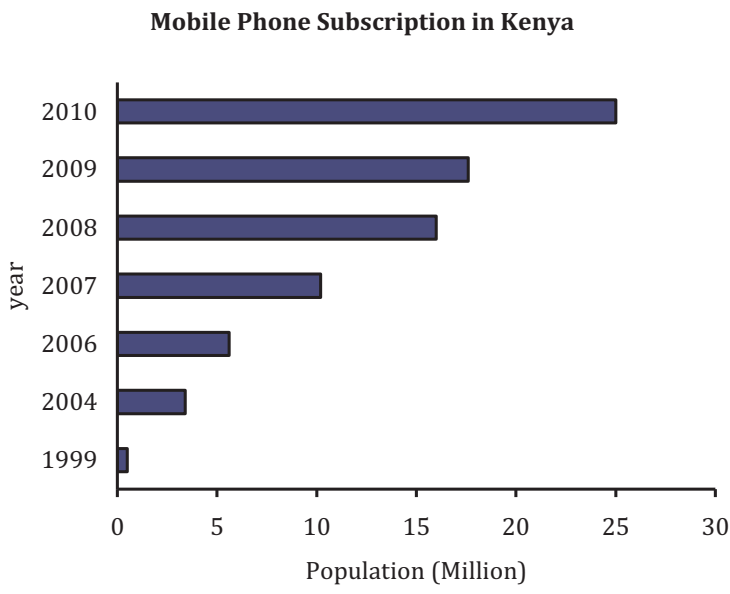
Popularly known as 'Safcom', Safaricom was introduced into the Kenyan market in 2000. It came in with a very aggressive marketing strategy and a lot of tariffs e.g. *Jambo, Taifa, Sema, Jibambie, Ongea, and Rudi tariffs,* which act as incentives.

It has a large number of active subscribers and is most of the time giving incentives thus encouraging more subscribers. The earliest tariff used to allow each subscriber to send three free *please call*

[23] Statistics are extracted from the Information Society Statistical profiles 2009-Africa. http://www.itu.int/dms_pub/itu-d/opb/ind/D-IND-RPM.AF-2009-PDF-E.pdf

me SMS per day, to cater for the needs of those who needed to communicate but had no airtime on their phones.

Graph 3: Growth of Mobile Phone Subscription in Kenya



Kenyans made use of this service and even made it a coded message used to assure their contacts that they were fine and just wanted to say hi, goodnight or just to keep in touch. Another popular practical tariff is the M-Pesa money transfer service aimed at mobile subscribers who do not have a bank account, typically because they do not have access to a bank or because they do not have sufficient income to justify a bank account. All these incentives have encouraged users to subscribe to the network thus increasing the number of the users' population.

2.2.3. Zain Kenya

Zain Kenya started off as Kencell in 2000, then changed to Celtel in 2004 but later rebranded to Zain Kenya in 2008. Its current popularity is due to the *vuka tariff* which allows its subscribers (prepaid and post-paid) in Africa and the Middle East using the same

Network to enjoy the benefits of being treated as a local customer wherever they are. They can make calls and send messages at local rates when communicating with a travelling Zain subscriber who then receives incoming calls free-of-charge and is able to make calls back home at local rates. Pre-paid subscribers are also enabled to top up their airtime with recharge cards bought from either their home country or in one of the countries with the network. The home based network service is automatically activated upon crossing the geographic border into one of the countries, with no prior registration required or sign-up fee. Other popular Zain tariffs include unlimited call (*Jiachilie*) and (*Club 20*) SMS tariffs.

2.2.4. Orange Mobile

The third mobile network launched in Kenya was Orange mobile in 2008. It is the first to offer GSM services from fixed-line incumbent Telkom Kenya. The launch of the Orange mobile service by Telkom Kenya followed France Telecom taking a controlling stake in the Kenyan company in 2007. The company has also launched its GSM network enabling it offer landline and Internet access as well as mobile services. Initially, broadband Internet and mobile offers were available in Nairobi and Mombasa only, but have by now progressively extended across the whole country. Undersea cables have also been launched to boost Internet access capacity. Orange is now the commercial brand for the mobile, broadband and fixed wireless services from Telkom Kenya.

2.2.5. YU

Yu is the brand name that the Econet wireless mobile network operator has taken. This provider entered the market as the fourth service provider in December 2008. It started with impressive low calling and SMS tariffs. So far its major incentive has been the offer to pay all subscribers 70% per minute for receiving calls from the other Kenyan mobile networks.

As clearly seen, Kenyans are benefiting from the mobile phone network companies trying to outdo each other by announcing cheaper and friendlier tariffs than the other. A majority of Kenyan

mobile phone subscribers have formed the habit of subscribing to several network providers at a time. The regular line is usually influenced by which operator friends, relatives and other contacts use because it is cheaper to call within the same network. The other providers are used depending on the tariff on offer and the cheapest possible option.

2.2.6. Internet

Compared to the rest of the world, Internet is still behind in its penetration into Africa. This could be because of the unstable supply of electricity coupled with the high cost of computers and the high rate of illiteracy. But the situation is changing with the evolution of the technology savvy teenagers and yuppies. Graph 4, graph 5 and graph 5 from the Internet World Stats usage and population statistics[24] clearly show this state of affairs.

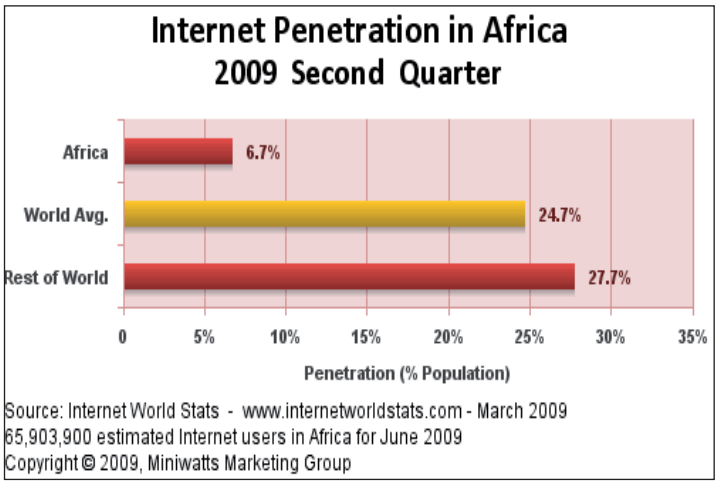
Internet first became available in Kenya in 1993. This was affordable to a very small group of technological enthusiasts and they accessed the Internet through a service known as Gopher which offered access to text-based information. This access was mainly through international leased lines. Mweu (2000) reveals that The African Regional Centre for Computing (ARCC), an NGO based in Nairobi, Kenya, was the first web-based Internet service by provider.

The connection to the global Internet backbone was via an analogue leased line. The first commercial Internet Service Provider (ISP), Formnet began operating in 1995. Soon after, it was joined by the entry of three other ISPs.

All the ISPs leased analogue or digital data lines from Kenya to the US to access the Internet backbone. Soon after, the number of ISPs grew and so did the pressure for bandwidth.

[24] http://www.internetworldstats.com/stats1.htm

Graph 4: Percentage of Internet Penetration in Africa



Mweu (2000) further explains that at this point the Kenya Posts & Telecommunications Corporation (KPTC) realised the need for an Internet access backbone which would make access to the Internet for ISPs cheaper, because of the local access.

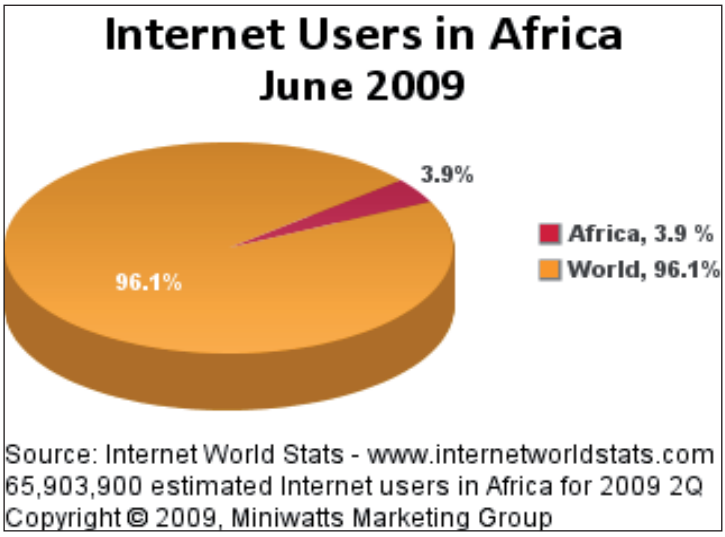
The backbone, EAFIX, was launched in December 1998, and together with it Jambonet, an access service for ISPs. Jambonet reduced the ISP cost to a quarter of what was initially paid internationally. This move paved way for the entry of more ISPs and the rise of market competition.

In July 1999 the Kenyan government officially liberalised the telecommunications market and formed the Communication Commission of Kenya (CCK) to regulate the sector. The CCK nominated Telkom Kenya which was formed from the telecommunications arm of the former KPTC and allowed it a monopoly to operate an Internet backbone for five years. Additionally, ISPs were officially acknowledged and authorised to operate after obtaining a licence from CCK. With the freedom of operation the number of ISPs increased to 50 by 2001.

Limo (2008) expounds that in 2007, a draft ICT bill was pre-

sented to the Kenyan Parliament presenting proposals on the establishment of universal access fund (UAC) in order to boost ICT exposure into the country.

Graph 5: Percentage of Internet Users in Africa



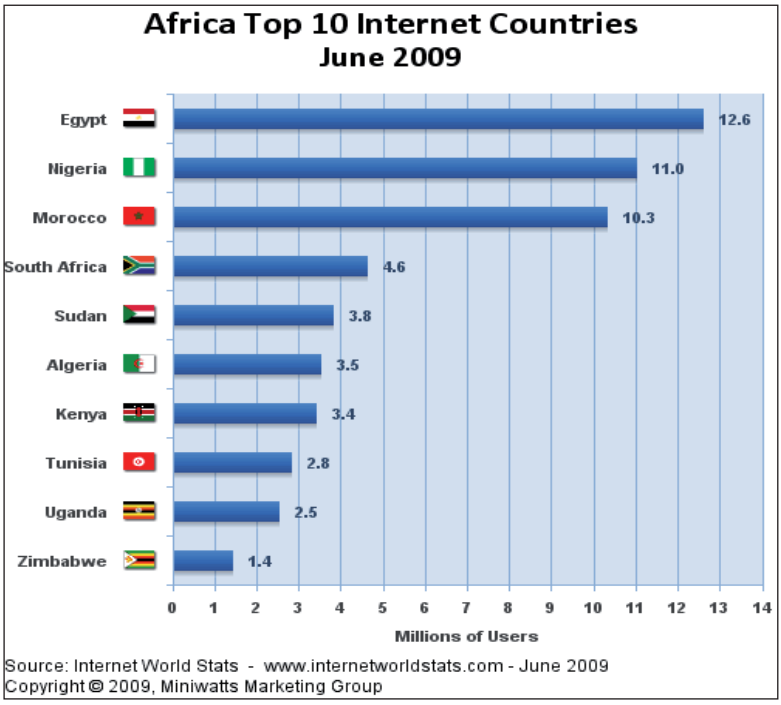
Interested parties in the sector like mobile phone firms, digital infrastructure developers, development partners and the Government would contribute to the fund, which is to be administered by Communication Commission of Kenya (CCK).

Even before the establishment of this fund, CCK came up with a good initiative in the name of a universal access plan of putting up community telecentres and school-based ICT training centres in the rural areas.

There are four such centres in Koibatek, Kitui, Makueni and Bungoma where one may go for the Internet, ICT training, typing, printing, etc. In addition, the ICT Board is setting up digital villages in every constituency as pilot projects to showcase the use of ICT. Although these access points are few and negligible on the national scale, they help in stimulating interest in ICT and are therefore an important step towards making Kenya an informa-

tion society.

Graph 6: List of the Top 10 African Countries in the Use of Internet

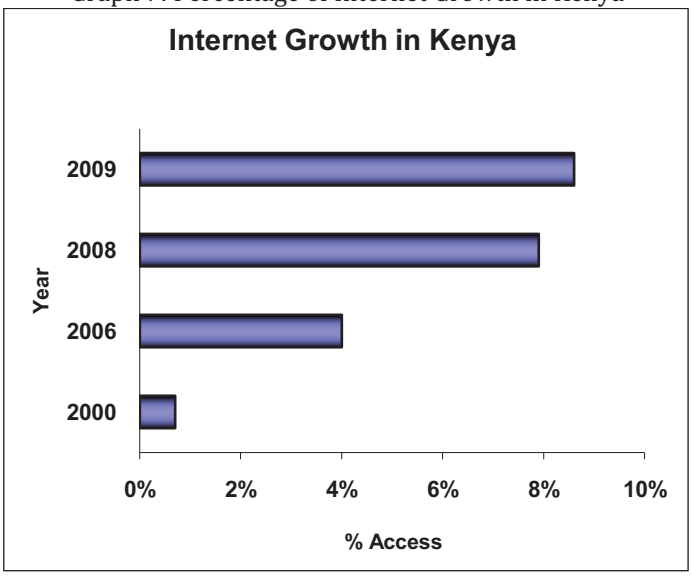


The cost of Internet connectivity is now more affordable after the completion of the landing of The East African Marine Systems (TEAMS) undersea to connect Kenya to the world through United Arab Emirates. The affordability and convenience has attracted more users.

A Kenyan survey by the market research company, Synovate, which was published online[25] in the Balancing Act News Update e-letter on 29th January 2010, shows that Kenya's Internet market is growing fast and currently has over 3.5 million users.

[25]http://www.balancingact-africa.com/news/back/balancing-act_489.html

Graph 7: Percentage of Internet Growth in Kenya



This growth in users is from both urban and rural areas and is predominantly amongst the young and well educated. It is claimed in the report that

> "In terms of age, 50 per cent of the respondents using Internet were aged 15-34 with 21 per cent in the 18-24 age bracket. The upcoming generation of Kenyans will be regular users of the Internet and it will form as much part of their lives as mobile phones. Over 56 per cent of the Internet using respondents were college or university educated. Therefore, those countries with better education levels in Africa will show markedly higher Internet penetration levels." (Synovate Report 29th January 2010).

According to the report, the five top uses of the Internet range between 40-50 per cent of the sample users are entertainment, games and music, social networking and instant messaging, Emails, general surfing, and job search.

The report concludes that the significance of its findings is that

Kenya is both one of Africa's largest mid-scale markets and is a guide in terms of technology adoption for other regions. The display of this is summarised in graph 7. This achievement of Kenya's Internet market is made possible by the availability of Internet services in professional institutions and companies coupled with the availability of Internet cafes in urban centres.

This is despite the regular setbacks such as Internet connection problems, power-cuts, variant keyboards arrangements, illegible keyboard symbols and the relatively high prices charged.

Chapter 3. Methodology

This chapter begins by discussing the general population group involved in the research, the scope and limitations of the study, the ethical considerations involved, and then finally, the methods employed in data collection for each CMC genre.

3.1. Population Group

The participants for this study were selected from Kenyan Universities, middle level colleges and Young Urban Professionals (yuppies). They were all required to be Kenyan and under 36 years. This group is considered to be representative of the average CMC user in Kenya. It was deemed that both groups, that is the students in higher education and the yuppies have free or affordable access to computers and Internet networks which they utilise at their institutions and in their work places at subsidised costs or free of charge. These groups also own cell-phones and find it cheap and convenient to send SMS for communication with their peers. In addition, these groups are eager to communicate using not only new CMC devices but also use 'new' language with their peers. Bodomo (2009:301) claims that people, especially the young, are far more rebellious in breaking the rules of standard writing, enjoying the freedom of speech with flexibility to express themselves. The group is at ease with reading and writing. They have a good command of at least two languages: English and Kiswahili. However, for the purposes of this research I selected those who have a command of at least three languages, a vernacular language, Kiswahili and English in order to capture the use of codeswitching data and creativity of multilingual individuals in the use of text CMC. The choice of my population group was motivated by the facts that

i. they are cross societal and cut across the different parts of the country
ii. this group is innovative in the use of CMC
iii. the group is accessible and cooperative in sharing data
iv. the group can access and use CMC devices with relative ease.

Invitations to participate in the research as informants were advertised in 5 universities, that is Masinde Muliro University of Science and Technology (MMUST), Egerton University, University of Nairobi (UON), University of Eastern Africa Baraton and Mt. Kenya University, and in 3 middle level colleges, Kisumu Polytechnic, Machakos Institute of Technology and Mombasa Polytechnic. These institutions sample 7 out of the 8 provinces in Kenya. A total of 104 students responded but 16 dropped out for fear of privacy violations. The remaining 88 students accepted but I excluded 3 who were not Kenyans and one who was 42 years old. It was agreed that the students would be paid for each message sent, that is 10 Ksh for each SMS, 20 Ksh for each Email, and 30Ksh for each chat sent. [26] This was to reimburse the incurred costs as well as to act as a motivator for their participation.

For yuppies, I asked friends and acquaintances to act as my informants, together with their own friends. 37 yuppies volunteered. These were all trained professionals working in companies and other institutions in urban areas.

It was hoped that the selected population group of 121 individuals fairly represented the intended population of the youth and yuppie Kenyans.

3.2. Scope and Limitations

The research only dealt with youthful informants. They were all educated. Those who were not yuppies or in educational institutions were not represented. Additionally, older people who use CMC for communication were not represented in the selection. Even so, this does not influence the results substantially because the general group that uses CMC in Kenya is represented by the selected group of informants.

It was also noted that the research mostly dealt with communication between peers. This means that the communicated messages were meant for recipients who somehow belong to the

[26] 1 Kenyan shilling (KSH) is equivalent to approximately 0.02 USD, 0.01 EUR and 0.01 GBP.

same group; SMS were meant for friends or family, Emails were also meant for friends and acquaintances, IMs were between close friends and SNS comments were for people with similar interests. This close relationship between the message senders and receivers provided a balanced setting to access informal messages from each genre.

The choice of the participants and the informal types of messages sent were therefore considered as the main limitations in the data collection.

3.3. Ethical Considerations

Ethical issues concerning the collection and handling of data were considered. Each of the participants of three genres SMS, Email, and IM was asked to give consent in order to participate in the research. They consented but were reluctant to have their names appear in their messages. They were assured that all names and any other identifying information would be removed and replaced with conforming pseudonyms before appearing in the database in order to protect their privacy. For example, after analysis, a message like

> *hi john's no. is 012345*

would be changed to read

> *hi jack's no. is 352104*

In addition, for Email and SMS, the participants were required to only forward to me the messages that they had sent, and not those that they had received unless they got consent to do so from the senders. For IM, I only used data from the participants who consented and forwarded their recorded chats to me, and also from the participants who chatted with me. It is important to note that although all these participants consented to have their messages used in the research, they were not informed on the kind of data that I was interested in. This helped to control the communication by keeping it raw and natural. For SNS, I collected comments which had been left by visitors to the sites. These comments are

already publicly available and the contributors had coined up pseudonyms so their identities were already protected. By concealing their identities through the use of these pseudonyms, users use language freely. In fact some users take advantage to misuse this opportunity to use obscene and abusive language in SNS.

3.4. Data Collection

As already pointed out, data for the research was collected in the form of the four genres of CMC:

- Electronic Mail (Email)
- Short Text Messaging (SMS)
- Instant Messaging (IM) or Chat
- Social Network Sites (SNS) and Video Network Sites (VNS) e.g. Facebook, YouTube, Kenyan Daily newspapers comments (The Daily Nation and The Standard newspapers), Kenyan discussion forums e.g. Mashada

The main equipment used were a computer with a reliable Internet connection, an iPhone and a USB port to enable the downloading of SMS. The iPhone was used owing to its huge storage capacity and ports. It was also practical due to its capacity to access the Internet, Emails, chat forums, YouTube and other Social networks.

The study sought for a large number of texts in order to realise the corpus and draw informed conclusions. Note that in some cases student participants were requested to send more messages than the yuppies. This is because the students were motivated by the pay while yuppies did it on a voluntary basis. A total of 5427 messages were collected cf. 3.7.

3.4.1. Emails

To acquire Email data, I asked each of the student participants to forward to me a minimum of 4 personal Email messages from their Email outboxes on any two occasions. *Personal* is used here to mean informal messages directed to friends unlike official mes-

sages that are composed and edited consciously. In order to reduce the issue of filtering of messages by the participants, they were asked to forward four of the last five personal Emails that they had sent. After a minimum period of eight weeks they were asked to send another four of the last five Emails in their outbox. The yuppies were also asked to forward me at least 2 personal Emails. Some participants forwarded more messages than the required minimum and I collected a total of 780 Email messages this way.

3.4.2. SMS

For SMS data, the students were asked to send copies of their SMS from the sent items folder for a period of 10 weeks. Each participant was asked to forward to me a minimum of 3 random SMS messages per week. The yuppies were asked to send at least 2 SMS. Some participants forwarded more messages. I collected a total of 2730 SMS messages this way.

3.4.3. IM

I asked the students to chat with each other through Instant Messaging during holidays and forward the chats to me for the research. It was thought that chats during holidays would be more realistic and practical since the students would be away from one another. The participants were funded to be chatting with each other at least once per week and 2 times a month with me. For chats, the yuppies were assured of their privacy and requested to forward as many chats as they could for the research. They availed 186 chats while the students provided 11 chats. Seemingly, chat is not a popular communication genre among students yet, perhaps due to the fact that it is expensive coupled with the fact that while on holiday, some return to areas where the necessary equipment and services are not available. I collected a total of 197 IM.

3.4.4. Social Network Sites (SNS)

I browsed and extracted texts from different public Kenyan online social forums provided their policies allowed it like on Facebook discussion forums, comments in reaction to the online Kenyan

Newspapers articles from the Daily Nation and the East African Standard and comments from YouTube which is often referred to as a Video Network Site (VNS). I extracted the data from pages with Kenyan videos. It was deemed likely that most of these informants were Kenyan yuppies. I used the assumption that for users to log on and contribute to the forums means that they are in position to access a computer, go online and are able to type- they therefore fit in the Yuppie category of informants. By use of SNS forums, both students and yuppies are more comfortable and come out openly with their views and in very relaxed natural language mainly because they use pseudonyms that are untraceable. Such forums are an arena to openly react to issues. I collected a total of 1720 SNS comments this way.

3.5. Variables

Having collected the corpus of data, I formulated a number of variables to enable the identification and classification of interesting features in the data. I organised the variables into 2 main themes relating to the technical and social affordances of CMC. This organization is inspired by Herring (2007) who realised these subdivisions in her faceted classification of CMC. She projected that

> "[...] the goal of the CMD (Computer Mediated Discourse) scheme is to articulate aspects of context – both technical and social – [...]. [...] However, as awareness of CMC spread with the popularization of the Internet, it soon became apparent that computer-mediated discourse was sensitive to a variety of technical and situational factors, making it complex and variable [...]." Herring (2007:761)

The language use in all the four genres is influenced by both social and technical affordances. The data reflects this. All the same, I must point out that not every variable can be attributed to one of the themes to the exclusion of the other. In fact some of the data discussions overlap within and among the themes for example the variable of phonological spelling under CMC as technologically motivated features overlaps within its theme with both pronoun-

ceable letters symbols and pronounceable numericals. Each of the overlapping subjects has been discussed under a theme and title that describes it best.

It is also worth mentioning that I used the *eDatax.mdb* program for the automatic counts for the variable's quantitative data. This program was developed by Mr. Maarten Hijzelendoorn who is a computational linguistics expert at Leiden University. More on this is in section 3.6.

The next section presents each of the variables by first defining it, followed by an explanation of its relevance in relation to the hypotheses in section 1.5. This is followed by a prediction of the expected findings. Finally an explanation of how the variable was tagged either manually or automatically by means of a specially made program that is explained in section 3.6. In some cases, an explanation of the arising complications and solutions is provided. In summary, the variables discussion is ordered as follows: a) definition, b) relevance, c) expected findings, d) tag and e) challenges. Data illustrations are also included.

3.5.1. Phonological Spelling

I defined the phonological spelling variable as instances in which words are presented in a way which more closely resembles pronunciation. For example,

(18) ***wen wil*** *we enda to that* ***ples****?* ***tel mi*** *so tht i* ***pripea ali***
Sw (go)
When will we go to that **place**? **Tell me** so that I **prepare early**.

I identified and tagged cases of this variable manually and entered them into the database for the total counts. I based the variable on the two principles:

- least effort: the sender types English words manifested as he thinks of them and does not bother to spell them in the standard way.
- mode limitation: phonological spelling may be used as a way of reducing the total characters of the message.

The expected findings are that all genres will register a presence of phonological spelling caused by the principle of least effort. Regarding mode limitation, the expectation is that SMS and SNS which are the genres with maximum character limitation will have a higher presence of the phonological spelling.

3.5.2. Pronounceable Letters

I defined pronounceable letters as the use of single letters to fill in for words based on the resemblance in pronunciation of this letter as a separate letter of the alphabet and the pronunciation of the word or part of it. An example is the use of *u* for *you*. These are a common occurrence in CMC in relation to the following principles:

- least effort: it takes less effort to type a single letter.
- rapid communication: it is quicker to type a single letter than the whole word.
- mode limitation: less space is required.

The expected results are that all the genres will register high counts of pronounceable letters because of the least effort principle. IM is expected to take the lead because of its rapidity. SMS and SNS are expected to be next because of their character limitation.

This variable was included to enable me to gain more insight into the usage of single letters to fill in for words in Kenyan CMC. The single letters were tagged automatically by the program (cf. 3.6.) as single letters that had spaces on both sides. An example of the use of this variable in a message is:

> **y r u** so quiet
> **Why are you** so quiet.

Other examples of letters symbols and the words they represent include;

> *b* [be, bee]
>
> *c* [see, sea, si]
>
> *n* [and]

r [are]

u [you]

m [I'm, I am]

x [ex]

y [why]

In order to avoid impeding the counts, the indefinite article *a* and the first person pronoun *I* were excluded because they occurred frequently and validly, not in place of words or word-segments. There were also some rare unclear cases which were tagged manually for example, the use of a combination of two or more adjacent pronounceable letters, without any spaces separating them. For example:

ru - are you
nu - and you?

Other infrequent cases were for instance where the single letter was used validly, for example:

(19) *she got a **b+** in de xam*
She scored a B+ in the exam

Such cases were tagged manually and excluded from the totals.

3.5.3. Pronounceable Numericals

I defined numericals as numbers that could be pronounced as words or parts of words and were used to replace the words/part of words. Similar to the pronounceable letters, this variable was also linked to the least effort, rapid communication and mode limitation principles with comparable expected findings (cf. 3.5.2.). This variable was included to find out how numerals are used in Kenyan CMC and what the most frequently used numerals are. Similar to Hård af Segerstad (2002), Frehner (2008) and Bodomo (2009) results, it is expected that number 2 and 4 are the most used.

Examples of use of pronounceable numbers are where numbers are used to represent their sounds for example,

*bye **4** now*
Bye **for** now.

*I'll tell you **2**morrow*
I'll tell you **to**morrow

Other examples of numbers with the words they represent are

1 [one, moja]
2 [to, tu, too]
4 [for, fo]

All numerals used in the message were automatically tagged by the program. The main limitation here was that some numerals were used validly for example in (23).

***6.30** Mbele Ya Lib ni poa. c u (6.30 pm at the front of the library is fine)*
or
***30** bob (30 shillings)*

In these examples, *6.30* and *30* are used in a valid way and ideally should not be counted as numerical cases. I therefore identified these cases manually and subtracted them from the computed totals.

3.5.4. Abbreviations

An abbreviation is defined as the ellipsis or shortening of a word or phrase by clipping or omitting parts of it. As a variable, it also reflects on the principles of rapid communication, least effort and mode limitation (cf. 3.5.2). This variable was included in order to observe the kinds of shortening patterns that exist in Kenyan CMC. All general cases of lexical compression are mutually inclusive in clipping and were manually tagged in this category. Examples include:

afte – prototypical final clipping for *afternoon*

uni – prototypical final clipping for *university*
ur- initial clipping for *your*
hse - medial clipping for *house*
mesg - mixed clipping for *message*

The expectation is that all genres will register a similar count of abbreviations because of the least effort principle. SMS is expected to have the highest count because of its additional feature of character limitation. IM should be ranked second because of its need for rapidity. Email is expected to have a lower count due to its more formal nature and lack of space limitations. The ranking of SNS is expected to be higher than Email because of its mode limitation.

3.5.5. Acronyms

An acronym is defined as a word formed from the initial letters of a name, or by combining initial letters or parts phrases. I set it as a variable in relation to all the principles i.e. rapid communication, least effort, mode limitation (cf. 3.5.2.) and informal communication where a given group is expected to understand the acronym in use.

I tagged this variable manually. An example of the use of an acronym is

(20) ***OMG****! i was kisd by joni's rmate!!*
Oh **M**y **G**od I was kissed by Joni's roommate.

The expected results are that all the genres will register an average count of acronyms because of the least effort principle. IM is then expected to have the highest count because of its need for rapidity followed by SMS because of its additional feature of character limitation. Email is expected to have a lower count due to its formal nature. SNS is likely to fall in the middle but with an additional feature of peer communication where peers use a lot of creativity in order to outdo others.

3.5.6. Exclusive Consonants

I regarded as exclusive consonants as all the words that only contained consonant letters. This variable was also linked to the rapid communication, least effort and mode limitation principles (cf. 3.5.2). It was included in order to find out the frequency and pattern of word abbreviation and possible reasons explaining it in relation to each CMC mode. Cases of this variable were automatically tagged by the program.

The letter *y* proved to be a challenge because it could not be considered as a clear consonant. For example; if it was listed as a consonant, then some words like *why, shy, by, my, cry, fry, dry, try,* etc. would mistakenly be registered as abbreviations although they are not. I hence did a manual check for occurrences of *y* and added them to the total count.

Examples of words tagged as abbreviations include

(21) *hi* ***Hd*** *a gud dy? Mine ok. Scand tha fotos* ***bt*** *send 2mr m bila airtym gdnyt*

hi had a good day? mine ok. Scanned the photos but send tomorrow Im bila airtime goodnight. (*bila* is Kiswahili for *without*)

Hi, did you have a good day? Mine was okay. I scanned the photos but I will send them tomorrow because I don't have airtime. Goodnight.

The expected results are that all the genres will register instances of exclusive consonant use because of the least effort principle. SMS is expected to have the highest count because of its character limitation and peer-to-peer style. This may be followed by IM owing to the peer to peer communication and creativity. Email may have the least because of their more formal nature.

3.5.7. Contractions

I have adopted Bieswanger's (2007) definition of contractions as combinations of two words that lead to a smaller number of characters than the spelling of the two words individually. In English, contractions are represented by an apostrophe to replace the omitted letters. Contractions usually consist of a pronoun followed by a form of the verb for example, *I'l* for *I will*. It is realised

that the use of contractions is not unique to CMC but nevertheless I included it as a variable mainly to observe,

- if it is used numerously since the CMC genres are mainly informal;
- if it is used with the contraction apostrophe or not;
- if there are new creative ways of combining words in a similar way to contractions.

The expected results are that there will be a high count of contractions for all the genres due to the least effort principle. Additionally, I expect that the majority of these contractions will not make use of the apostrophe to mark the contraction point. It is also possible that there are novel kinds of contractions.

3.5.8. Misspelling and Typographic Errors

I defined the variable 'misspelling and typographical errors' as the accidental use of non-standard spelling. It is related to the rapid communication and least effort principle. This variable was important in finding out which CMC genre registers the most spelling errors and the possible causes and the impact of this on the communication. Another level of misspellings identified is the recurrence of variants of English whereby some texts were written the way they are pronounced due to the influence of the vernacular language for example,

(22) *Tafadha**ri priss**...*
Tafadha**li** p**l**ease... (*tafadhali* is Kiswahili for *please*)

This message is a variant which was most likely written by a Kikuyu speaker since the language does not make use of the lateral consonant. Occurrences of misspelling were tagged manually as the program did not contain dictionaries of any of the languages involved. It was also very tricky to distinguish between words that had been misspelt and those that had been deliberately written unconventionally like clippings. For example, a word like *listen* could have been deliberately written as *listn*. In order to minimise this, I looked for obvious misspellings like where a letter was repeated e.g. *assiggnment [assignment]*, where letters were interchanged e.g. *langauge [language]*, or where an additional letter

was used e.g. *hostekl [hostel]*. It was easier in chats because in many instances, the communicators made corrections after typos. It is expected that IM would have the most errors because of the rapidity involved in typing. SMS is expected to have the least because it is composed and typed more attentively. Emails and SNS are expected to be average.

3.5.9. Capitalisation

The variable of capitalisation in CMC was observed at two levels:

- excess capitalisation
- missing capitalisation

I defined excess capitalisation as any letter that occurred adjacent to another capital. The first capital was overlooked by the program because it was likely to be valid. Relative to the mode limitation principle, this variable was included in order to observe the creativity in the use of excess capitalisation. Excess capitalisation was tagged by the program. For example, in the word *PLEEEEz* the excess capitals were given a count of five. The main limitation of this was that in some cases, the counts incorporated valid capitalisation for example in abbreviations such as *NGO (Non-Governmental Organization), AOL (American (Africa) on Line)* etc. In contrast, excess capitalisation for example M, Y, L, B and T in the following examples

(23) *6.30* ***M****bele* ***Y****a* ***L****ib ni poa. c u*
It is fine (to meet) in front of the library at 6.30.

(24) ***M****ay****B****e,*

(25) *wha****T***

were not counted since they were not adjacent to each other, and it would prove difficult to identify them automatically in isolation and avoid valid initial capitals at the same time. However, there were not many such occurrences of the described cases, nevertheless, I tagged such manually by either addition or subtraction from the main counts as was required. I also made note of the number

of messages that contained excessive and exclusive capitalisation. The expected result for excess capitalisation is that it is used creatively in all genres and mostly in IM as a way around the technological deficiency in the presentation of on-going feelings.

Absent Capitalisation was defined as the missing of the initial capital letter at the beginning of each sentence. This variable was associated with the least effort principle. It is important in the observation of the deviation that CMC has from standard writing. Absent capitalisation was tagged automatically at the sentence level, for example:

(26) ***m****y God, wat wil apen. I hv felt lyk kulia.* ***n****iko salon nakam*
My God, What will happen? I have felt like crying. I am in the salon but I am coming. (*kulia* is Kiswahili for *to cry*, *niko* for *I am in*, *na*-kam for *I am* –coming)

In this case the program would indicate that the message above had 2 missing capital letters in the highlighted letters. The main challenge was that the program could not identify absent capitalisation at the word level like on proper nouns for example in:

(27) *Bring us ngwaci from* ***n****yeri*
Bring us sweet potatoes from **N**yeri (*Ngwaci* is Kikuyu for *sweet potatoes*)

It would be counted that there is no missing capitalisation whereas in fact *nyeri* is a proper noun and should be *Nyeri.* It was not possible for the program to identify such cases automatically. I therefore restricted the variable to the sentence level. The expected results are that all genres will not make much use of the initial capital letter because of the least effort principle.

3.5.10. Punctuation

I defined punctuations as standard signs set to regulate texts. It was defined at 3 levels:

- The use of punctuation(s) in each CMC message in relation to its use in standard written language.
- Missing termination marks: any sentence without a closing punctuation.

- Excess punctuations: any occurrence of more than one adjacent punctuation mark.

The first level was associated with the informal nature of CMC texts in comparison to standard texts. Missing termination marks were mainly related to the principle of least effort. Occurrence of excess punctuations was linked to mode limitation which leads users to resort to the use of punctuations to express themselves.

The punctuations tagged automatically were from among the following

- hyphens
! exclamation marks
? question marks
... ellipsis
. full stop
, comma
' apostrophe
"" quotation marks
: colon
other (any other that was not listed)

For example:

(28) *hi sema, uko je? mi poa, just chckin u. baadayez*
Hi, how are you? I am fine, just checking on you. See you later

would be counted as 2 commas, 1 question mark and 1 full stop. This has a total of four punctuations and the comma is the most frequently used form of punctuation for this message. For excess punctuation, the first punctuation was deemed as valid but any other adjacent ones were regarded to be excess. This excess punctuation(s) could be identical or varied as the following examples indicate:

(29) *wow!!!!!* (4 identical excess punctuations)

(30) *hows the going so far??!.* (3 varied excess punctuations)

Note that both the general occurrence of the apostrophe and its absence in contractions and possession marker are counted. Similarly, the exclamation mark, the question mark, the ellipsis mark and the full stop are dealt with as part of termination marks expected to occur at the end of the message. These are tagged manually.

The expected results are that because of the mode limitation principle, the use of punctuations in CMC is different from their use in standard written language. They are likely to be used in more ways than only regulating texts. There would be many messages without the final termination punctuations especially in IM which is a form of continuous conversation and SMS because of the least effort and mode limitation principles.

3.5.11. Graphics: Smileys and Emoticons

This variable is defined as the occurrence of graphics to communicate feelings and emotions. It reflects an aspect of the mode limitation principle since it is difficult to express emotions through text. This variable was set to find out more about the general use of Smileys and Emoticons in CMC.

The Microsoft Access program could only read the ASCII characters and only recognised basic Smileys composed of these characters. It was unable to read graphics and pictures used for Smileys and Emoticons. Therefore I composed a list of the common recurrent Emoticons and created a code for each, such that whenever the program came across the code, it computed it as the Smiley being represented.

It was expected that IM would register the highest number of graphics because of the underlying assumption that it is the closest to verbal communication. This would be followed by SNS, then Email and finally SMS which is the most brief.

The composed lists of Emoticons and Smileys with their codes are shown in table 2 and table 3.

Table 2: List of Recurring Emoticons and their Codes

Emoticon Picture	Meaning	Code
	Smug face	[Smug]
	busy working	[busy]
	Agreement[27]	[Yes]
	Disagreement[28]	[No]
	Sleepy[29]	[Slp]
	Embarrassed	[Emb]
	Angel Face	[Angl]
	Evil Smile	[Evl]
	Nervous (sweating)	[Nrv]
	Playful	[Play]
	Love	[Lv]
	Ill	[Ill]
	Attention -Pointing /wagging finger	[Att]
	Tears	[Cry]
	Thinking	[Thnk]
	Cool	[Kul]
	bye	[bye]
	Beer drinking	[celeb]
	Any other that is not listed	[other]

[27] Head nod

[28] Vigorous head shake

[29] Droopy eyes, taking pillow, snoring etc

Table 3: List of Recurring Smileys, their Emoticon counterparts and Codes

Smiley		Emoticon picture	Meaning	Code
☺ :-)	:->		Basic Smile	[Sml]
:-]	:->>		Happiness, Laughter[30] LOL	[LOL]
;-]	;->> ;-)		Naughty Wink	[Wink]
☹ :-(:-[	;-(;->		Unhappy	[Sad]
;-[	;-(;-<<		Very Unhappy	[Sad]
:-x	;-x		Kiss	[Kis]
:-0			Shock, Horror[31]	[Hor]
;-P			Joking[32]	[Jok]
			Any other that is not listed	[Other]

3.5.12. Other Symbols

Other symbols were defined as all the other possible graphics that were not punctuations, letters, numbers or Smileys. This variable was included to keep track of other symbols that are may be used in each of the genres and their meanings. This variable was automatically tagged in the same way as the punctuations. Examples of the other symbols include

& [Ampersand]
% [Percent]
+ [plus]

[30] Animated laughter

[31] Horrified face with open mouth & head shake

[32] Joking Smile with tongue hanging out

 # [Octothorpe, number]
** [Star]*
@ [at]
other any other e.g./ slash, = equals, $ dollar, £ Sterling pound, ( ) brackets etc.

It was expected that these symbols would be present in the data although it was not clear whether they would be used in a standard way.

3.5.13. Salutations

The term salutation was set to encompass both greetings or sign in and valediction or sign out. I defined a greeting as an opening from the sender to the recipient including pre greetings and post greetings. I defined a valediction as a form of farewell or sign out by the sender of the message. This variable is worth investigating in order to keep track of how these salutations are used to set conversations apart in CMC. I manually identified the presence/absence of greetings and valedictions in each message. The number of these salutations was counted per message, and then later the general total presented per CMC genre, in order to discover the frequency of use of these salutations in each CMC genre. The language of the salutation and valediction was also tagged in order to identify the most frequently used language for these salutations for example:

(31) *sasa, tumefika safely thanx b 2 God*
Hello, we have arrived safely thanks be to God.

This message is analysed as follows:
Greeting: 1 Sheng (*sasa)*
Valediction: 0

(32) *mambo supuu! hp u rem leo 2ko na wewe!@tao nw, 2 arrive approx 2230hrs! c u!!*
Hello beautiful girl! I hope you remember that we will be with you today. I am in town now and will arrive at approximately 22.30 pm. See you!

Greeting: 1, Kiswahili (*mambo)*

Valediction: 1, English (*c u)*

It was expected that just like in standard conversations, salutations would be common in all the CMC genres. IM would have the highest count because of its conversational setting while SNS would have the least count since it is considered as an interruption to put a salutation in the middle of other people's posts.

3.5.14. Language Choice

This variable made counts of the language used. I manually tagged uses of Kiswahili (Sw), English (Eng) and the vernacular languages (vrn) at the message and sentence level when they used one language. The tags were entered in the program for the total counts. This variable is important in the description of the use of language amongst peers in CMC. It was anticipated that there would arise many instances of codeswitching and therefore this variable was further split into interword and intraword codeswitching to tag cases which involved more than one language.

3.5.15. Interword Codeswitching

Interword codeswitching was defined as the use of words from different languages in the same message. This is interesting in CMC as it will enable me to find out if there are any underlying patterns in the CS, mainly in relation to the informal communication principle discussed in section 1.5.

I tagged cases of codeswitching manually by identifying the language of individual words per message. This was done in order to minimise doubts and uncertainties of whether a sentence, phrase or expression is in Sheng or in the individual languages. For that reason I defined codeswitching at the vocabulary level. This was effective because I was able to single out some of Sheng's own unique vocabulary e.g. *njaro* [secret plot/plan] see more explanation in section 2.1.4.

According to my definition, a clear case of codeswitching is in a message like;

(33) *Neatia? Why so quiet? Umenyamaza sana ama jobo ndiyo mob nway Gud day.*
vrn Eng Sw Eng+Sh Sw Sh Eng
How are you? Why are you so quiet? You are too quiet, is it because you have a lot of work? anyway, have a good day.

The counts were done at the word level e.g. in the message:
1 vernacular language (Kikuyu) [neatia]
6 English [why, so, quiet, nway, gud day]
4 Kiswahili [umenyamaza, sana ama, ndiyo]
2 Sheng [jobo, mob]

The languages were categorised into different groups titled, English, Kiswahili, Sheng and the vernacular language category. A message like:

(34) *chamge! Maze siwezi kulenga.U really mean a lot to me, av a tyt 9t.*
vrn Sh Sw Sh Eng
Hello, surely I cannot ignore you. You really mean a lot to me. Have a good night

was given a count of

11 English, [*U, really, mean, a, lot, to, me, av, a, tyt, 9t.*]
1 Kiswahili [*siwezi*]
2 Sheng [*Maze, kulenga*]
1 vernacular [*Chamge!*] Nandi.

3.5.16. Intraword Codeswitching

I defined Intraword CS as the use of morphemes from different languages in the same word. Just like interword codeswitching in 3.5.15, this variable was associated with the informal communication principle. I manually tagged cases of intraword codeswitching and entered them into the program database. For example, in words like

(35) *sikurespond*
siku+respond
Sw Eng
I did not respond

This involves a codemix of 2 languages: Kiswahili negative prefix

si and negative past *ku* and the English verb *respond.* Another example that also involves 2 languages is

(36) *dropishad*
drop-i-sha-d
Eng Sw Eng
Dropped them (Gave them a car ride to their destination)

The main interference faced in identifying cases of intraword codeswitching is that in many cases, there is no clear distinction between cases of codeswitching and Sheng due to the fact that the matrix language of Sheng is Kiswahili and/or English. If Sheng is treated as a separate language, which is justifiable in many respects, it is no longer a possible language in the tagging of codeswitching. This did not present a complication because my main interest was the identification of the different language combinations in the words and not the product language. I tagged all the vocabulary that involved codeswitching in a separate category.

3.6. The eDatax.mdb Program

I used the *eDatax.mdb* program for the automatic counts for the variable's quantitative data. This program was developed by Mr. Maarten Hijzelendoorn who is a computational linguist at Leiden University. The program greatly assisted in the automatic identification of the data owing to the large corpus. It did the calculations in 3 phases ranging from the message level, to the sentence level and finally, the word level.

In the first phase, the raw data was entered manually into a window in form of a complete message e.g. an SMS, Email, IM or SNS comment. The only modification here was the coding of Smileys and Emoticons in order to avoid the program coding Smileys as excess punctuations and at the same time only identify text and not pictures of Emoticons.

The next phase involved splitting the message into sentences which were also entered manually in the next window with each sentence being put on a different line in the sentence window. At

this stage, the program automatically calculated counts related to sentences for example the capitalisation of the initial letter of first word in the sentence.

The third level involved the splitting up of the sentences into words. The words were entered into a third window and automatically counted in terms of clippings, contractions, language(s) used, etc. After all these entries, the program then did a general computation based on the whole message.

Logically the program could not identify all the variables automatically and therefore I was left with the task of manually checking each message and making necessary adjustments. For example, the program could not recognise the required data in variables like codeswitching, phonological spelling, affixation etc. Similarly, in other cases, it could not identify all the data that was in the same category, for example in abbreviations, it could only identify abbreviations composed of only consonants e.g. *stn [station]* and not those with vowels e.g. Msa [Mombasa]. Conversely, for some variables, the program identified data that did not exactly fit into the variable. For example, identifying a valid abbreviation as having excess capitals e.g. NGO.

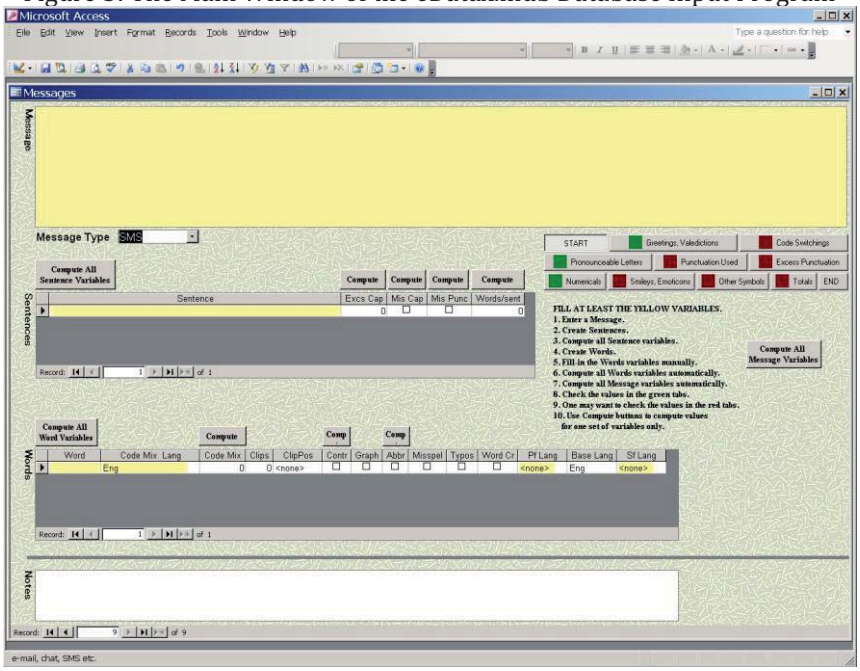
In order to maintain reliability and consistency in the total counts, I had to remedy these limitations of the program by manually determining and adding all the data that could not be identified automatically or subtracting all the excess data automatically identified by the program. This has been explained for each of the variables in section 3.5.

Figure 5 shows the visual presentation of the input form of the database program.

3.7. Procedure of Data Analysis

I selected an overall of 5427 messages for all genres distributed as shown in table 4. I then went through the data and extracted the necessary illustrations for the qualitative description and discussion of the features.

Figure 5: The Main Window of the eDatax.mdb Database Input Program



The qualitative description of the features presented is general and does not focus on each of the genres independently unless where indicated.

In addition, some of the illustrations used are excerpts from longer (entire) messages. I have only extracted and presented the sections that serve the purposes of the discussion. Each descriptive point in the discussion is presented accompanied by an illustration from a CMC message and an English or semantic translation where necessary.

From the overall message total of 5427, I selected 579 messages for the quantitative analysis of the variables using the *eDatax.mdb program*.

Table 4: No. of Total Messages Collected

Genre	Total Messages
IM	197
Email	780
SMS	2730
SNS	1720
Total	5427

This was not only prompted by the time limit but also by the fact that I considered the 579 messages as a balanced and representative sample for quantitative comparison. The selected messages per genre were distributed as shown in table 5.

Table 5: No. of Total Messages against those Selected Per Genre

Genre	Total Messages	Selected Messages
IM	197	32
Email	780	35
SMS	2730	300
SNS	1720	212
TOTAL	**5427**	**579**

In a bid to minimise bias in the selection of the 579 messages, I selected every fourth message in each genre until the total number of words per genre added up to between 4325-4350. The motivation for choosing this number of words was that I considered a total of 15000-20000 words a representative sample.

Table 6 gives a complete breakdown of the number of messages, sentences, words and characters selected for this study. Noticeably, there appears to be a high disparity between the number of messages per genre with IM at 32 and SMS at 300. This is explained by the fact that some genres like IM and Email do not have total character limits and are longer.

Table 6: Selected Messages Per Genre

Genre	Selected Messages	Sentences	Words	Characters
IM	32	555	4325	23221
Email	35	240	4347	21370
SMS	300	817	4347	20173
SNS	212	553	4349	24965
TOTAL	**579**	**2165**	**17368**	**89729**

Similarly, the number of sentences has Email at 240 while SMS has 817. As a result one Email or IM may have an average of 17 sentences, 135 words and 725 characters while SMS and SNS are shorter because they have a character limit with one message registering 2 sentences, 14 words and 117 characters on average. Hence it would not be representative to give an equivalent total number of messages per genre. I have used the total number of words as the uniformity target. Because of the balanced sample that is used for the quantitative research, I can make statements that compare the relative use of a particular variable across genres.

Chapter 4. Technologically Motivated Features of CMC

Bodomo (2009:22) points out that there is a causal relationship between new technology and new forms of language. Therefore knowledge about the intricate relationship between new technology and language use is essential. Hård af Segerstad (2002:1) comments that communication that is mediated by technology seems to affect the written language. It is clear that the relationship between language and technology cannot be denied.

This chapter presents CMC to the extent that it is technologically driven. It reveals some features of language use that are mainly influenced by technology. These features result when users adapt the language of their message to suit the technology involved leading to what many distinguish and refer to as CMC language. The chapter includes the discussion of features like phonological spelling, pronounceable letters and numericals, and different forms of lexical compression like abbreviations, consonant spelling, contractions and clipping. I proceed to describe additional features; for example relaxing spelling standards, changing use of punctuation and the increasing and innovative use of graphics in CMC texts, all influenced by the technology.

All the features that are discussed in this chapter have been identified from the variables presented in section 3.5. I have split the results into features that are influenced by the technology which I present in this chapter and those that are socially driven which will be presented in chapter 5. As already noted, some features fall into both groups and it is not immediately clear where to place them. Such features have been placed where they may have more correlation. Note that this is only an organisational matter since all the occurring features and their motivations are discussed.

4.1. Phonological Spelling

> *dint mek it 2 chach coz gyz wok up2 runin batls n so steid indoz. Bt its pikin up sloli I kan c ppl wokin. m sef dnt wori. mis u*

This study defines phonological spelling as a feature of CMC language in which words are written in a way that reflects their

phonetic transcription. According to Frehner (2008:52) these are words that are spelt in such a way that they approximate their phonological value. She refers to them as non-conventional spelling.

In an attempt to give a possible description for this phenomenon, I support the earlier assertion by Frehner (2008:52) and Barasa & Mous (2009), on phonological spelling which claims that CMC is mostly written in the way in which it is spoken. The data also gives evidence of the fine line between the written and spoken. Users of CMC compose and write messages mainly based on speech, and in turn, in order to understand the received messages, the receivers read the messages actively.

Table 7: Distribution of Phonological spelling

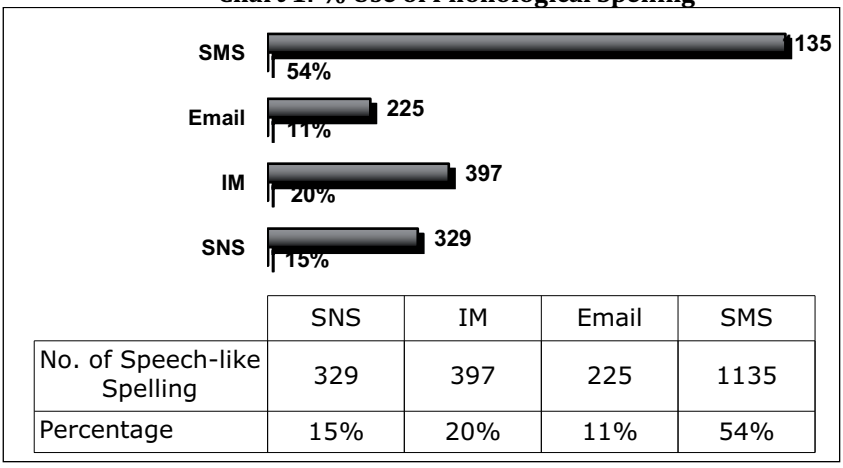
Genre	Phonological spelling
IM	397
Email	225
SMS	1135
SNS	329
Total	**2086**

Phonological spelling in Kenya is closely influenced by Kiswahili. Kiswahili orthography stays very close to pronunciation. Kiswahili phonology is characterised by a CV syllable structure. It has a five vowel system [a], [ɛ], [i], [ɔ] and [u] represented as *a, e, i, o, u*. Users adapt Kiswahili spelling on English words. This leads them to write words in relation to their sound. Nevertheless, it is important to note that in many cases the relationship of the word to the sound presented is not exact but an approximation. Examples of these are in illustration (43) *wea* for *where* and (44) *kliaring* for *clearing*,

Instances of the phonological spelling in the current study's re-

sults are distributed among genres as shown in table 7. As expected, SMS at 54% has the highest partition of the phonological spelling instances. This is attributed to the principle of least effort coupled with the mode limitation principle. It seems easier to use phonological spelling while writing an SMS than to focus on the standard spelling.

Chart 1: % Use of Phonological Spelling



	SNS	IM	Email	SMS
No. of Speech-like Spelling	329	397	225	1135
Percentage	15%	20%	11%	54%

The majority of these phonological spellings also lessen the number of keystrokes and the space used. This serves for both least effort and mode limitation. Examples of this are in illustration (37)-(39).

(37) *BEWARE U AV BEEN* ***XPOZD*****-----**
Beware you have been **exposed**. [SNS]

(38) *Hi* ***gud aftanun****. Ope u having a* ***gud*** *day.* ***Miting*** *was ok. Jus* ***tok*** *of* ***rili*** *gen stuff a speaker 4rm min of education* ***kem****.*
Hi, **good afternoon**. Hope you are having a **good** day. The **meeting** was ok. They just **talk**ed of **really** general stuff. A speaker from the Ministry of Education **came**. [SMS]

(39) ***hav*** *u* ***chekt*** *on the* ***injinia*** *in the* ***vilej*** *or u only* ***saspekt*** *??*
Have you **checked** on the **engineer** in the **village** or you only **suspect**? [IM]

In illustration (38) which is an SMS, *aftanun* uses 12 keystrokes while *afternoon* has 20 keystrokes on the phone keypad. In (39), apart from *saspekt* all the phonologically spelt words in the illustrations contain less key strokes when compared to the standard

spelling. For example on the computer keyboard *chekt* uses 5 keystrokes while *checked* uses 7 keystrokes. Phonological spelling is done to save effort. It is also done to save space where mode limitation plays a role. However, some words in illustration (38) like *stuff, speaker* and *education* do not make use of the phonological spelling. It is likely that the word *stuff has* not been spelt as *staf* in order to avoid its misinterpretation from *staff*. The fact that *speaker* and *education* are not spelt in a phonological manner shows that it is unlikely that CMC users consider each word consciously for economic spelling.

IM and SNS account for 20% and 15% of the phonological spelling respectively. Email has the least contribution at 11%. The expected findings were that all genres would register the presence of phonological spelling because of the least effort principle. However, SMS and SNS which are the genres with character limitation were expected to have a higher presence of the phonological spelling than Email and IM. The findings show that SMS has the highest phonological spelling occurrences while the Email has the least occurrences as expected owing to its more formal nature in comparison with the rest of the genres.

Surprisingly, SNS has a relatively low count of phonological spelling despite expectations that it would have a high count. One of the reasons for this is related to mode limitation in that the SNS character limit is much easier to work with than in SMS, therefore the typing in SNS is done in a more standard manner without too much pressure to save space. Additionally, for some official forums, the SNS degree of formality in relation to spelling is higher than that of SMS and IM because of its public nature. IM has a lower count on phonological spelling when compared to SMS because of the absence of mode limitation thus it is not a major advantage to shorten the spelling for this reason. The hypothesis suggestion derived from these results is that phonological spelling has always been a practice in SMS and it has currently extended into the other CMC genres.

A point to note is that in some cases, words are written using the phonological spelling without reflecting any shortening. An example of this is in illustration (39) where the word *suspect* is spelt as *saspekt*. Notably, this does not involve any shortening to support the least effort and mode limitation principles. This shows that although phonological spelling may have been exclusively used in least effort and mode limitation contexts, it is currently a 'standard' CMC style. The use of phonological spelling is a common feature and is registered in previous researchers like Bodomo (2009:70) who groups it under approaches to shortening. Frehner (2008:104) refers to it as phonological approximation. She claims that this spelling hardly occurs in Emails and is exclusively used in SMS. From this study's findings, phonological spelling mainly occurs in SMS due to the mode limitation and least effort principles. However, it has been introduced into the other CMC genres. Additionally, the English phonological spelling in Kenyan CMC is advanced by the nature of Kiswahili writing which closely reflects pronunciation.

This section provides some examples of the different types of phonological spellings from the data. The phonological spelling in the original message is highlighted. The message is followed by a translation into the conventional writing. Note that some messages have used both Kiswahili and English words. The general English translation has been presented below it.

(40) *I* ***kof*** *@* ***ua thot, sniz*** *@* ***ua smel n*** *cry* ***wen u smyl*** *@ me* ***coz u r 2*** *much* ***4*** *me*
I **cough** at **your thought, sneeze** at **your smell and** cry **when you smile** at me **because you are too** much **for** me. [SMS]

In illustration (40), besides the phonological spelling, the writer has incorporated the symbol @ for *at* and the use of single pronounceable letters like *n* for *and, u* for *you* and *r* for *are*. Further, numbers have been used in place of words that share their pronunciation for example *2* for *too* and *4* for *for*. The ensuing illustrations show the informal nature of such spelling.

(41) ***M orait*** *niko ofisi tu. iv* ***oredi*** *gt* ***enuf*** *jobo!!*
Sw Sw Sw Sh

I'm alright and just in the office. I've **already** got **enough** work. [SMS]

(42) *haaiiyaa!!! I 4got my **pas baq hom**!*
Oh no! I forgot my **purse back** at **home** [SMS]

In illustration (41), it is notable that *al* in *alright* has been changed into its pronounced form *o* in both *alright* and *already*. However, *enuf* takes on a different trend whereby it is expected to be *inaf* but the sender retains some of the original letters in the word or uses alternative English spelling principles and not Kiswahili ones. Illustration (42) clearly shows phonological spelling for *pas* and *hom* as a result of least effort. Interestingly, instead of *k*, *q* has been used in the word *back*. The possible explanation to this is the creative show off. Other examples of phonological spelling are:

(43) ***wea u gaiz hedin 2nait? am lukin 4*** *a hot plot*
Where you **guys heading tonight**? am **looking** for a nice plan. [SMS]

(44) ***shez kliaring*** *the* ***cos nest*** *yr.* ***hop shel*** *get* ***wak afta!!***
She's **clearing** the **course** next year. **Hope** she'll get **work after**wards [SMS]

(45) *Send the **rekodz ova** pls*
Send the **records over** please. [IM]

In illustrations (43) and (44), the words *wea* and *kliaring* do not exactly reflect the Kiswahili pronunciation. In the same illustrations, *gaiz* and *kliaring* have the same number of characters as their standard spelling *guys* and *clearing*. In the SMS genre, the spelling of g*aiz* involves fewer keystrokes than *guys*. This shows that the phonological spelling in this case may not due to mode limitation but least effort. The spelling of *kliaring* has the same number of characters and uses the same number of keystrokes similar to the standard spelling *clearing*. This is a case whereby phonological spelling is not only determined by least effort or mode limitation. The possible explanation is that the message composer wants to maintain uniformity in using phonological spelling for the whole message. In illustration (45) the words *records* and *over* have been spelt in a phonological way unlike in the standard spelling. Additional illustrations of phonological spelling are:

(46) *R u **kamin** 4 de recoln kesho? **Op** 2 c u*
Are you **comin(g)** for the recollection tomorrow? **Hope** to see you. [SMS]

Comparable to the phonological spelling in *saspekt* and *kliaring*, illustration (46), spells the word *coming* with a *k*. This spelling which could be due to influence from Kiswahili reflects its phonetic transcription better. The final *g* is also omitted since the writer considers it as silent and hence omissible. Similarly, *h* and *e* have been left out in the word *hope* as they are considered silent.

(47) *jus wishn u a qt **dey***
Just wishing you a quiet **day**. [SMS]

Illustration (47) has shortened all its words and then written out the word *day* based on its close phonetic transcription.

(48) *Heeey!!! **Wats** with the **sylens**? a **thot** we r **frenz**?? Beat it **gal**....*
Hey, **what's** with the silence? I t**hought** we are **friends**? Beat it **girl.** [SNS]

In illustration (48) all the words with phonetic spelling are shorter than when using conventional spelling. Notably, the word *thought* has been spelled as it is pronounced.

y for/ai/ in English words

As part of using phonological spelling, people also use letters whose standard pronunciation represents the intended pronunciation (cf. 4.2). This involves the use of a pronounceable letter as a spelling symbol for a specific sound. It is worth noting that this is not a practice in Kiswahili spelling which in fact spells words through their closest phonetic transcription. Examples of this involve the mechanism of using *y* for */ai/.* This use is interesting because it involves two processes. First is the choice of the phonological spelling *y* whose standard represention should be the sound */wai/.* Secondly, there is the deletion of the initial */w/* leading to */ai/*. These are in the following illustrations,

(49) *stay a little **wyl** longer.he **myt** change wit **tym***
Stay a little **while** longer, he **might** change with **time**. [SNS]

Others examples include,

(50) *airtym*- airtime
(51) l*yf* is *tyt*- life is tight
(52) *gyz* - guys
(53) *tryd*- tried
(54) *nyc* - nice
(55) *syds lyk*- sides like
(56) *gudnyt*- goodnight
(57) *wyz* – wise

4.1.1. Determiners and Demonstrative Pronouns

Phonological spelling for least effort and space economy in CMC is also manifest in form of the reduction of determiners and demonstrative pronouns. The formula is to convert the voiced dental fricative /th/ into an alveolar plosive /d/ as shown in the data (58) - (61).

(58) *hey, u'r jst **de** 1. thnks sana*
Hey you are just **the** one. Thanks a lot. (*sana* is Kiswahili for *a lot.*) [SMS]

(59) *I'll tek u 2 **dat** kajoint 4 nyam chom! u'll lov it!!!*
I'll take you to **that** little joint for roast meat. You will love it! [IM]

(60) *hi, can i kam 2 pik **de** car **dis** wknd?*
Hi, can I come to pick **the** car **this** weekend? [SMS]

(61) ***dey** got a baby boy! im so happy 4 **dem** yaani finally!!*
They got a baby boy! I'm so happy for **them**, that is they have finally managed! (yaani is a Kiswahili word for *that is.*) [IM]

(62) *I don think **doz** jamaaz r serious….*
I don't think **those** guys are serious. [SNS]

The two characters in the voiced dental fricative /th/ are converted to one character in form of the alveolar plosive /d/ whose pronunciation approximates the voiced dental fricative.

4.1.2. Use of /z/ and influence of Engsh in CMC

de sch advanced me sm chumz 4 de tiko b4
de official releaz of our saloz

The school advanced me some money for the ticket before the official release of our salaries

The principle of informality in CMC texts is clearly seen in the data through the phonological spelling of Engsh. This kind of spelling in the data reveals some remarkable recurrent features that are likely to have stemmed from Engsh.

It is important to reiterate that Sheng and Engsh codes are unavoidable when dealing with the youth and communication in Kenya. However, their manifestation has mainly been in speech (verbal) form. Previously, it was very uncommon to find Sheng or Engsh in written forms. Therefore it is notable that currently with the emergence of CMC, Sheng and Engsh have invaded the written scene.

Users have comfortably adopted the structure of Engsh into CMC texting. Prior to the ensuing discussion on Engsh, I wish to state the provision that many times, it is difficult to establish what actually Engsh is and what is just based on phonetic influence. For example considering the lexeme *easy,* if it is written as *eazy* one cannot be sure if the writing is based on pronunciation or if it is intended to be Engsh or if it just an alternative norm in spelling.

For the purposes here, I propose that the recurrent features under discussion may have been influenced by the way the words are pronounced in Engsh. This section presents the data detailing and describing these recurrent features and makes an attempt to explain them.

/z/ Suffix in Engsh[33]

Many messages contain words that use /z/ at the suffix position. This /z/ appears in different contexts as described in this section. It is important to note that from a linguistic view as discussed in section 5.5.7, some of the words have the /z/ suffix solely as an Engsh marker. For example,

[33] The use of slashes indicates orthographic representation of letters and not linguistic units.

(63) anywayz (anyway) [IM]

(64) sawaz (sawa - ok) [SMS, SNS, IM, Email]

(65) hukoz (huko - there) [SMS, Email]

This conclusion is reached based on the observation that in many such words, the /z/ suffix does not hold any linguistic function. It acts as a filler to mark the word as being Engsh. In spite of this, there are many other cases as described here where the /z/ suffix acts as a replacement for a suffix written as *s,* e.g. *carz* for *cars*. This follows speech but has now been integrated into CMC spelling not only to replace the suffix *s* but also to mark the word as Engsh. More data on this is presented in table 13.

Use of /z/ as a Plural Marker

To start off, I present the use of the /z/ as a plural marker in English, Kiswahili and Sheng words with -z suffix.

Creation of Engsh words from English

The data in table 8 shows a recurrence of the use of /z/ to mark plurals. For English nouns, this /z/ is used in the orthographic position of the *-s* suffix in English regular plurals. It is noted that in the cases where the /z/ only marks the plural, the nouns lose it (/z/ suffix) in their singular forms.

Therefore, it is clear that it is only in the plural usage that these words are recognizable as Engsh (cf. 5.5.6). In the singular forms they are identical to their form in English (or Kiswahili).

Creation of Engsh words from Kiswahili and Sheng

The /z/ suffix is also used as a plural marker in some words of Kiswahili origin and also in words that are purely Sheng as shown in table 9:

Table 8: English Nouns with /z/ Suffix

CMC Noun	English
66. kidz	kids
67. gyz	guys
68. skillz	skills
69. fingaz	fingers
70. prayerz	prayers
71. frenz	friends
72. Kenyanz	Kenyans
73. buddyz	buddies
74. lolz	Lols[34]

Table 9: The /z/ suffix in Kiswahili and Sheng Words in Engsh

Noun+/z/ suffix	Origin	Sheng Singular	Meaning
75. *mamaz*	Kiswahili *mama*	*mama*	mothers/women
76. *jamaaz*	Kiswahili *jamaa*	*jamaa*	people
77. *shidaz*	Kiswahili *shida*	*shida*	problems
78. *mamboz*	Kiswahili *mambo*	*mambo*	news
79. *wagondiz*	Sheng *gondi*	*mgondi*	thieves
80. *wasupuuz/supuz*	English *super*	*msupuu/suupu /supu*	attractive girls
81. *junguz*	Kiswahili m*zungu,*	*jungu*	caucasians
82. *zunyez*	Kiswahili *mzungu*	*zunye*	caucasians
83. *odieroz*	vernacular (Luo) *odiero*	*odiero*	caucasians
84. s*kumz*	Kiswahili *sukuma wiki*	*skunje*	kales (a type of vegetable)

In the Sheng examples (79) and (80) *wa* which is the Kiswahili plural marker (noun class 1/2) is used in combination with the /z/ suffix which assumedly takes over the function of the English

[34] This is used as the plural form of lol in order to mean *lots of laughs*.

/s/ plural marker. Therefore in essence, there is a double plural marking in the words.

/z/ Plural Marker Suffix preceded by /o/ in Ellipsis

Another Sheng/Engsh process that is reflected in nouns in CMC is the addition of */o/* suffix to clipped words. Bosire (2009:83) refers to it as the truncation word finally with a final dummy suffix. In such cases, the word is clipped and an /o/ added to make it Sheng, followed by a /z/which functions as the plural marker just like /s/ in regular nouns in English. Illustrations of this are on table 10.

Table 10: /o/ + /z/ Plural marker

CMC Plural Noun	***Singular***	**English**
85. storoz	*storo*	stories
86. campoz	*campo*	campuses
87. joboz	*jobo*	jobs
88. estoz	*esto*	estates
89. juoz	*juo*	juice
90. satoz	*sato*	saturdays
91. saloz	*salo*	salaries
92. kidoz	*mkido*	kids
93. haoz	*hao*	houses
94. cuzoz	*cuzo*[35]	cousins
95. thayoz	*thayo*	thighs

Notably, the /z/ plural marker suffix is also added to both Kiswahili and English nouns that originally contain /o/ as the final vowel. For example in some cases, the Kiswahili word *mambo* becomes *mamboz*.

(96) ***Mamboz***??
Hi. (Mamboz is used as a greeting but it literally means news or issues.)

(97) uliweka wapi zile **fotoz** zangu?

In the case above, the English word *photo/foto* becomes

[35] In this example, instead of *cusoz,* /s/ has been replaced by /z/ to form *cuz.*

photoz/fotoz.

Another occurrence to make note of is that there are some Engsh words which do not exist without the -z. Although this -z has its origin as a plural marker, in these cases it does not serve as an element of plurality. Illustrations of this are presented in table 11. In illustration (99), s*hagz* (rural home) *is* derived from the earlier form *ushago* (rural home), and in (100), *chumz* (money) is derived from *chumas* (pl. coins).

Table 11: Compulsory non plural marker /z/ suffix

CMC Noun	Unacceptable	Origin	Meaning
98. bunduz	*bundu	-	middle of nowhere
99. shagz	*shag	Sheng *ushago*	rural home
100. chumz	*chum	Kiswahili *chuma* (coin)	money
101. dig(s)z	*dig	English *dig*	home
102. mdabuz	*mdabu	-	father

Note that the * symbol is used to indicate ungrammatical illustrations.

/z/ Suffix in adverbs

Other words (adverbs) that appear with the /z/ suffix are included on the list in table 12. The main interest here is to illustrate that /z/ is an Engsh marker.

Table 12: /z/ Suffix on Adverbs

CMC Engsh	Kiswahili	English
103. laterz	baadaye	later
104. baadayez	baadaye	later
105. hukuz	huku	here
106. nyumaz	nyuma	back
107. keshoz	kesho	tomorrow
108. bilaz	bila	without

The /z/ suffix in the illustrations in table 12 is not from the phonetic influence, and neither is it a plural marker like in table 10. Bosire (2009:82) explains a similar dummy affixation of /o/ and

/sh/ in Sheng. He expounds that such a suffix is a dummy and does not add any new meaning other than alter the phonological shape of the words. The dummy suffix /z/is therefore a tool fashioned by speakers of Engsh to manipulate the shape of lexical items and give them a foreign sounding character. The suffix is therefore present only as a marker of Engsh (cf. 5.5.6).

/z/ as a pronunciation spelling in English words

For lexemes with English origin, the *-z* suffix is added due to the influence of pronunciation of the words as shown in table 13.

Table 13: /z/ suffix on English words

CMC Sheng	English Equivalent
109. athaz	others
110. therz	there is
111. ... iz where doez thiz	... is where does this
112. wat iz diz	what is this
113. sinz	since
114. itz	its
115. hiz	his
116. lotz	lots
117. thoz	those
118. nowadayz	nowadays
119. advaiz	advice
120. waz	was
121. bcoz / coz	because
122. plz / pliz	please
123. bizna[36]	business
124. bizzy/bizi	busy
125. huzzy/huzy/[37]*/hubby*	husband

As already noted, the addition of the *-z* mostly occurs word finally. It is feasible that this practice first started off as a way of pronouncing the plural *-s* marker in the American English way. Over

[36] Use of the pronunciation spelling /*biz*nes/ and the replacement of *es* with *a*.

[37] This example also illustrates the process of clipping the final characters *-(s)band* and the insertion of the *i* and ending up with a shorter word *huz(z)y* compared to the original *husband*.

time it was generally used to replace instances of *s* that occur in the plural position which is at the end of the word. It is currently starting to be used as a general replacement for s in the root of words. This is an instance of new conventions developing in CMC.

4.1.3. Conclusion

The revolution of CMC in communication has promoted deviations from the standard orthography. The use of phonetically influenced spelling in phonological spelling is the deviation that has been discussed in this section. CMC users write messages closer to pronunciation. In fact, in view of this, a new orthography is emerging much more based on the spoken word because of the merger of oral and written communication in CMC.

One of the conclusions derived from the findings here is that indeed phonological spelling is mainly found in SMS primarily because of the least effort and mode limitation principles. These principles encourage the use of direct pronounceable spelling by reducing the effort of ensuring the use of the correct standard spelling and also ensuring the reduction of the characters used. This phonological spelling variable is also to a lesser extent reflected in the other genres due to the least effort principle, rapidity and the expression of peer identity. The genre's degree of formality also contributes to the occurrence of phonological spelling with the more formal genres registering a lower count than the highly informal genres. It is worth noting right at the beginning that SNS forums are divided into two; the more official SNS forums use language in a more conventional way while the relaxed forums use it in a more informal way.

Another conclusion is that the consistent use of the /o/ and /z/ suffix which is a typical marker of Sheng and Engsh respectively, points to the informal and orally based nature of CMC language. Unlike before, Sheng and Engsh have now expanded their reach and entered into writing through CMC. Some of the processes that compose Sheng/Engsh have been retained in addition to new ones which have been invented in order to fit into the needs of the CMC texting scenario. It is also important to point out that since Sheng

and Engsh have previously been oral, many spellings are not consistent yet. A user tends to use what suits him/her best in terms of space, effort, and so long as the intended meaning is achieved. An example of this from the data is the Engsh word *huzzy/hubby* for *husband* which registered different spellings including, *hazzy/habby, hazzi/habbi, hazy/haby* and *hazi/habi.* Similarly, *dudettes* (female equivalent of dudes)-*doodets(z), dudets(z), dudeits(z), dudeyts(z)* etc. It may prove difficult to standardise the spelling because even some words in English have several different spellings in CMC for example; *guys-guyz, gyz, gais, gaiz*. Another example is, *goodnight – gnight, gudnight, gdnight, gnyt, gudnyt, gdnyt, gnite, gudnite, gdnite, gnt, gudnt, and gdnt.*

4.2. Pronounceable Letters

vp? y r u let?? l b on q @ de dh 4 de g c u
Vipi? (what's up?) why are you late? I'll be on queue at the dining hall for the grand (Kshs 1000/=) see you.

The English alphabet contains 26 letters made up of 21 consonants and 5 vowels. The Kiswahili alphabet is also Latin based and is similar to the English alphabet except for the fact that Kiswahili language does not make use of *q* and *x*. Each of the alphabet letters has a pronounceable sound that is based on the standard pronunciation of single letters as learned in school for example [ar] for *r*. Some of the alphabet letter pronunciations are identical or close to the pronunciation of some words. The substitution of these words with the single letters that sound similar is very common in CMC text. This practice is especially common in SMS. These letters are what I refer to as pronounceable letters. All the genres make use of the pronounceable letters as demonstrated in table 14 and chart 2 and 3.

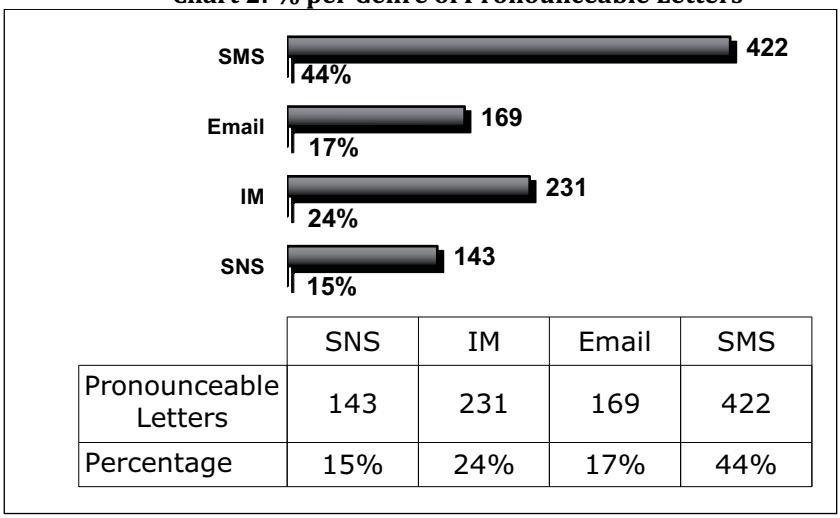
The results in table 14 show that the vowel *u(you)* is most popular in SMS. The consonant *b(be)* is mostly used in Email and SMS.

Table 14: Distribution of Pronounceable Letters among Genres

Genre	u	b	c	m	n	r	x	y	other	total
IM	65	16	13	22	53	43	9	7	3	231
Email	39	24	18	23	29	19	11	6	0	169
SMS	185	19	43	35	60	53	10	8	9	422
SNS	37	13	10	9	17	36	9	7	5	143
Total	**326**	**72**	**84**	**89**	**159**	**151**	**39**	**28**	**17**	**965**

The consonants *c(see)* and *m(I'm)* are popular in SMS while *n(and)* and *r(are)* are mostly used in SMS and IM respectively. A general justification to these totals could be that the pronoun *you* which happens to be represented by the vowel *u*, is a common part of all CMC messages. Similarly, the conjunction *n(and)* and auxiliary verbs *m(I'm)* and *r(are)* are also common parts of messages.

Chart 2: % per Genre of Pronounceable Letters



	SNS	IM	Email	SMS
Pronounceable Letters	143	231	169	422
Percentage	15%	24%	17%	44%

SMS (44%) leads with close to half the occurrences of pronounceable letters, followed by IM (24%). Email and SNS have the least part with 17% and 15% respectively. These results are different from the expected results which were that all the genres would register high counts of pronounceable letters because of the least

effort principle. IM was expected to top the list because of its additional feature of rapidity and then SMS and SNS because of their character limitation.

In these results, SMS has the highest count of pronounceable letters. This implies that the role of mode limitation and least effort surpasses that of rapidity in their influence on the use of pronounceable letters in CMC. These principles have led to the use of pronounceable letters over time until they have taken root in SMS. Another main principle that contributes to the popularity of pronounceable letters is informality and competition in innovation among youth peer communication. The use of letters in a pronounceable manner gives the message a typical SMS appearance which is considered acceptable among innovative youth. This acceptability is important as it leads to further acceptability of the user in these circles. This aspect of youth identity has popularised the use of single letters to represent a word or part of it as SMS language. The results of IM in the use of pronounceable letters are related to its rapidity where it is easier to type a single letter to represent a whole word. In regard to Email, these results are similar to Frehner (2008) who found a wide margin between Email and SMS use of pronounceable letters. In her results, Email had 264 cases while SMS had 1237 cases, which translates to 18% and 82% respectively. In my data, although there is a disparity, the margin is not as wide suggesting that the participants are more innovative in Emails or more conservative in SMS. Frehner's results show that her population group which also constituted university students hardly made use of pronounceable letters in Emails.

In comparison to SMS and IM, the low contribution of Email in my results is because of its more formal and conventional nature. The lack of character limitation also does not promote the use of pronounceable letters in Email. The surprising result is that SNS has the least count yet it also has a form of mode limitation like in SMS. The reason for this is that the character limitation in SNS is less restrictive than in SMS. In addition, SNS is not as rapid as IM and thus has no immediate urgency to shorten words. In contrast

to Email and the other genres, SNS is less personal and is posted publicly and mostly to strangers who may not necessarily understand the intended message if they are unable to decipher what the letters represent. This explains to some extent why it does not apply as much pronounceable letters and numbers as Email (cf. 4.3).

Chart 3: % Use of each Letter

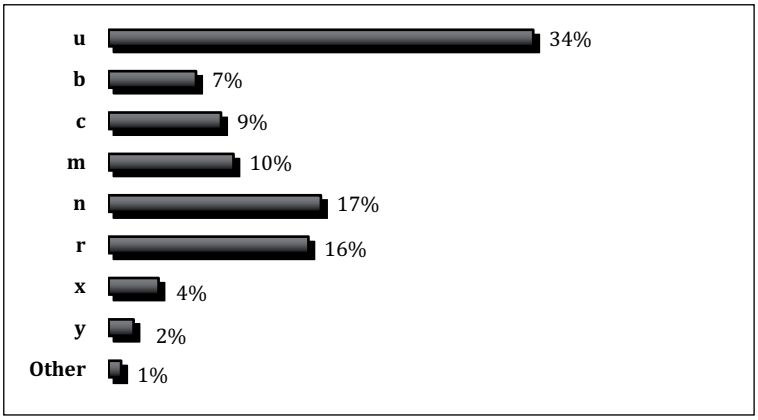


Chart 3 displays the use of each pronounceable letter. Similar to Frehner (2008:98) my results further reveal that the *u* vowel for the pronoun *you* is the most used. This can be attributed to the multiple uses of *u* as a personal pronoun for the 2nd person in both singular and plural for both subject and object positions. It is also used as both a generic and reflexive pronoun. It is therefore likely to appear in many messages. The consonants *n(and)* and *r(are)* are commonly used as well.

The consequent section discusses pronounceable letters that only occur in place of another word in CMC. The standard occurrences of single letters are not discussed. Therefore a vowel like *a* in (126) is not regarded as a pronounceable letter in the data.

(126) *pls buy **a** pckt 4 me*
Please buy a packet for me. [IM]

Similarly, *I* in (127) is not classified as a letter homophone when it acts as a personal pronoun.

(127) ***I*** *may go there nxt wk*
I may go there next week. [Email]

The discussion begins with vowel homophones and moves on to consonant homophones.

4.2.1. Pronounceable Vowel Letters

u [you]

The only vowel used as a letter to mark a word is *u* for the pronoun yo*u*. It is the only vowel symbol for which the alphabet pronunciation in the English alphabet starts in a consonant. Its pronunciation results in an existing English word *you* and it is only used for that word. There is no Kiswahili word *yu*. The closest there is, is the Kiswahili third person singular human prefix *yu e.g. yuko wapi?* (*where is he/she?*). There and nowhere else would I expect the use of the homophone *u* in Kiswahili. Nevertheless if it were used in that context, it becomes *uko wapi? (where are you?)* which exists with a different meaning and pronunciation.

The first column in table 14 indicates that the *u* homophone is the most commonly used pronounceable letter in all genres. It is used 326 times (34%). It is most popular in SMS.

In the data, the vowel *u* occurs in place of the pronoun *you* both singular (128) and (129) and plural as demonstrated in (130) and (131).

(128) *how r* ***u*** *doin?*
How are you doing? [IM]

(129) *I stil lv* ***u***
I still love you. [SMS]

(130) ***U*** *guyz r lost sana* (used to show that they haven't been seen in a while)
You guys are really scarce. [Email]

(131) *our car has a puncture!! can* ***u*** *jamaaz hlp us out?* [SMS]
Our car has a puncture! Can you guys help us out?

4.2.2. Pronounceable Consonant Letters

The usage of consonant homophones in the corpus is as follows:

b [bi]

The consonant letter *b* is used as a replacement for *be,*

(132) *Il stil **b** at work*
I'll still be at work. [IM]

(133) *jus **b** strong 4 me* [Email]
Just be strong for me.

(134) *2mefika, thnx **b** 2 God*
We've arrived safely thanks be to God. [SMS]

It is possible to use *b* to replace *bee* but this was not observed in the data. There was no mention of the word *bee* for observation. These homophones are only used for certain lexical items or there is generally a strong tendency to use such homophones for a single lexical item. The use of *b* as a homophone in Kiswahili is not common because Kiswahili does not have a self-standing word as *bi* apart from the ellipted *bibi* (lady, wife) in *bi.* There was one message with this kind of use.

(135) *alifukuza **bb** jana*
bibi
He sent away the wife yesterday. [IM]

c [si]

c is used as a replacement for the English word see. It is also possible to use it to replace sea although the use of the word did not occur in the data.

(136) *I cant **c** it*
I can't see it. [SNS]

(137) ***c** u*
See you. [SMS, IM, Email]

c is also used to replace the Kiswahili negation marker si. In Kiswahili, the grammatical morpheme for the negation marker for

the 1st person singular is si-. This is replaced with c in some instances,

(138) *imagn **c**na sahii*
Imagine **si**na saa hii.
Imagine I don't have it right now. [SMS]

(139) ***c**kuja*
sikuja
I did not come. [SMS]

Letter *c* can also be used to represent the Kiswahili word *si* which literally means *why don't*.

(140) ***c** u kam?*
Si you come?
Why don't you come along? [SMS]

(141) ***c** ujipange 2mlipe*
Si ujipange tumulipe?
Why don't you get organised then we pay him at the end of the month. [SMS]

k [kei]

Letter k is used in place of okay. Another possibility would have been to replace the word key, but it was not found in the data.

(142) *hp kila mtu is **k.***
I hope that everyone is **okay**. [IM]

m [em]

Letter *m* is used as a replacement for, *I'm* which is a contraction of *I am*. Note that this could also be interpreted as a replacement of *am,* leaving out the pronoun from *I am*.

(143) ***M** ok.*
I'm ok. [SMS, Email, IM]

(144) ***m** of 2 de vilage*
I'm off to the village. [SMS]

(145) *Hi 2 brok 2 kol bt **m** ok*
Hi, I'm too broke to call but **I'm** ok. [SMS]

There were no illustrations of its replacement use in Kiswahili.

n [en]
n is used as a replacement for *and, an* and *any*. The following are illustrations of this:

and [ən(d)]

(146) *I stil lov u **n** no 1wil eva change dat*
I still love you **and** no one will ever change that. [SMS]

(147) *They ok **n** say hi*
They are ok **and** say hi. [IM]

(148) *hi. niko hom **n** kila ki2 is fyn*
Hi niko home **and** kila kitu is fine.
Hi I'm home **and** everything is fine. [Email]

an [an]

(149) *dint get **n** answer*
Didn't get **an** answer. [IM]

(150) ***n** angry man is a hungry man*
An angry man is a hungry man. [SNS]

any [ɛni]
The letter *n* is used as a pronouncible approximation of the word *any* as shown in the following examples.

(151) *Do u hav **n** class kesho?*
Do you have **any** class tomorrow? [IM]

(152) *Il find out if Iv **n***
I'll find out if I have **any**. [IM]

(153) ***n** job is ok 4 me*
Any job is ok for me. [Email]

q [kju]
q is used in English to replace *queue and cu(kyu)* in *cute*.

(154) *wuuuui hiyo **q** ya regstr ilikuwa mrefu!!!*
Oh dear! the registration **queue** was too long! [SMS]

(155) *Thats **q**t...*
That is **cu**te.... [IM]

Interestingly *q* in illustration (156) is used to replace the Kiswahili *kiu (kiuno)* and *ki (kinywaji)* in (157) despite the fact that the letter q does not exist in the Kiswahili alphabet. This means that one has to read the word with the English q sound in order to make out the meaning. This construction can be attributed to show-off or competition. The youths try to outdo one another in inventing new forms.

(156) *Huyo manzi amebeba **QNO** ;-)))*
Huyo msichana amebeba **kiu**no.
That girl has a big ass. [SNS]

(157) *Hope uta-unleash **qy**nywaji*
Hope uta- unleash **ki**nywaji
I hope that you will provide a drink. [IM]

r [ar]

This is used to replace the auxiliary verb *are*.

(158) *HI man, we **r** in kk kamataring buzz*
Hi man, we **are** in Kakamega drinking. (Kakamega is a city in Western Kenya.) [SMS]

(159) ***R** u now usd to the hostel?*
Are you now used to the hostel? [Email]

(160) *m ok n op u **r** 2.*
I'm ok and hope you **are** to. [SMS]

s [es]

This is used to replace the auxiliary verb *is/'s*. For example,

(161) *m helping out a frnd in a fix. Kesho **s** k tho.*
I'm helping out a friend in a fix. Tomorrow **is** okay though. [SMS]

t [ti]

The data presented the following instance where *t* is used to re-

place tea.

(162) *tel her 2 weka sm **t** we r kujaring in ful swing*
Tell her to prepare some **tea**. We are coming in full swing. [SMS]

x [eks]

Some of the illustrations with *x,* e.g. *xmas (christmas)* and *x (kiss)* have a standard orthographic symbolic usage that existed before CMC and has been interpolated in it. It is worth noting that these illustrations only use the letter *x* to approximate the pronounciation of the intended sound. A similar use of *x* related to the word *xmas* is used to replace the name *Chris.* It is a relatively new manifestation of creativity. In illustration (169), it is interesting that *x* has been used to represent *ki's* in the term *Kibaki's*. This is clearly a show of creativity amongst youths.

x has been used to replace *εx, chris*, *ks* and *ki's*

ex

(163) *that's his **x** don't invite her!!*
That is his ex. Don't invite her! [Email]

(164) *ati she's an **x** nun;-))*
That she's an ex nun. [IM]

Chris

(165) *imagin **x** ameadmitiwa…*
Imagine **Chris** has been admitted into hospital. [IM]

The replacement of *x* for */ks/* in the following illustrations is another such example.

(166) ***thanx***
Thanks. [IM, SMS]

(167) *it **stix** sana*
sticks
It lasts for long. [Email]

(168) *his **2chix** r evn worse of!!* [SNS]
tu+chicks (*tu*-Kiswahili diminutive prefix)

His petite girlfriends are even worse off.

In illustration (166)-(168), x is used to incorporate the final *k* and *s* sounds in the words thanks, sticks and chicks. Both sticks and chicks are in Engsh and are used to mean spending a long time and girlfriends respectively.

ki's

(169) *am sprisd he stil has the audacity to criticize **kibax** govt*
I'm surprised that he still has the audacity to criticise **Kibaki's** government. (Kibaki is Kenya's 3rd president). [SNS]

y [why]

y is used to replace why.

(170) ***y** r u so gloomy nowadays?*
Why are you so gloomy nowadays? [IM]

(171) *she s so gone kwani **Y** did u mk ha wait tht lng?*
She's gone. **Why** did you make her wait that long? [SMS]

(172) *If luv's so nyce tel me **y** it hurts 2.....*
If love is so nice, tell me why it hurts to… [SNS]

The absence of *w* and *z* is conspicuous. These letters do not have an equivalent or near equivalent homophone in both Kiswahili and English and thus cannot be easily substituted.

Others

A further example of pronounceable consonant letters worth mentioning is the use of *vp* for the word *vipi* in Kiswahili. The combination of these two letters leads to a standard abbreviation of Vice President (VP). Yet in this context it is used to represent the word *vipi (how)* in Kiswahili.

(173) *muziki huu nikama wa kikisii au **vp***
vipi
This appears like Kisii music or what do you think? (Kisii is one of the Kenyan indigenous vernacular languages.) [SNS]

It is important to note that the use of some of the pronounceable letters that are used to represent words may be operating on the

lexical compression rationale where initial letters, especially vowels are clipped from the word. Examples include the use of *l* and *t* in the following illustrations:

(174) *Sory i thot **I** mek it*
Sorry i thought **I'll** make it. [SMS]

(175) ***t** was already spoilt just lyk that:-((*
It was already spoilt just like that (displeasure). [IM]

I have discussed these occurrences as a form of lexical compression with exclusive consonant spelling in section 4.5.2.

4.2.3. Conclusion

Pronounceable letters are a very common phenomenon in CMC. The general occurrence of the pronounceable letters in all the genres is influenced by the least effort principle. They are mainly used in SMS. This is associated with the mode limitation and the long term usage which has made their use frequent in SMS. Informality and peer identity also play a role in popularising the use of the pronounceable letters. In addition to these, for IM the main possible influence is rapidity. IM registers the second highest count. Email and SNS register relatively lower total counts which can be explained by the higher level of formality, lack of rapidity and lack of mode limitation for Email, and the public nature and lack of rapidity for SNS. Table 15 is a summary of the letter symbols used to represent words and sounds.

In terms of individual letters, SMS makes the most use of all the letters except *b* and *x* which occur the most in Email although by a very low margin. The reason for this is not clear. The findings reveal *u* as the most commonly used pronounceable vowel while *n* and *r* are the most popular consonants.

This is explained to the regular appearance of the words represented by these letters in communication contexts.

Table 15: Summary of Pronounceable Letters

Letter	Representation in English	Representation in Kiswahili
u	you	-
b	be	bi-bi
c	see	si
l	I'll, will	-
m	I'm/am	-
n	and, an, any	-
q	queue, cu,	kiu, ki
r	are	-
t	tea, it	-
x	ex	k(i)s
y	why	-

For example, the pronoun *you (u)* is used widely in different contexts e.g. as a personal pronoun in both singular and plural and as both subject and object. It can also be used as a generic pronoun. The different uses give an advantage over the other letters.

4.3. Pronounceable Numericals

2dy @ 8 sm1 shud b thea 2 c u b4 u go...2ko pa1 ;-)
Today at 8 someone should be there to see you before you go. Tuko pamoja (we are together).

The use of numeral symbols (numericals) in texts is considered as a major pointer to CMC language. It commonly occurs in SMS. Bodomo (2009:24) regards number homophones as a form of shortening or acronymy at the phonological level. As explained in 4.2, CMC users form shortenings by replacing a longer combination of letters or words by shorter combinations and words which are either the same as or near to the pronunciations of the longer words (Bodomo 2009:70). In the case of pronounceable numericals, this replacement is done by numericals to represent words or parts of words.

The data in table 16 reveals percentage use of each numerical based on the analysed data. From the discussion on variables, the expected results were that all the genres would register high

counts of pronounceable numericals because of the least effort principle. It was expected that IM, SMS and SNS would have the highest count because of the additional features of rapidity for IM and mode limitation for SMS and SNS.

Table 16: Use of Pronounceable Numericals

Usage of Pronounceable Numericals per Genre										
Genre	**No. 1**	**No. 2**	**No. 3**	**No. 4**	**No. 5**	**No. 6**	**No. 7**	**No. 8**	**No. 9**	**Total**
IM	43	84	6	60	4	2	1	24	33	**257**
Email	26	49	3	56	2	1	0	17	14	**168**
SMS	62	164	14	73	3	1	2	31	51	**401**
SNS	17	32	3	41	2	0	1	23	12	**131**
Total	**148**	**329**	**26**	**230**	**11**	**4**	**4**	**95**	**110**	**957**

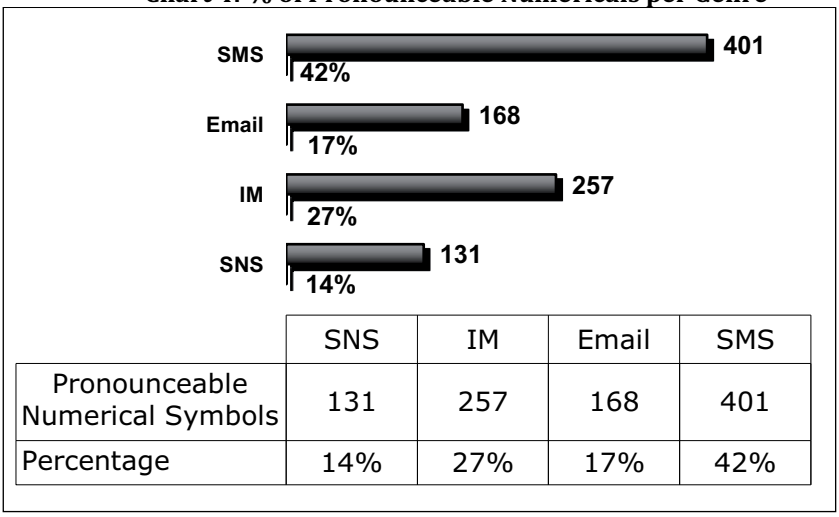
The results show a total of 957 cases of numericals. As expected, SMS uses the most numericals at 42% because of least effort and mode limitation. For example, typing the word *for* in an SMS takes 9 key taps in comparison with typing the equivalent numerical *4* which requires only one slightly long press on the key or six individual key taps. In addition, the use of the numerical 4 is economical on space because it only uses one character while the word *for* uses three characters.

Next to SMS, IM has 27% which stresses the role of rapidity in the use of pronounceable numbers. Email and SNS are at 17% and 14% respectively. This ranking is generally comparable to Frehner (2008) whose counts on numericals resulted to Emails at 5.87 while SMS had 68.48 per 1000 words. This roughly translates to Emails at 8% and SMS at 92% respectively. All the same, Frehner's range between the SMS and Email count is higher than mine which shows that pronounceable numericals are penetrating into Emails at a fast rate at least in my data.

Email was expected to have the least count because of its more formal nature, yet in my results it has a higher count than SNS.

The reasons for this are that the space allotted for SNS under mode limitation is relatively high thus there is not much pressure to use numericals to shorten the message. In contrast to Email and the other genres, SNS is less personal and is posted publicly and mostly to strangers. The higher degree of impersonality and publicity may explain why it does not apply as much pronounceable letters and numbers (cf. 4.2) as Email in order to be understood easily by everyone.

Chart 4: % of Pronounceable Numericals per Genre

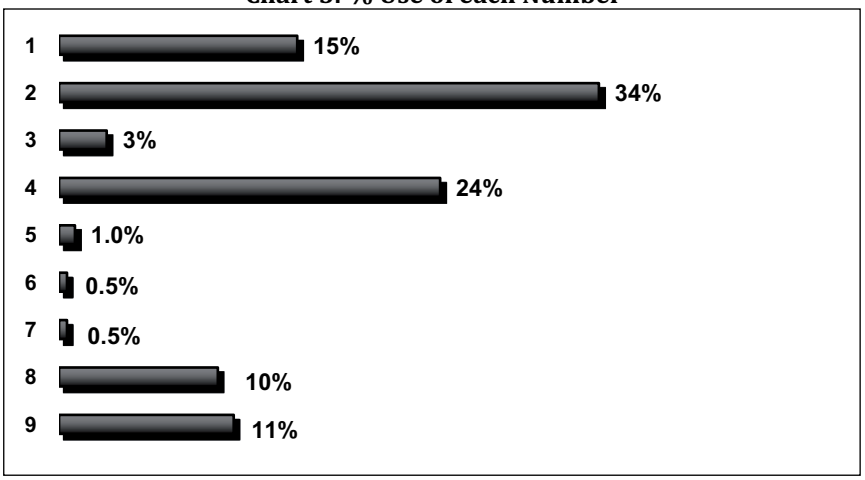


	SNS	IM	Email	SMS
Pronounceable Numerical Symbols	131	257	168	401
Percentage	14%	27%	17%	42%

The use of individual numericals is summarised in chart 5. Number *2* is the most frequently used in my data and *4* is the second most used. One of the facts explaining this is that these numericals can be used to replace independent words which serve as common prepositions *to(2)* and *for(4)*. These prepositions have high chances of being used in messages. Another explanation is that the numericals *2* and *4* are composed of a single syllable (CV) which enables them to be easily integrated in the Kiswahili words.

Numericals *6*, *7* and *5* are the least used. *1*, *2* and *9* are popular in SMS and IM. *4* and *8* are averagely used in all genres.

Chart 5: % Use of each Number



This section discusses the general use of numericals in the CMC corpus of data. The data reveals that numericals in CMC are used in three main ways as illustrated in cases a, b and c. I will now present these cases but exclude case (a) which is the standard use of nominals and ordinals. Apart from the case in (c), the data shows cases like in (b) where numericals are used for part of a word.

(a) Numericals are used to denote the number count or the time e.g.

(176) *kulikuwa na only **3** guys*
There were only 3 guys. [IM]

(177) *it was at **6*** [IM]

(178) *hey, u'r just de **1**. thnks sana*
Hey you are just the **one**. Thanks a lot. [SMS]

This use will not be discussed since it is the expected standard use of numericals.

(b) Numericals are used as near homophones where they stand in for parts of a word like in;

(179) *He **1td** his bk bck **b4** 12*
He **wanted** his book back before 12. [SMS]

(c) Numericals are used as homophones where they stand in for a complete homophonic word like *2(to)* in illustration (180).
(180) *I'l go **2** tao*
I'll go **to** town. [IM]

4.3.1. Bilingual use of Numericals

In the corpus data, the numericals are used as homophones in both English and Kiswahili for example number *1* may represent *one* or *moja* (cf. illustrations (183) and (184). Table 17 presents the Kiswahili and English names of the numerals 0-10 and their homophones. The dashes (-) show the numerals not found in the data as homophones. The homophones are given as spelling replacements and not in their phonetic form. As shown in column 3, with number 2 (*tu*) and 4 (*fo*), some English homophones are used within Kiswahili texts but this use is not reciprocated by the Kiswahili homophones in English texts.

Table 17: Kiswahili and English Homophonic uses for Numerals

No.	Kiswahili	Homophones used in Sw.	English	Homophones used in Eng
0	sifuri, sufuri	-	zero, nil, none	-
1	moja	moja	one	one, wan
2	mbili	mbili, tu	two	to, too
3	tatu	tatu	three	-
4	nne	- fo	four	for, fore, pho
5	tano	tano	five	-
6	sita	sita	six	-
7	saba	-	seven	-
8	nane	-	eight	ei
9	tisa	-	nine	ni, ny,
10	kumi	-	ten	ten

I now will discuss the use of each of the numerals as interpreted from the data.

Use of 1

The numeral denoting *one* is used in several ways in the data as

shown.

a) It is used in place of part of a word that contains the word *one* or *moja* e.g. English pronouns like. In all the instances from *(181)*-*(185*), the numeral *1*, carries the meaning of the word *one*.

(181) ***no1*** *knowz*
No **one** knows. [Email]

(182) ***sm1's*** *phone is now included in the list of 1st appearances!!*
*some**one's*** [SNS]

(183) ***evry1*** *is talking bout it*
Every**one** is talking about it. [IM]

(184) *lets mt @ 5 n' go* ***pa1***
pa**moja**
Let us meet at 5 and go together. [SMS]

(185) *Si unibuyie* ***ka1*** [SMS]
ka**moja**
Please buy me **one** drink. (Uni-buy-ie is a codemix between Kiswahili & English).

It seems that *moja* in Kiswahili is so far not used to replace the sound *[moja]* in a word other than where it means 'one'.

b) It is also used in place of part of a word that contains the form *[wʌn].* Refer also to example (179).

(186) *I* ***1nd****a y???* [SMS]
I **won**der why?

(187) ***1****dia my dear*
Wandia my dear. (Wandia (*1dia*) is a female's name in the Kamba language.) [SMS]

Use of 2

As pointed out earlier, 2 is used as *two* and *mbili* and additionally as [tu] for (Sw *tu*, Eng *to* and *too*). First I discuss the usage of 2 for [tu] in Kiswahili (a-d) and then *to* in English (e-f) and finally discuss the use of 2 for [mbili] in Kiswahili (h).

a) It is used as part of words containing the Kiswahili pronoun *tu-* 'we'.

(188) *uga!! mambo? **2kopoa 2nakumis 2** much*
tukopoa **tu**nakumis too much
Sw. tu Sw. tu
Hi, we are fine although missing you too much. [SMS]

(189) *ok **u2pate** digz sa 8*
u**tu**pate
Ok, meet us at home at 14.00. [SMS]

(190) *so **2kutane** wea*
***tu**kutane*
So where do we meet? [IM]

b) It is used in place of *tu* which forms parts of other words (not as a pronoun).

(191) *Ill fika **ki2** arnd 4 cant wait*
I'll fika ki**tu** around 4 cant wait
I'll arrive at around 4. I can't wait. [SMS]

c) It is used for the Kiswahili particle *tu* which means 'just'.

(192) *sawa **2***
*sawa **tu***
It is okay. [IM]

(193) ***2ko 2***
tu**ko **tu
We are just around. [SMS]

(194) *Unaniboo **2** sana*
unaniboo **tu** sana
You really bore me. [SNS]

d) It is used for the Kiswahili plural diminutive marker

(195) *She took all her **2kids** with her!*
She took all her **tukids** with her
She took all her little kids with her. [Email]

e) It is used as part of English words containing the form [tu], e.g. the orthographic *to* in English time expressions such as *to-night*.

(196) *I'l go **2nyt***
tonight
I'll go tonight. [IM]

(197) *pls read **2days** nation*
today's
Please read today's Daily Nation newspaper. **[**SMS]

(198) *Wil b kamin **2mro*** [IM]
Will be coming **to**morrow.

Another interesting use is to replace the form [tu] word internally as in (199). Interestingly the numerical appears syllable internally even though it only forms part of the onset.

(199) *bk 4 me a bars**2**l*
Book for me a bars**too**l. [SMS]

f) *2* is also used in place of the English preposition *to* and the adverb *too.*

(200) *we op **2** visit u guys soon.*
We hope **to** visit you guys soon.
[Email]

(201) *he's now **2** much!!!*
too
He's now too much! [IM]

(202) *I miss u **2** lv*
I miss you **too** love. [SMS]

The numerical *2* can also be repeated to achieve stress as *too* in (203) and (204*)*.

(203) *Wow! ua sis ni kasupuu.She looks **2'2'2'** pretty.*
Wow! your sister is pretty. She looks **too, too, too** pretty. [Email]

(204) *Its **222222222222222** much!!!!!!*
too, too, too, too, too... [IM]

g) *2* can also be used as part of Kiswahili words containing *mbili*

(205) *they ki2ad the fuel*
*they ki**mbili**ad the fuel!!!*
They rushed for the fuel. [SNS]

Case (205) is not easy to interpret because the sequence *mbili* is in the middle of the root *ki**mbili**a* 'to run'.

One of the most intriguing use of the numerical *2* is in what I call the '*numerical codeswitching*' illustration in (206).

(206) *jinga kama **22***
As stupid as a monkey. [SMS]

Where the initial *2* is to be pronounced in English as *two* [*tu*]. Note that unlike the Kiswahili *tu* that signifies the pronoun *we*, in this case the *2* is used as an initial part of a noun. The final 2 is to be pronounced as the Kiswahili *2* which is *mbili*. When the two 2s are combined, then the end result is the Kiswahili word for a type of monkey.

tu + mbili = tumbili
jinga kama 2 2
foolish like a monkey.

However, if there was spacing between the twos then it would have also been read as [tu mbili] 'two two's (*2* two's)'. This construction leads one to ponder on what would be the case when the writer wants to construct such a formation about two monkeys in this code. The sensible presumption then would be: *22 2*

(207) *22 2*
tumbili mbili (wawili)

if it is in Kiswahili or in codeswitch, it would be

(208) *2 22*
two tumbili

This can be described as an illustration of the numerical play that the symbol *2* brings about.

Use of 3

a) One of the most common uses of *3* as a CMC numerical is as part of words ma*three* or ma*tatu* '*3*'.

(209) *we pakd the **ma3** @ the cop stn..*
We parked the ma**tatu**/ma**three** at the police station. [SMS]

In (209), *ma3* can be interpreted as either *mathree* or *matatu* or clipped to *mat* - these words are synonyms for a public transport van with the former in Sheng and the latter in Kiswahili. These synonyms originate from the time of the introduction of the public transport vans. The initial name *matatu* is in Kiswahili and was coined from *tatu* which originated from the standard fare of 3 ten cent coins. Sheng later on manipulated the name into *mathree*.

(210) *don't u dare..ull NEVER get a nice job wit a **3u** on ua neck!!ama u vaa poloneks*
Don't you dare, you'll never get a nice job with a **tattoo** on your neck unless you always wear a polo neck. [SNS]

The use of *3u* for *tatoo* in illustration (210) is interesting because it not only uses the Kiswahili word for *3 (tatu)* amidst an English construction, it also includes a *u* vowel at the end of the word. In essence it would seem that the word is spelt as *tatu+u*. This additional *u* is added in order to lengthen the final vowel and achieve stress.

b) Used as a replacement for [thr]

(211) *I'm now **3u** comin soon*
I'm now **thr**ough. I'll be coming soon. [SMS]

(212) *we met **3ice***
We met thrice [IM]

(213) *they hid in the **ba3m***
They hid in the ba**thr**oom [SNS]

Illustration (213*)* is an instance where the numerical is across syllables. Note the vowel [u] is not represented.

Use of 4

The numerical 4 is only used in its English equivalent *four* in the data corpus. So far, the Kiswahili equivalent *nne* has not been used in numerical representation.

a) In the data, *4* has been used to replace *for*:

(214) *Hi msupuuuu what tym **4** 2nite?*
Hi pretty, what time **for** tonight. [SMS]

(215) *b thea **4** hm*
Be there **for** him. [IM]

It is used in place of part of a word that contains the forms *fo*, *for* and *pho*.

b) It is used to replace orthographic *fo:*

(216) *a heart tht loves iz **4**rvr young*
A heart that loves is **for**ever young. [Email]

Illustration (217) is the only instance found in the corpus where the numerical *4* is used in a Kiswahili expression to replace *fo*.

(217) *finally katoi kamelala **444** phew!!!*
fofofo
Finally the kid has fallen into deep sleep. Phew! [SNS]

It is worth pointing out that the expression *fofofo* is not a common lexical item in Kiswahili but a sound symbolic interjection.

c) It is used to replace the orthographic *for:*

(218) *i **4gt** 2 tel*
I **for**got to tell. [SMS]

(219) *this is one of my fave songs on it...kikuyu **4eva**...*
This is one of my favourite songs on it…Kikuyu **for**ever. [SNS]

(220) *respect 4 u will always b **4thkamin**.*
Respect for you will always be **for**thcoming. [Email]

(221) ***4wrd** ever backward neva*
Forward ever, backward never. [SNS]

(222) *Will **in4m** u*
I will in**for**m you. [IM]

d) It is used to replace the orthographic *fore:*

(223) *He left jst **b4** u arvd*
He left just be**fore** you arrived. [IM]

e) It is used to replace the orthographic *pho:*

(224) *his **4n** was stolen*
His **pho**ne was stolen. [SMS]

(225) *pls send me de **4toz***
Please send me the **pho**tos. [Email]

f) It is used to replace orthographic /f/:

(226) *had kam **4rm** kericho*
I had come **f**rom Kericho. [IM]

Use of 5

The numerical 5 is only used in a word to represent the number 5 as it is in Kiswahili *jumatano* 'Wednesday', literally the fifth day of the week (227). Therefore it replaces the Kiswahili word *tano* for 5 and it is strictly speaking not used as a pronounceable numerical.

a) It is used to replace *tano:*

(227) *kuja **juma5***
Juma**tano** (Wednesday)

Come on Wednesday. [SMS]

Use of 6

The data do not show any instances of using 6 to replace the English *six*. The numerical is used in Kiswahili to replace the sound sequence *sita*, '6' in Kiswahili in the form of

(228) ***6ki*** *ujinga*
sitaki
si-ta*k-i*
/1sg:neg-want-neg/
I don't want nonsense. [SMS]

This usage is remarkable because it replaces both a prefix and part of the following root at the same time. The sequence *si-ta-* for /1sg:neg-fut/ that is a common prefix sequence in verbs has not been reported to be replaced by 6. In addition, the stress on *sita (6)* is on *si* while it is on *ta* in *sitaki* in (228).

Use of 7

The numerical 7 has been used to replace the both the Kiswahili *saba* and the English *seven*.

a) It is used to replace *saba:*

(229) *ati? kwa* ***7bu*** *ya???*
saba*bu*
What? Why? [IM]

b) It is used to replace *seven*(i):

(230) *Hehehehe imagine* ***m7*** *semad ati the disputed Migingo Island is Kenyas bt the water round it is 4 UG!!!*
*Mu****seven****i (Name for Uganda's president)*
Imagine Mu**seven**i said that the disputed Migingo Island is in Kenya but the water surrounding it belongs to Uganda! [SNS]

A point of interest is that it seems that the stress of the word does not make any contribution to the usage of the number to replace a segment of the word. For example in (229), the stress is on *sa* in *saba*, while it is on *ba* in *sababu*. This implies that the use of nu-

mericals in CMC text involves both orthographic and pronunciation in order to make out the meaning.

Use of 8

Similar to numerical *6*, the data do not show the use of *8* to represent its Kiswahili form *nane*.

a) It is only used to replace [eit]

(231) *We are all well, kk was **gr8** we **8** alot!!;-)*
We are all well. Kakamega was gr**eat**. We **ate** a lot. (Kakamega is a town the Kenyan western province) [Email]

(232) *pls **w8** abit! Im kamin **str8** there, the ma3 was 2**l8***
*w**ait*** *str**aight*** *too l**ate***
Please wait a bit! I'm coming straight there. The mathree[38] was too late. [SMS]

Case (233) contains the repetitive use of 8 for stress.

(233) ***888888888888888888888888888888888888888****thnks* [SNS]

I am unaware of what the numeral 8 represents here. The guess is that it is a creative representation of the expression *great* with the first part *gr* omitted. In such a case the numerical 8 would stand in for [eit]. The repetition indicates stress to show a lot of gratitude. The message would read as *great thanks*.

Use of 9

The numerical *9* is only used in its English equivalent *nine* in the data. So far, the Kiswahili equivalents *tisa* and *kenda* have not been used in numerical representation. It is most popularly used in SMS valedictions to represent the first part of the word *night* in *good night*.

Here the illustration of 9 is used to replace [nai]

(234) ***g9t**, twas **9c** 2 mt u*

[38]. Cf. illustration (209) for an explanation of mathree.

Good**ni**ght it was **nic**e to meet you. [SMS]

Use of 10

The numerical *10* is only used in its English equivalent *ten*. Its Kiswahili form *kumi* is not represented in the data.

(235) *Id **4g10** the mtng*
I'd forgotten the meeting. [SMS]

(236) *lv me **10der,** lv me swt darl….*
Love me tender, love me sweet darling… [SNS]

4.3.2. The L33tspeak Phenomenon[39]

Leetspeak also known as 1337 or 31337 (Elite) is an alphabet used primarily in CMC. It has been described as a written-only dialect of English. It uses various combinations of numericals and other non letter symbols to replace Latin letters. It is a standardised use of numerical substitution for alphabetical letters that closely resemble the form of character of the numerical symbols (Mitchell 2006). This usage is independent of spoken language because it is solely based on the form of Latin alphabet symbols and Arabic numericals. It is a unique spelling system which is still mainly reliant on the keyboard and has much room for creativity. The data exhibits the usage of numbers to replace letters. As observed in the data, the main reason for the use of this is to disguise obscenities. In fact it is likely that leetspeak was probably developed as a way to outplay vulgarism-censoring filters. Another reason of its use in contexts without obscenities is perhaps just for fun. Examples include,

the replacement of *0* for *o:*

(237) *u f**00**1*
You f**oo**l. [SNS]

(238) *kweli alish**00**ta!!!*
ali-**shoot**-a (*shoota* is a vulgar Sheng word for *to pass wind*).
Did he really pass wind? [SNS]

[39] More on this phenomenon can be found at: http://www.bbc.co.uk/dna/h2g2/A787917

the replacement of *1* for *I:*

(239) *m**4**v**1** ya kuku*
M**a**v**i** – Obscene Kiswahili word for faecal matter.
Chicken shit. [SNS]

the replacement of *2* for *s:*

(240) *But for those with monster a**22**es* (a**ss**es). [SNS]

(241) ***2**enye*
senye-Vulgar Sheng word for female genitals. [SNS]

and the replacement of *3* for *E*

(242) *Ñíç**3** h**3**LpfûL fri**3**ñd*
Nic**e** h**e**lpful fri**e**nd. [SNS]

(243) *w**3**w**3**!!!*
wewe!
You. [SNS]

In illustration (242) and (243), the number *3* is used uniformly to replace *E* with no obscenities involved. These examples show the use of leetspeak for creativity and fun.

The replacement of *4* for *A:*

(244) *h**4**g**4***
haga (hAgA)-Vulgar Sheng word for the behind. [SNS]

(245) *m**4**t**4**k0*
matako- Obscene Kiswahili word for the behind. [SNS]

(246) *m**4**1**4**y**4***
*m**A**LAY**A***
prostitute- Obscene Kiswahili word for prostitute. [SNS]

The replacement of *5* for *s:*

(247) pu**55**y

Pussy-Vulgar word for a female genital organ.. [SNS]

(248) ***5****h3nz1 w3w3!!!*
Shenzi wewe! (Kiswahili)
You fool (stupid). [SNS]

The replacement of *6* for *b:*

(249) *haki ali nyam****6****a!!*
nyam**b**a –vulgar word for 'pass wind'.
I swear he passed wind! [SNS]

The replacement of *7* for *T* and *8* for *B:*

(250) ***7****0m****8****4*
T*om****B****A-Vulgar word describing a sexual relation.* [SNS]

The replacement of *9* for *g:*

(251) *di****9****ida*
*di****g****ida-Vulgar Sheng word for the behind*. [SNS]

It is worth noting that leetspeak in the data is mostly used in SNS. This seems related to the 'stranger effect' in the relationship among SNS users in the public discussion forums. These users do not seem embarrassed to introduce vulgar words amongst strangers but at the same time feel uncomfortable to present these obscenities freely and are compelled to mask them.

4.3.3. Conclusion

Similar to the pronounceable letters, the use of numericals in CMC is associated with SMS and mainly influenced by least effort, mode limitation and rapidity. Comparably to pronounceable letters, the findings reveal that SMS and IM make the most use of numericals while Email and SNS have the least. Numbers *2* and *4* are the most commonly used in Kenyan CMC. The result of number *2* as the most frequent is similar to Frehner's (2008:98) findings.

It is clear that the CMC numerical use discussed in this section is not only based on oral language but on orthography too. In some cases the numericals are simply substituted to the similar orthographic forms while in other cases they are substituted to resemble their pronunciation. The data shows that the meaning of numericals not only unveils itself in the use of numericals for their sounds in independent words e.g. *(2 for to)*, but also in parts of words like *(3u for through)*. Notably, in the data, there is no instance of Kiswahili pronunciation in English text[40] but there are quite a number of English number pronunciation cases in Kiswahili. The numericals are distributed in various positions e.g. at the beginning and end of root; as a morpheme inside a word and across morpheme boundaries and across syllables.

It appears that the higher the frequency of the use of a numerical for its sound in an independent word, e.g. *2* for *to* and *4* for *for,* the more likely it is that the numerical will be used as a substitute for its spelling and sound in other words.

The data also presents cases of the l33tspeak phenomena. This involves the use of numericals for letters based on their similarity in shape.

Pronounceable numericals and letters are comparable in their use in CMC. Both are used creatively either independently or as part of a word to represent a sound or spelling of the word. The main difference observed is that the letters are mostly used to represent independent words e.g. *b* for *be* while numericals not only represent independent words but also get attached to other parts of the words e.g. *7bu* for *sababu*. This kind of word internal attachment is only shown in letter *c, q* and *x*.

[40] Such a putative case would be *co4ct* for *co**nne**ct* with the Kiswahili word for 4, **nne**. Such cases do not occur.

4.4. Pronounceable Symbols

There are several cases in the data worth mentioning where symbols are used to represent words. The most common is the recurrent use of the symbol @ to represent the preposition *at.*

@ (at)

(252) *the mtng is on mon* ***@*** *2*
The meeting is on Monday **at** 14.00. [SMS]

(253) *having mingi chics* ***@*** *colle can stress u*
Having many girlfriends **at** college can stress you. [SNS]

(254) *I doubt if its a good idea* ***@*** *all...*
I doubt if it is a good idea **at** all.
[Email]

(255) ***@*** *her place*
At her place. [IM]

It is an interesting observation that the symbol @ is used in SMS. It involves six keystrokes in comparison to the combination of the letters *a* and *t* which involve only two keystrokes. This case shows that it is not only least effort which plays a role. It could be concluded that the use of the @ symbol in the SMS messages above serves to save space albeit by only one character. Yet, the messages are not close to the maximum limit thus there is still space for more characters. The conclusion is that in SMS, this symbol is mainly used as a style.

Other symbol occurrences include the ** (star)* and *# (ampersand)*. The ampersand is used to denote (sharp). These appear in several messages.

* (star)

(256) *being a * doesnt make you rich*
Being a **star** does not make you rich. [SNS]

(257) *Hehehehehe i saw ****!!*
(Laughter) I saw **stars**! (shock) [IM]

(258) *boxers have proved that you can punch your way to* ****dom***

Boxers have proved that you can punch your way to **stardom**. [SNS]

(sharp)

(259) *He's a # dude!!!*
He is a **sharp** dude [SNS]

(260) *You may be # but it doesn't mean you will sound #*
You may be **sharp** but it doesn't mean that you will sound **sharp**. [SNS]

4.5. Lexical Compression

Sm jms dn gt wt ts all 'bout ROFL☺!!!!!
Some jamaas (people) don't get what this is all about (rolling on the floor laughing)

A renowned practice in CMC that has been revealed in many CMC studies is the compression of words in order to present them in a shorter form. Hård af Segerstad (2002:201) regards this as a means to reduce text in messages. Frehner (2008:52) refers to the compression as lexical reduction and says that they abound in Emails and SMS. Crystal (2008) also discusses the compression and focuses on abbreviations in SMS. Bodomo (2009:78) generally terms lexical compression as spelling errors. He concludes that the reasons for the occurrences of spelling errors in chats may be attributed to the low awareness of the participants and the high consciousness of speed during message transmissions. He declares that no embarrassment is caused from this because there are no expectations of accuracy.

An interesting point that I would like to comment on is related to some views like Gaines (2008)[41] who claims that it seems that users put in more effort in consonant writing like in abbreviations, acronyms, contractions and other compression styles. According to him, first of all, these users need to figure out which alphabetical letters are representative, and then type them, the receiver on the other end also has to figure out what each set of letters represent before deciphering the message. Gaines suggests that it would be easier, and require less effort to just type in the whole

[41] http://randomstuffihate.wordpress.com/2008/03/06/people-who-write-words-in-shorthand-by-removing-vowels-or-using-acronyms-suck/

word or even use T9 (predictive text input) for SMS if it is available. Yet this seems not to be the issue from the results of this research. Basing on my results which show that compression is a popular practice in CMC, my view is that the inception of the compression is enjoyed as a form of creativity and it is not felt as the application of unnecessary effort. After acquisition, not much effort is applied in the composition and deciphering of the message. It is a practice that is close to automatic and does not require much effort once it is mastered. Further studies should be carried out in order to ascertain this.

This section discusses the features of CMC texts that compress lexemes and present them with fewer characters than in the standard spelling. The lexical compression features under discussion have been grouped under abbreviations, acronyms, consonant writing, contractions and other forms of character clipping.

Table 18: Distribution of Lexical Compressions

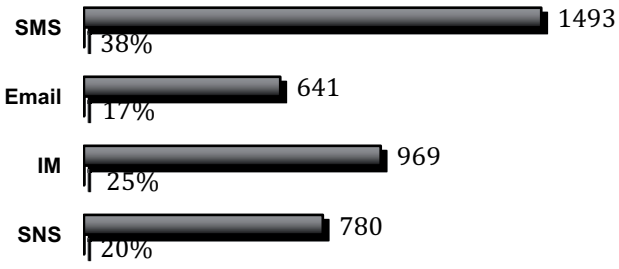
Genre	Abbreviations	Acronyms	Consonants	Contractions	Other forms of Lexical compression	Total
IM	193	209	162	213	192	**969**
Email	99	87	125	176	154	**641**
SMS	223	174	188	224	684	**1493**
SNS	167	129	156	186	142	**780**
Total	**682**	**599**	**631**	**799**	**1172**	**3883**

The detailed results are shown in table 18 and chart 6. The next sections discuss each type of lexical compression listed.

The data reveals that lexical compression is used evenly in all the genres because of least effort but abounds more in SMS at 38% due to the addition of mode limitation and the general adapted style which marks the SMS genre.

IM has 25% which can be attributed to rapidity. Owing to its relative degree of formality and availability of space,

Chart 6: % of Lexical Compressions per Genre



	SNS	IM	Email	SMS
No. of Lexical Compressions	780	969	641	1493
Percentage	20%	25%	17%	38%

Email has the least at 17%. SNS has 20% which is lower than SMS (38%) because its mode limitation has a higher allowance than that of SMS. Another considerable reason is that SNS is typed in a more relaxed manner like Email but it is more public which has an influence in reducing the level of lexical compression. The general impression from the results is that lexical compression is now a feature of all genres.

4.5.1. Abbreviations

An abbreviation is the shortening of a word by omitting parts of it, for example *Dr* for *doctor*, *Swa* for *Kiswahili*, *Eng* for *English* or *Engineering* etc. Abbreviations also known as clippings include all forms of word shortenings and clipping of which sections of the word are left out. I use this definition and concur with Rúa (2005) who explains that abbreviation is not only the deletion of letters at the end of a word, but includes forms that show letter deletion at the front, middle, or in different places in the same word. Accordingly, all forms that are shorter than the original word and preserve some of the original letters without adding extra letters that do not belong to the original word are abbreviations.

An example of a CMC message with abbreviations is

(261) ***coz*** *its fun 2 b a* ***celeb****?? n get comments like* ***urs****??!!*
Because it is fun to be a **celebrity**? and get comments like **yours**? [SNS]

In the message, *because* is abbreviated as *coz*, *celebrity* as *celeb*, and *yours* as *urs*. The use of abbreviations is common in CMC and is considered a culture in SMS. Some of the omissions or clippings leading to abbreviations are influenced by the pronunciation of the words like in the phonological spelling system.

The expected results of abbreviations were that SMS would have the highest count because of least effort and its additional feature of character limitation. IM was to be ranked second because of its rapidity. Email was expected to have a lower count due to its formal nature. SNS was expected to be average.

The results of abbreviations in the data are not different from the general results of lexical compressions. SMS has 33% followed by IM at 28%. This shows that the mode limitation in SMS plays a larger role in the creation of abbreviations than rapidity in IM. Abbreviations in SMS undoubtedly save space. An example is the abbreviation of the word *people* to *ppl* which reduces the number of characters from 6 to 3 or the word *important* (9 lettes) which is abbreviated to *impt* (4 letters). The percentage in SNS is average at 24% which is influenced by mode limitation and its public nature. Email on the other hand has the least count at 15% owing to its relatively high degree of formality with no influence from rapidity or mode limitation. Abbreviations in the corpus have been grouped as follows:

Clipping of initial vowel *e* suceeded by *x*
This involves the clipping of the initial vowel *e* which is succeeded by the letter *x* in words. The assumption is that the letter *x* represents the omitted vowel plus *x*. Illustrations of this include,

(262) *hes faced sm bad* ***xpriens***
He's faced some bad **experience.** [IM]

(263) *I* ***xpect*** *it kesho*

I **expect** it tomorrow**.** [SMS]

(264) *Kuna **xtras**?*
Are there any **extra.** [IM]

(265) *wens the **xam**???*
When is the **exam**? [IM]

Additionally in illustration (262), the word *some* is compressed to *sm*. In the word *experience*, the initial part *expe* is abbreviated as *xp*.

Clipping of the Glottal Fricative

One form of the abbreviations involves the omission of the glottal fricative *h* in the initial parts of words with the vowels as the ensuing sound as shown in examples (269)-(272).

(266) ***av** got it*
I **have** got it. [SMS]

(267) *kwani wat **apend?***
What **happened**? [IM]

(268) *the prinsi didn't even hint **ow** much.*
The principal didn't even hint **how** much. [Email]

(269) *2dy's my last dy in **otel**.*
Today is my last day in the **hotel**. [Email]

(270) ***op** u're ok*
I hope you are ok. [SMS]

The glottal fricative is also clipped in some Swahili words and prefixes. This effect is similar to the way that they are pronounced for example the *h* is often dropped in Kenyan Swahili like in the next illustrations:

(271) *Mbona **ujamshow?*** (Sw+Eng: ***h**ujamu-show*. The resulting word takes a Sheng meaning of '*told*').
Why haven't you told him? [Email]

(272) ***ebu** tel me???* (Sw: ***h**ebu* means *hey*)
Hey, tell me. [IM]

(273) ***anaitaji*** *help* (Sw: ***h****itaji* means *needs. Without the glottal fricative, the root left is itaji*)
He/she needs help. [SNS]

It is important to note that the clipping of the initial glottal fricative *h* in Kiswahili is only done in cases where it is considered as a dummy character and the meaning can be conveyed without it. It cannot be omitted in some cases like in the simple future tense where it is a prefix that serves as a negation marker. For example given the verb *lipa (pay)*, its negation form in the future tense would be,

h*-a-ta-lipa* (he/she will not pay)
neg-sbj-fut-vrb

The clipping of the glottal fricative which is the negation marker would result in the complete opposite meaning - *atalipa* (he/she will pay). Therefore CMC users do not omit the glottal fricative in such cases because it would result to the opposite of the meaning.

Clipping of the final part of the word.
Another form of abbreviation is the clipping of the final part of the word e.g. the word *prinsi* for *principal* in (268) or *celeb* for *celebrity* in (261). Notably this clipping of the final part of the word and leaving it with a stressed final vowel syllable is a type of truncation common in Sheng (Bosire 2009:83). This is also a common practice in CMC where the final part of words is clipped such that the word is left in its shortest recognizable form. Illustrations of this kind of final part clipping are highlighted in (274)-(276).

(274) *R u in* ***lib***
Are you in the **library**? [SMS]

(275) *utakam kwa* ***lec*** *(lecture)*
uta-come kwa lec
Will you come to the lecture? [SMS]

(276) *av nt bn 2 Internet 4eva thats y havnt* ***comm****.*
I have not been to the Internet forever, that's why Ive not **communicated** [Email]

Examples like *lib*, *prinsi* and *celeb* are also used in the same form in Engsh (cf.2.1.4).

Clipping of final *g* in verbs with progressive tense.
This is another form of CMC abbreviation that is based on the way the verbs are pronounced in Engsh without the final *g*. It is found in all the genres. Illustrations are in (277) – (280).

(277) *im **lengarin**[42] that saga*
I'm **ignoring** that saga. [Email}

(278) ***approachin** tao.*
I'm approaching town. [SMS]

(279) *Am **sendin**2u now*
I am **sending** to you now. [IM]

(280) *thnx 4 **calin***
Thanks for **calling**. [SMS]

The following illustrations involve the clipping of final *e* in past tense verbs:

(281) *he was **confirmd***
He was **confirmed.** [Email]

(282) ***blokd** my line sato til nw*
Blocked my line from Saturday until now. [IM]

(283) *how come uv **confusd** my no?*
How come you've **confused** my number? [SMS]

Clipping of the final *e* succeeding some consonants like *v, m* and *n*.

The final *e* is omitted from the spelling because the preceding consonant contains the intended sound. Illustrations of this occurrence with a preceding *v* are in (284)-(290).

---

[42] *lengarin* is an intraword codeswitching of Sheng *lenga (ignore)* & English suffix *–ing..*

(284) *wd b grt 2* ***hav*** *u wit us*
Would be great to **have** you with us. [SMS]

(285) *I got de msg, si u can* ***giv*** *us his no. ?*
I got the message, si you can **give** us his number? (in this illustration, s*i* is Kiswahili for ` *please)* [SMS]

(286) *Bn kold 2* ***solv*** *story wich I cant*
I've been called to **solve** a case which I can't. [SMS]

(287) *u wont* ***bliev****!!*
You won't **believe**! [IM]

(288) *kip ur* ***lov*** *brning 4eva*
Keep your **love** burning forever. [SNS]

(289) ***we'v*** *realy misd ya*
we've really missed you. [SMS]

(290) ***I'v*** *managed 2 get hold of my frnd*
I've really managed to get hold of my friend. [Email]

(291) *we'd* ***gon*** *2 heng wit de* ***hol*** *grp!!*
We had **gone** to dance with the **whole** group! [SMS]

(292) *utakuwa* ***hom*** *wen??*
When will you be **home**. [SMS]

Clipping of the vowel in the tense, subject and object marker syllable

The CMC data contains messages with Kiswahili verbs that currently show a trend of omitting vowels in the tense, subject and object marker syllable. The omitted vowels mainly occur between two consonants and are predictable from the context such that their absent does not interfere with comprehension. The following is a description of this feature.

Omission of the past tense marker vowel

The *i* vowel in the past tense marker *li* has been omitted in illustrations (293) - (294).

(293) *mu****l****present?*
mu-**li** + present

Sbj- Pt Vrb
Did you present? [SMS]

In illustration (294) the verb *tlkunywa* exhibits a double omission of two vowels in the subject marker *tu* and in the past tense marker *li*.

(294) ***tlk****unywa viexcess!*
tu-li + kunywa
Sbj- Pt Vrb
We drank excessively. [IM]

Similarly in (295) the verb *nilmcheki* has a double omission of the vowel *i* in the past tense marker *li*.

(295) *I think ni****lm****cheki online last wkend.*
ni-**li-mu** + check
Sbj-Pt-obj Sh Vrb (cheki – see)
I think I saw her online last weekend. [IM]

Omission of the present tense marker vowel

The *a* vowel in the present tense marker *na* has been omitted in illustrations (296) and (297).

(296) *raila a****n****kam kufungua show!!*
a-**na** + kam
Sbj-Prst +Eng Vrb (come)
Raila will come to open the show. [SMS]

(297) *wa****n****cheza na mimi!!!!*
wa-**na** + cheza
Sbj-Prst + Sw Vrb (cheza - play)
They are playing with me! (threat) [Email]

Omission of the perfect tense marker vowel

The *e* vowel in the perfect tense marker *me* has been omitted in illustrations (298) and (299):

(298) *a****m****sema aje*
a-**me** + sema
Sbj-Pfct + Sw Vrb (sema - say)
What did he say? [IM]

(299) *tu**m**confirm the deal*
tu-**me** + confirm
Sbj-Pfct + Vrb (confirm)
We have confirmed the deal [SMS]

Omission of the future tense marker vowel

The *a* vowel in the future tense marker *ta* has been omitted in the following illustration:

(300) *it looks like ni**t**chukua tu*
ni-**ta** + chukua
Sbj-Fut + Sw Vrb (chukua-take)
It looks like I will just take. [IM]

Omission of the subject marker vowel

The data also has illustration (301) and (302) which generally omit the vowel *i* in the 1st person singular subject marker *ni*. This vowel occurs between two consonants.

(301) ***n**ko na doo zako..utakamia?*
ni-ko
Sbj-Sw vrb (have)
I have your money. Will you come for it? [SMS]

(302) *ati **n**tafika huko kitu 5 hivi*
ni-ta + fika
Sbj-ft + Sw Vrb (fika - arrive)
I'm told that I will arrive there at about five. [SMS]

This omission of *i* from *ni* is likely to have originated from the regular omission of *i* before *n**i**na – nna*.

Omission of the object marker vowel

The omission of the object marker vowel examples include,

(303) *alnchkulia juzi*
a-li-**ni**-chukulia
Sbj-pt -obj+ Sw Vrb + prep (chukua + Prep – pick for)
He collected it for me the day before yesterday. [IM]

(304) *nimezbuy*
ni-me +**zi**+ buy
Sbj-pfct + obj Vrb (buy)

I have bought the things. [SMS]

Many of these vowel omissions seem related to the fact that the contexts of the omitted vowels can be easily predicted because no other vowel can fit in. For example in a present tense verb like *wanacheza*, it is only the vowel *a* (out of the 5 Swahili vowels) that can fit next to *n* to complete the syllable marking the present tense. This is the case for the rest of the verbs.

4.5.2. Exclusive Consonant Spelling

Consonant writing is the omission of vowels and exclusive use of consonants to represent a word. Bodomo (2009:124) calls this practice the shortening by omission of vowel indicating letters.

The findings are not very different from the findings in the general lexical compression results. They show SMS at 188 (30%), IM at 162 (26%), SNS at 156 (24%) and Email at 125 (20%). This range is very close in comparison to the other variables under lexical compression. It is exactly as the expected results which were that all the genres would register an average count of exclusive consonant use because of the general least effort principle in CMC. Email has a slightly lower count because of its more formal nature while SMS has the highest count because of its restricted mode limitation. Initially the consonant only style of spelling was associated with SMS. It was influenced by the character limitation. For example the compression of the word *goodnight* into exclusive consonants as *gdnt* not only lessens the keypad presses from 17 to 5 but also reduces the number of characters from 9 to 4. In fact this is according to Hård af Segerstad's (2002:257) findings where consonant writing only appeared in the SMS data. Based on my results, I can conclude that consonant writing has currently become a general CMC culture and is now used in all genres even in messages that are brief and without any of the character limit. Examples of this are in messages (305), (306) and (307).

(305) *m* ***hvn*** *my* ***lst cls***
I'm **having** my **last class.** [SMS]

(306) ***wd*** *b* ***grt***
Would be great. [IM]

(307) ***Wch*** *kcb* ***brnch****??*
Which KCB (Kenya Commercial Bank) **branch**? [SMS]

(308) *She gave* ***brth*** *2* ***ths*** *cute litl thing. Thank God the dad's genes* ***rn't vsbl****.*
She gave **birth** to **this** cute little thing. Thank God the dad's genes **aren't visible**. [IM]

(309) ***Slmz plz flsh*** *me if u* ***gt*** *dis* ***msg.thnk*** *u.*
Salaams (Greetings) **please flash** me if you **get** this **message**. **Thank** you. [SMS]

Slmz comes from Sheng/Engsh *saalamz* which in turn originates from the Kiswahili word *salamu* and means greetings.

In addition to the exclusive consonant writing, the corpus shows consistent cases of retaining the initial vowel in many words that begin with a vowel letter. This vowel is then succeeded by exclusive consonant letters. These retained initial vowels are important in the word identification. For example:

(310) *hebu we* ***orgnz*** *4 a lift...*
Hey let us **organise** for a lift (free car ride). [SMS]

(311) *the* ***aplctns*** *r rdy*
The **applications** are ready. [SMS]

(312) ***othws*** *m ok*
Otherwise, I'm okay. [Email]

Similar to the compression of *i'm* to *m* in illustration (312), some lexical forms are compressed into one consonant. Examples include,

Abbreviation of *I will*

The contraction of the phrase *I will (I'll)* is abbreviated to *l* as shown in the following illustrations.

(313) ***L*** *confirm tht l8r*
I'll confirm that later. [IM]

(314) *tel her* ***l*** *go nxt wk*
Tell her **I'll** go next week**.** [Email]

The illustrations also contain other forms of exclusive consonant writing like *tht (that) nxt (next)* and *wk (week)*.

Abbreviation of *It*

Similar to *l*, the pronoun *it* is compressed and abbreviated to *t* as follows:

(315) *Gv **t** 2 her.*
Give **it** to her. [SMS]

(316) ***t**shld be rdy in 4 dys.*
It should be ready in 4 days. [SMS]

additional forms of exclusive consonant writing in the illustrations include *gv (give) tshld (it should)* and *rdy (ready)*.

The main observation in consonant writing is that the consonants are used both as an orthographic and pronunciation clue. One has to process their pronunciation before being able to decipher the anticipated meaning. This makes the consonants representative such that they are able to reveal the intended word and meaning when pronounced.

4.5.3. Acronyms

An acronym is defined as a word formed from the initial letters of a name, or by combining initial letters or parts of phrases[43]. CMC acronyms are referred to by many as *geek speak* acronyms. According to Bodomo (2009:50), *BTW (by the way),* and *ASAP (as soon as possible)* are the most commonly used CMC acronyms.

From the results (cf. table 18) all the genres make use of acronyms because it employs least effort. IM at 35% makes the most use of acronyms because of its synchronicity and rapidity which makes the message to be as quick as possible yet loaded with meaning. The synchronic chat style of IM encourages the use of acronyms that quickly describe ongoing actions and situations. Examples of these include acronyms *like bbl (I will be back later)*, *bios (the boss*

[43] Refer to http://www.netlingo.com/acronyms.php for more information on acronyms.

or supervisor is over my shoulder looking) and *rtfl (I'm rolling on the floor laughing)*. Another reason is the fact that IM allows for easy access of the acronym definitions either by asking the sender directly or by using Internet search engines. This has resulted in the use of acronyms being generally regarded as a style of IM. SMS is second with 29% because of mode limitation which encourages users to use acronyms in order to save space. On the contrary, unlike in abbreviations, rapidity as an influence to the use of acronyms surpasses mode limitation. More acronyms are used in IM than abbreviations. The acronyms in IM not only serve to keep the dialogue fast but also as natural as possible. Acronyms are also used in SNS (22%) and Email at (14%).

Table 19: Distribution of Acronyms

Complete Phrase	Acronym	IM	Email	SMS	SNS	**Total**
laughing out loud	lol	52	12	39	33	**136**
oh my God	omg	39	15	27	23	**104**
by the way	btw	19	8	14	12	**53**
as soon as possible	asap	17	6	6	14	**43**
in my humble opinion	imho	6	1	2	26	**35**
be right back	brb	22	1	3	8	**34**
	Total	**155**	**43**	**91**	**116**	**405**

More data regarding the individual acronyms that are commonly used among the genres is in table 19. Note that the acronyms presented in the table are the most recurrent in the data. They all fall into the group that Bodomo (2009:46) refers to as 'more universal' acronyms. They are used in a standard way in CMC across the world.

The top two most recurrent acronyms describe reactions, i.e. humour (*lol*) and surprise (*omg*) in a compressed manner.

This is one technique adapted by CMC users to communicate feelings. Baron's (2004:29) gender based study on the use of IM among college students also ranks *lol* as the most frequently used

acronym. She explains that it is used as a phatic filler and has basically taken over the role of *ok, really* and *yeah* in conversations.

Emails apparently do not make use of acronyms in comparison to the other genres. This is due to their unlimited space, unhurried typing and the higher level of formality that is maintained. It can also be deduced from the data that the most recurrent acronyms appear more in IM and SNS. This implies that although these acronyms are a convention for IM, SNS has adapted them and they are easily understood in both genres. From the general results in table 18, the style of acronyms has also been adapted in SMS.

An interesting observation in the use of acronyms in the data is that both capital and small letters are used indiscriminately without any rules as to when to use only capitals or small letters. Examples of the acronyms used in the data are

(317) ***brb*** *(I'll be right back).*

X: Hi !
I hope you had a good weekend!
Y**:** how're you?
hey
ya
brb
X: How's everything?
Y: *gimme 2 min*
X: ok! Its ok if youre busy...was just saying a quick hi!!
Y: am back [IM]

(318) ***LMAO****! (**L**aughing **M**y **A**ss **O**ff*[44]*) Wah! i was on the floor chekaing*[45]

(319) *Ditto.* 😢*I had a break down* ***BTW....LOL***[46]
Ditto. I had a break down **B**y **T**he **W**ay... **L**aughing **O**ut **L**oud. [IM]

(320) *Pls let me know* ***asap*** *b4 its too dark*
Please let me know **A**s **S**oon **A**s **P**ossible before it is too dark. [SMS]

---

[44] LMAO can also mean "**L**eave **M**e **A**lone **O**k?"

[45] *Chekaing* involves intraword codeswitching between the Kiswahili lexeme *cheka* (laugh) and English progressive suffix *–ing.*

[46] *Lol* can also mean *lots/loads of love*, among others.

It is apparent from the data that most of the recurrent acronyms are universal. The few that are uniquely formed or adopted from the Kenyan context include the acronym illustration in (321) which is an example of an unconventional Kiswahili acronym.

(321) *imagine the guy wants **tkk*** (**T**oa **K**itu **K**idogo)
(remove something small) –bribe. [SMS]

The data also exhibits the use of some standard Kenyan acronyms which are not specific for CMC. This shows that CMC also uses conventional acronyms for example,

(322) *Hv u applied 4the **ECK** job yet* (**E**lectoral **C**ommission of **K**enya) [Email]

(323) *Hey cutie! C*[47] *u beep me wen u pata this text. U r so mteja n I don elewa y u hatad 2days **CAT**.* (**C**ontinuous **A**ssessment **T**est)
Hey cutie, please beep me when you get this SMS. You are unreachable and I don't understand why you missed today's CAT. [SMS]

4.5.4. Contractions

Contractions as defined by Bieswanger (2007) are combinations of two words that lead to a smaller number of characters than the spelling of the two words individually. Besides the assorted forms of lexical compression, the category of contractions is the most common phenomenon of lexical compression in CMC (cf. table 18). In spite of this, it is worth noting that unlike the standard contractions that make use of the apostrophe mark, most of the CMC contractions appear without the mark. The lack of the apostrophe mark in contractions is a general practice in CMC contractions. Further discussion on the use of the apostrophe mark in CMC is presented in 4.7.8. The findings on the use of contractions and the distribution of the apostrophe are summarised in table 20. As expected, all the genres have a high count of contractions due to the principles in the hypotheses presented. SMS has 28%, IM 26%, SNS 24% and Email 22%. The explanation to this is similar to the one on exclusive consonant writing in section 4.5.2.

[47] *c* is Swahili *si* whose meaning in this case is *why not.*

As already pointed out, the results reveal that owing to the least effort, rapidity and the mode limitation principles, majority of the common contractions do not contain an apostrophe to mark the contraction point. SMS has the least number of contractions with the apostrophe.

Table 20: Distribution of the Apostrophe in Contractions

Genre	Contractions with Apostrophe	Contractions without Apostrophe	**Total**
IM	89	124	**213**
Email	66	110	**176**
SMS	61	163	**224**
SNS	74	112	**186**
Total	**290**	**509**	**799**

This is because in some cases, the use of the apostrophe in SMS leads to more effort and time in accessing and scrolling through the list of symbols in order to select it. Even though it is easier to access the apostrophe on the computer keyboard for IM, Email and SNS, it is not used in most cases. The main explanation to why the apostrophe to mark contractions is not used in CMC messages is that it is considered obvious and that its presence would not add any valuable information to the contracted word. More on the use of the apostrophe in CMC is discussed in section 4.7.8.

Illustrations of contractions found in the data include

(324) ***y kudn't*** *she come??*
Why couldn't she come? [IM]

The contraction in (324) is also a form of phonological spelling with the initial *ku*.

(325) *Hi hpe* ***u'r*** *fyn.*
Hi, I hope you are fine. [SMS]

(326) ***t'was*** *a plot-thanks sana*
Sh Sw
It was very nice. Thanks a lot. [SMS]

The contraction in (326) has the apostrophe at a different place from the expected standard.

The contraction in (327) requires the context coupled with pronunciation for it to be understood.

(327) *wish u a **g'de***
Wish you a good day. [SMS]

Illustration (328) is a contraction that uses codeswitching, such that the stem of the word is in Kiswahili and the contraction form of *is ('s)* is in English. It is thus a standard contraction but on codeswitching.

(328) *Kila **kitu's** ok.*
*Kila **kitu is** (Sw: Kila kitu - everything)*
Everything is ok. [SNS]

In the data, the contraction of the auxiliary verb *I have* occurs in two general forms, either as *Iv* or *av* as follows:

(329) *As **iv** told u b4 i dont want any1 2 go thro wat i did.*
As **I've** told you before, I don't want anyone to go through what I did. [IM]

(330) *now **av** got 2 xplain again!*
Now **I've** got to explain again! [Email]

Instead of being used as an auxiliary verb, the contraction of *have* in (331) and (332) is about the main verb. This use is different from the traditional expression of the contracted auxiliary.

(331) *Just cald **2 've** a rainy chek on u*
I just called **to have** a rain check on you. [SMS]

(332) *tuk my kids 2 de show jana didn't **they'vefun**.*
I took my kids to the show yesterday, didn't **they have fun.** [IM]

Illustration (333) is unique because the contraction completely omits the auxiliary verb *will* and is only left with the personal pronoun and the main verb but is still marked by the contraction symbol.

(333) ***I'kol*** *her kesho*
I will call her tomorrow. [SMS]

4.5.5. Other Forms of Lexical Compression

In the current data, the other forms of lexical compression by indiscriminate character clipping results have SMS at 58%, IM 16%, Email 14% and SNS at 12%. This is similar to the general lexical compression results. The SMS high number of clippings is due to the combination of mode limitation, least effort and a high level of informality. The rest of the genres are similar in their lower results in clippings. It seems that in comparison to the other genres, SNS at 12% does not make much use of lexical compression in this category. Its public nature might contribute to this because indiscriminate character clipping may make words to be generally incomprehensible. The results from Email show that it is not conducive to the clipping of characters from words. At 16%, IM results indicate that although rapidity plays a comprehensively important role in all forms of lexical compression, it is not the main factor that leads to clipping. Mode limitation is the main determiner to the indiscriminate clipping of word characters.

The data contains many other random omissions whose only pattern is to barely leave the word in its barest recognised form either structurally or phonetically as shown next.

(334) *R u kamin 4 de* ***recoln*** *kesho? Op 2 c u*
Are you coming for the **recollection** tomorrow? I hope to see you there. [SMS]

(335) *We ok just coughd* ***usku***
usiku
We are okay. We just coughed at night. [SMS]

(336) *dei ya mazishi* ***2neza*** *feil kuishia*
tunaweza
On the burial day, we may fail to make it there. [IM]

4.5.6. Conclusion

This section has discussed the use of abbreviations, acronyms, exclusive consonant spelling and contractions as forms of lexical compression in CMC. The general results have all genres recording

counts in lexical compression because the principle of least effort plays a major role in message construction. Except for acronyms, SMS has the most counts in all the other forms of lexical compression discussed. This proves that the additional effects of mode limitation lead to the compression of words. IM leads in the use of its genre specific type of acronyms and is second in the other forms of lexical compression. This demonstrates that rapidity which is a common factor in IM plays an important role leading to lexical compression, especially in the use of acronyms. SNS and Email generally register lower counts. This is because rapidity and mode limitation are not factors in Emails. The public nature of SNS coupled with its more generous character limit than in SMS makes it to utilise less lexical compressions in general.

The goal of lexical compression in CMC is to leave the word or phrase in its shortest yet recognisable form. This is clearly seen in abbreviations, acronyms and in consonant spelling. In relation to informal language, the data also shows the use of contractions as a common practice in CMC. However, most of the contractions used do not have the apostrophe, or for some, the apostrophe is not put at the contraction point. More of this is discussed in section 4.7.8.

4.6. Relaxing Spelling Standards in CMC Texts

> *The mistakes that we make are governed, at least in part, by the technology we use to make them. The point is so obvious as to border on banal, but I still think it's worth considering for a moment.*[48]

Crystal (2001:111) explains that misspelling (spelling errors) in CMC whether intentional or not, are a natural feature of CMC text messages and "occur regardless of the educational background of the writer, in any situation where there is fast typing and a lack of editorial revision". Frehner (2008) claims that misspelling in CMC is largely due to the rapidity with which a message is composed. She explains that one tries to type as quickly as one thinks while knowing that everything that is keyed in can be deleted, modified or corrected easily. In fact pausing to correct spellings or rewrit-

[48] http://everything2.com/node/1970257

ing sentences only interrupts the typing and the thinking process. In contrast with Crystal (2001), as already pointed out, Bodomo (2009:78) attributes spelling errors to the low awareness of grammar of the senders who communicate with friends through IM. In my view, which is similar to Crystal (2001:111), the lack of grammatical awareness or intelligence does not feature much in CMC. The consensus among most researchers is that much of the spelling errors are influenced by the rapidity and least effort coupled with the informality of the whole CMC setting. Unconventional spellings are very common in CMC texts and yet remarkably, they are not treated as strictly and in a judgemental way like in more formal communication media.

Table 21: Distribution of Misspellings and Typos among Genres

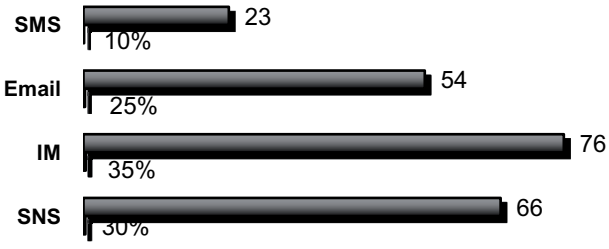
Genre	Misspelling	Typos	Total
SMS	13	10	**23**
Email	31	23	**54**
IM	50	26	**76**
SNS	29	37	**66**
Total	**123**	**96**	**219**

It is important to make the distinction between misspellings and typos. Misspellings are due to misunderstanding of how to spell a word. Typos on the other hand are usually the result of hitting the wrong keys (especially when in a hurry). Here I generalise everything as spelling errors since that is the end result. I also attempt to point out the typos and the misspellings in the illustrations.

The results are presented in table 21 and chart 7. The results are as expected with IM recording the highest number of misspellings due to the combination of a high level of rapidity, least effort and informality. SNS and Email rank next since majority of them are not proofread for spelling.

SMS involves more careful typing and proofreading because of mode limitation, it therefore has the least count of misspellings among the genres. IM, SNS and Email involve thinking while typing and the main concentration is in framing the thought thus in most cases, the spelling is not given a lot of attention.

Chart 7: % Distribution of Spelling Errors in CMC



	SNS	IM	Email	SMS
No. of Spelling Errors	66	76	54	23
Percentage	30%	35%	25%	10%

The possible reasons for spelling errors include word confusion, small keypads and rapidity.

- **Word Confusion**

Donovan (2009) explains that some words sound similar when pronounced out loud but have different spellings and meanings. Misspelling of these words is a clear indication that CMC texts and oral speech are closely correlated. The users type out the words as they would pronounce them orally. When the user is confronted with a word, he will pronounce the word orally (even though subtly) and type it out with the initial spelling that comes to mind, and this proves tricky for near homophones as demonstrated in the following illustrations.

(337) *The president should not threaten to **suck** them just do it!*
The president should not threaten to **sack** them he should just do it! [SNS]

(338) *m here 2 wish u best of **lack***
I am here to wish you best of **luck** [IM]

(339) *Will **flush** u. Its noisy sana hia*
Will **flash** you. It is noisy too much here [SMS]

(340) *Do us a favor, **nest** time*
Do us a favour, **next** time [Email]

- **Small Keypads/Boards**

Another cause for spelling errors known as typos in this case arises from the small keypads and keyboards. On keypads, these typos are mainly in form of a different character replacing an intended double letter spelling while on keyboards the typos are in form of repetition or transposing of characters. Typos could be attributed to human error coupled with small crammed key pads/boards combined with haste. Not only are the keypads small but they are also complicated to use. For example in mobile phones keypads, besides tapping a character once, it is not very clear on how long to hold onto it to produce another character because the same key is usually given for a minimum number of 3 characters.

Examples from the SMS keypad typos include

(341) *tek the **uqer** rd*
upper
Take the upper road. [SMS]

(342) *pls* pu5 *a st 4 me*
pull
Please pull a seat for me. [SMS]

In illustration (341) the pressing of the character *p* twice in quick succession or pressing it and holding on it a bit longer results in *q* instead of a double *p*. Similarly in (342) pressing *l* in quick succession or a bit too long results in the numeral *5* instead of double *l*.

Examples from the computer keyboard typos include

(343) *i agree with you on **transfffer** of power to parliament.*
I agree with you on **transfer** of power to parliament. [IM]

(344) *Lakini **whhere** is the mandatory Bible found in most hotel rooms? **hhe** he he he he*

But **where** is the mandatory Bible found in most hotel rooms? (Laughter). [SNS]

(345) *Omera, that one will **sssually** find himself in White House!*
Mr. that one will **surely** find himself in White house! (Omera is a Luo address word for Sir/Mr.) [Email]

Illustrations (344) and (345) are examples of both misspelling and typo together. In (345) not only has the sender made a mix up between s*urely* and *usually* but has also added a typo resulting from the repetition of the characters. Other instances of this consecutive repetition of character are not very clear whether they are intentional or typos.

I also found out that most intentional repetitions involve the use of at least 3 or more consecutive repetition of the same character to achieve stress. This involves both consonants and vowels mostly in the middle or at the end of a word. Repetition of the initial character is relatively uncommon because it produces a stuttering or stammering effect. The following are examples of how the word 'love' is used in different instances in the YouTube data.

(346a) ***looooveeeeeeeee** this song* [SNS]

(346b) *i **looooooooooooooveeeeeeee** dis song* [SNS]

(346c) *I **lllloooovvvveeee** uuuu guuuyzzz* [SNS]

(346d) *i **looove** this song.* [SNS]

(346e) *i **luvvvvvvvvv** the song* [SNS]

This character repetition is obviously intentional.

As already mentioned, the issue of small keyboards also leads to typos from transposing letters (hence the infamous *teh*) or adjacent keys being pressed by mistake. This could include not just letters, but *numb3rs* and *caPSLOCK* as well[49]. Examples include,

(347) *unaproble**n** gani msee*

[49] http://everything2.com/node/1970257

What is your problem dude? [SMS]

(348) ***VERU*** *CORRUPT animals*
Very corrupt animals. [Email]

(349) *When i c u litterin in tao i kindly do point it out...1 hindi* ***giu*** *tried 2 lenga me n i was with 4 pals...hakujua akiichukua aki. I hate litterers*
When I see you littering in town I kindly do point it out...one Indian **guy** tried to ignore me and I was with four pals...he didn't know how he picked it up. Gosh. I hate people who litter. [SNS]

In these illustrations, it can be interpreted that there is a mix-up in the keys when typing. The *n* and *m* keys are neighbouring on the keyboard, thus *n* was typed instead of *m* in example (347) while in (348), instead of typing *y*, the user typed *u*. In (349), instead of typing the word *guy*, the user typed *i* instead of *u* and then *u* instead of *y* and ended up with *giu*. All these are neighbouring characters on the keyboard and can easily lead to such mix-ups. Note that illustration (347) is from an SMS (keypad) while (348) and (349) are from Email and SNS (computer keyboard). Letter swapping typos occur on computer keyboards more often than on phone keypads probably because many users type SMS with not more than two fingers. This is unlike on a computer keyboard where more fingers are used thus increasing the chance of transposing letters.

- **Rapidity**

Speed also results in spelling errors. As hypothesised in the principle of rapid communication, on many occasions, users are caught in a time bubble and need to send the message very quickly. They therefore type in haste and have no time to revise messages before posting them. This is clearly seen in IM where in most cases after the text has been sent, there immediately follows a correction or an explanation of what is meant. This indicates that the sender only gets time to read the message when it is already on the receivers screen or after it has been sent. Extract (350) illustrates this. It is an IM between two friends X and Y.

(350) *X. Amazinnnnnnnnnnnnnnnng!! Love Naiarobi!!*
oooooops Nairobi
Y. hata ***mmi*** *nai damu;-)*
****mimi****

X. Amazing! I love Nairobi.
Y. Me too, Nairobi is in my blood ;-) (Nairobi is the capital city of Kenya.) [IM]

It is interesting to note that most IM users make corrections by enclosing them between stars ** or brackets () or [] in order to distinguish the correction text from the natural flow of the IM. Baron (2004:29) corpus of IM data also reveals the use of corrections immediately after posting the message. There is a probability that in IM this may be influenced by the speed between the posts whereby one realises a mistake only after the message has been sent and thus makes haste to correct it.

4.6.1. Missing Characters

It is worth acknowledging the existence of some spelling errors which occur from missing characters. However, it is very difficult to clearly distinguish them in CMC due to the fact they could easily be intentional as phonological spelling, ellipsis, acronyms, pronounceable letters etc. For example:

(351) *m @ stage frm ksm 2 buy hshold **stuf**.*
I'm at stage from Kisumu to buy household stuff. (Kisumu is a city in western Kenya, Nyanza province located on the shores of Lake Victoria.) [SMS]

(352) *I agree wit u cause kibaki **destroed** everythin*
I agree with you because Kibaki **destroyed** everything. [SNS]

In examples (351) and (352), it is difficult to rate the words *stuf* and *destroed* as misspellings or intentional.

Other misspellings just involve typing errors for example,

(353) *Plz tek the **sucurity** money*
Please take the **security** money. [SMS]

(354) ***Butb** whats wrong with selling maize to sudan if it fetches more money?*
But what's wrong with selling maize to Sudan if it fetches more money? [SNS]

(355) *asked **my** the judge*
Asked **by** the judge. [IM]

(356) *that was **TAXE.***
That was **TAX**. [SNS]

In these errors, there is no keyboard connection whatsoever between the keys. In (353) the writer might have anticipated the latter *u*, while in (354) he might have stayed longer on the *b* key typed earlier. Illustration (355) can be explained by inattentiveness resulting to a mix up between *my* which maybe a more frequent word compared to *by*. In (356) it is possible that the writer might have been misled by the word *taxes* thus resulting in the mix up.

4.6.2. Exclusion of Spaces

Spaces are gaps that are used to separate words in texts to facilitate reading and comprehension. The data reveals some instances of typing without spaces. For most cases in SMS this is done deliberately to save space. Its occurrence in the other genres which have no maximum character restriction is attributed to typing errors. It seems to be triggered by the fact that the typing speed is high coupled with lack of proofreading. This makes the messages to be sent without some necessary spacing. For example,

(357) *The **governmenthave** suddenly realized the need to take 'development' to this part of **Kenya.The** world have noted.* [SNS]

(358) *Ala...* *Enyewe we **havecharacters inthisworld**..... LMAO ... hah hah hah hah*
ala - expression of amazement, *enyewe* - actually [IM]

(359) *sharrap u fool! shenzi type. **thisthread** sayz new stuff, then u come up wit ol' skul shyt then **ukiambiwaunavimba***
ukiambiwa unavimba
Shut up you fool, uncouth type. This thread says new stuff, and then you come up with old stuff then when you are told you sulk. [SNS]

An exclusion of spaces style recurrent in SNS is similar to the traditional telegram with the only difference being that the STOP sign in SNS is a graphic dot or a comma. It seems deliberate and does not impede reading since the words are not completely joined but each is enclosed between dots or commas. It is a sign of relaxing spelling rules as in the following illustrations.

(360) *We understand your **frustration!!!!!we** ll never support Raila come what may...Take our son to the **hague,jail** him, do whatever you wanna do..but we ll stand by **him.Did** anyone hear him calling for mass action or **executions.We** are not buggers and you ll see us do it in the years to come..* [SNS]

(361) ***mambo.nimesota*** *videadly!!!*
Hi, I'm totally broke. (*sota* is the Sheng word for financially broke) [SMS]

(362) *There is no escape someone will have to pay for their role in the post election* ***violence,innocent*** *childrens blood will never let you sleep in peace..* [SNS]

Another form of exclusion of space in the marked words in the messages illustrated below is a style closely linked to how they are viewed or interpreted as one word or one compartment without a pause.

(363) ***bytheway****, the bday card hasnt fikad.*
By the way, the birthday card hasn't arrived. [Email]

(364) *sasa* ***howru?*** *we miss you sana*
Hi **how are you**? we miss you a lot. [SMS]

(365) *from da moment* ***imet*** *u gal*
From the moment **I met** you girl. [Email]

It is interesting that I did not come across any illustrations where the exclusion of space involves words in different languages.

4.6.3. Letter Repetition

Letter repetition is a mechanism used in the messages to achieve speech-like qualities like stress, tone and the general desired emotional impact in what Frehner (2008:56) refers to as emulated prosody spelling as a form of extra emphasis. The intended impact is similar to oral speech. This is not necessarily a case of writing in a phonetic way but rather in the standard way but with letter/word repetition to achieve emphasis.

Vowels, consonants and even in some cases words, and punctuation marks are repeated in messages in order to achieve emphasis. For example in message (366), the words *wacha wewe* have used vowel repetition in order to achieve emphasis.

(366) *waachaa weeeweee;-))*
Wacha wewe.
You stop it. [IM]

Letter repetition is used creatively as a way around the technological deficiency in the presentation of feelings especially stress. From the recorded results, SNS has the highest count of 41% followed closely by IM at 40%. SMS has the least count at 4% and Email at 15%. The reason for the high presence of letter repetition in SNS is linked to its discursive nature where participants feel compelled to stress their points and they do it by letter, capitalisation and punctuation repetition (cf. table 28 and 22). The synchronous nature of IM requires the users to emulate real life communication scenario, where they also use similar methods like in SNS to communicate feelings. Email has a lower count of letter repetition. This is contrary from Hård af Segerstad's (2002) findings where Email had no letter repetition. But the reason for this disparity is the fact that her Email corpus consisted of official Emails to authorities. SMS has the least count because of the mode limitation principle where it is deemed as using up a lot of space for repetition. Additionally for SMS, this practice also involves more effort in typing the extra repetitive characters. The following are examples which use vowel repetition,

(367) ***niceeeeeeeeee videooooooooooooo.....***
Nice video. [SNS]

(368) ***had soooooooooooo much fun***
Had so much fun. [IM]

(369) ***wooooooooooooooow***
Wow! [Email]

(370) **whaaaaaaaaat?????**
what? [IM]

(371) ***byeeee***
Bye. [SMS]

Similarly, the following examples display the repetition of consonants in order to achieve emphasis.

(372) ***fffffffffyyyn***
Fine. [Email]

(373) *it's* ***verrrrryyyyyy*** *good!!!!!*

It's very good! [SNS]

The data also reveals other instances of non-standard orthography related to representation of the oral reality which is to capture actions like in illustrations (374) and (375).

(374) ***mwahhhhhhhhhhh*** (Kiss)
[Email]

(375) ***hahahhahahaha*** (Laughter) [SNS]

4.6.4. Capitalisation

As already discussed under acronyms, capitalisation can be used in acronyms and other forms of abbreviations. In another turn, this section deals with the manifestation and uses of capitalisation in CMC which varies from the conventional use. Hård af Segerstad (2002:143) confirms that indeed capitalisation in CMC is a form of shouting or emphasis. Frehner (2008:56) also terms this as emulated prosody spelling and categorises it under unconventional spelling.

Excessive Capitalisation

I define excessive capitalisation as any capitalised letter that occurs adjacent to another capital letter in the same word or in adjacent words.

Table 22: Number of Excessively Capitalised Characters per Genre

Genre	Excessively capitalised characters
SMS	169
Email	127
IM	465
SNS	420
Total	1181

The data exhibits a total of 1,181 of excessively capitalised characters which occur in 218 messages. I begin by presenting in table 22, the results of the number of characters that appear as excessively capitalised.

Just as predicted, the results indicate that IM has the highest number of excessively capitalised characters. Next is SNS followed by SMS. Email has the least score. Note that these are counts of the number of characters with excessive capitalisation and are distinct from the number of messages that contain them presented in table 23 as excessively capitalised messages.

Exclusive Capitalisation

Exclusive capitalisation is described as the use of capital letters for the entire message. Table 24 presents results of messages that are exclusively typed in capital letters.

SNS has 61 exclusively capitalised messages, while SMS has 17. Email and IM have none. Despite the differences in the population group and the nature of the collected data, Hård af Segerstad's (2002) results also have Email without any cases of exclusive capitalisation. This is interesting in two ways with the first being that her population group which sent Emails was of an older generation than mine.

Table 23: Excessively Capitalised Messages per Genre

Genre	Excessively capitalised messages
SMS	33
Email	9
IM	30
SNS	146
Total	218

This age difference may have played a role in their Email presentation style. Secondly is the fact that her data consisted of Emails that were meant for the local government. This would likely result to the use of a very formal style.

Table 24: Exclusively Capitalised Messages per Genre

Genre	Exclusively Capitalised Messages
SMS	17
Email	0
IM	0
SNS	61
Total	**78**

The Emails in my data on the other hand are between youth peers who would have comfortably used exclusive capitalisation without any qualms and yet they did not. The absence of exclusive capitalisation shows that there is some sensitivity over its use in Email and IM messages. It is interesting that our results differ in IM where Hård af Segerstad's population group is similar to mine comprising youth peers. In her case, IM between youth peers also registers the use of exclusive capitalisation.

From the results on exclusive capitalisation summarised in table 24, it can be deduced that only the SNS genre makes persistent use of exclusive capitalisations.

Expressive Capitalisation

Since it is not very easy to single out intentionally expressive capitalisation in messages, my closest attempt is to subtract exclusive capitalisation from excessive capitalisation in messages and then regard the result as intentionally expressive capitalisation. The

result is presented in table 25, table 26 and chart 8. It is indeed possible that some of the exclusively capitalised messages are intentionally expressive but this is not readily clear because it could also be a case of having the caps lock key on accidentally.

Table 25: Message Capitalisation

Genre	Excessively Capitalised Messages	Exclusively Capitalised Messages	Expressively Capitalised Messages
SMS	33	17	16
Email	9	0	9
IM	30	0	30
SNS	146	61	85
Total	**218**	**78**	**140**

The interpretation of the results of Email and IM in table 25 is that although these genres contain cases of excessive capitalisation in many of their messages, none of them has any case of exclusive capitalisation.

The reason for the high presence of IM messages that make use of expressive capitalisation in IM is linked to its synchronous nature which requires the users to emulate real life communication scenario.

Therefore they resort to use capitalisation to communicate feelings. Similarly, SNS is characterised by its commentary/discursive nature. This makes it totally acceptable for users to show their strong opinions, emotions, stress and emphasis using capitalisation. SMS may also use capitalisation as a form of expressivity[50] but to a lesser extent as it is not conventional in the genre.

[50] Note that not all use of capitalization is meant for expressiveness. Some occurrences are simply not explainable.

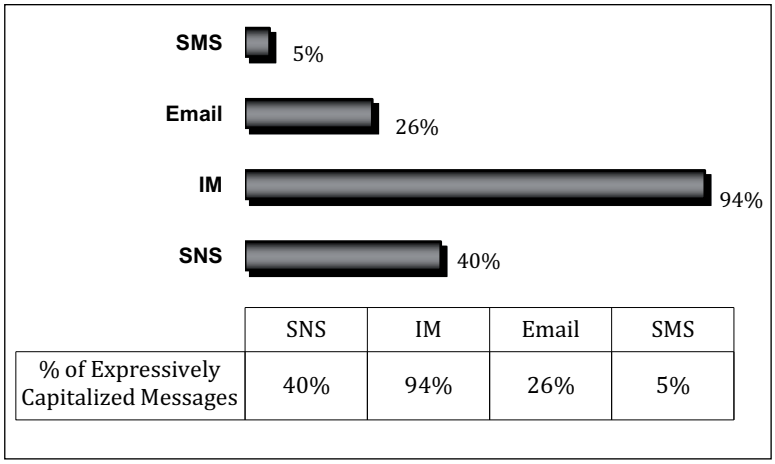
Table 26: Expressively Capitalised Messages per Genre

Genre	Expressively Capitalised Messages	% of Expressively Capitalised Messages
300 SMS	16	5%
35 Email	9	26%
32 IM	30	94%
212 SNS	85	40%
579 Total	140	24%

Besides being expressive, another possible reason for the infrequent use of capitalisation in SMS may be due to least effort, whereby the user is lazy to turn the capitalisation key on. Turning the capitalisation key on is considered not only as effort consuming but also as an interruption to the flow of the message composition and input process. Email has a lower count of apitalization which may be influenced by the fact that Email affords time and alternative choices like the use of bold, highlight, bullets or even the use of different colours of text to achieve stress.

My results contradict Frehner's (2008:95) findings that unlike in Email, many SMS users still regard wholly capitalised messages as legitimate. My population group regards the use of capitalisation as 'shouting' for emphasis. They mainly use it to achieve this end in SNS where it is prevalent. If it is used in a neutral Email, then it is considered as a typographical error. Similarly, if it appears in a neutral SMS, then the user is judged as not being technically savvy. In fact it is interesting that users in my population group would rather type the whole message in lower case as it is considered more legitimate than use of capital letters exclusively.

Chart 8: % of Expressively Capitalised Messages



	SNS	IM	Email	SMS
% of Expressively Capitalized Messages	40%	94%	26%	5%

The use of capitalisation in the messages is described hereafter. In the following examples, it is clear that the capital letters are used as a way to catch the reader's attention. They are meant to leap out right into the reader's vision and show different emotions as follows:

Disgust

(376) ***Y'DO U ACT LYK A KID???GROW UP!!!*** [SMS]

(377) *Waki sema chungwa/ndizi we should reply* ***UNGA.....KAZI....AFYA****....*
When they say Orange/Banana, we should reply, FLOUR, JOBS and HEALTH[51]. [SNS]

(378) ***JINGATYPE!! TSK.. GO FIGURE!!!***
Stupid! (click) Go figure! [SNS]

Happiness and Goodwill wishes

(379) ***CONGRATULATIONS****!!!!!!!!!!!And jubilations!!!!!!!!!!!!!!!* [SNS]

(380) ***U R 2 SWT 2 B 4GO1O C U SOOOOOOOOOOOON XXX***
You are too sweet to be forgotten, see you soon (kiss). [SMS]

[51] Orange/Banana were slogans for the ruling political parties that have formed a coalition government. The message writer is disgusted at the corruption and is suggesting that the public should challenge the government to provide, food, jobs and health services.

(381) *GO! GO! OBAMA WE ARE PRAYING FOR YOU SON OF OUR SOIL.....OBAMA JUU*
juu refers to *top* in Kiswahili. [SNS]

Surprise and Disbelief

(382) ***WACHA!!!*** *hahaha now wat? is smtng goin 2 cut?*
Really! Now what? is something going to take place? (Will you have an affair?) [IM]

(383) *haiyaaaaa!!!! waacha.....* ***NUMBER 1?????***
What! Really? Number one? [IM]

(384) ***ATI WAT???*** *kumbe its* ***TRUE****!!!*
That what? So it is true! [SMS]

Pleas

(385) *Sori i deleted ua Msg by mistk.* ***PLEEEE****z sms again??????.*
Sorry I deleted your message by mistake, please SMS again. [SMS]

(386) *haki* ***NISAIDIE*** *2 plz-****IM*** *2 2 2* ***BROKE!!!***
Please just help me -Im too broke. [SMS]

Despair and Solidarity

(387) ***OOOH NOOOO.... HAKI THATS HORRIBLE****:-((*
Oh no! That is horrible! [IM]

(388) ***WOOOOIYIE-****saa utado?*
Oh Gosh, what will you do now? [SMS]

(389) ***HEEEEEEEEEEEEELP!!!*** *I cant find my key-saidia me wit the spare*
Help! I cant find my key. Please give me the spare. [SMS]

Stressing a Point

(390) *Seriously! So, Im* ***DONE*** *with Java. Fake ass wanna-be's. Never will I be seen in one****! N.E.V.E.R!*** [SNS]

(391) *a heart tht loves iz 4rvr young n love itslf iz lyk heaven bt it cn hurt lyk* ***HELL!***
A heart that loves is forever young and love itself is like heaven but it can hurt like hell.
[SNS]

(392) *threatening to quit is* ***OK*** *but decamping the RVP supporters is* ***NOT RIGHT.***

Threatening to quit is ok but decamping the Rift Valley Province supporters is not right. [SNS]

(393) PLEASE, SEND ME KIBAKI'S PHOTO URGENTLY. WE ARE PLAYING CARDS AND HAVE NO JOKER! [SMS] [52]

Other cases of capitalisation appear not to be expressive language but rather a kind of style by the user, for example in (394):

(394) *da tym of* ***JuStIcE.If yOu kNow yOu Rn't amOng*** *da justice becuse ulipora mali ya wanai ....*
The time of justice. If you know you aren't among the justice because you stole the public's wealth. [SNS]

It seems that capitalisation in CMC is mostly used to denote emotions. In normal conversation, the words/phrases would normally be stressed, or spoken in a different tone i.e. higher tone (shout/scream) or a very low tone (disgust).

Absent Capitalisation

Another point noted on capitalisations is that many messages do not make use of a capital letter at the beginning of the message or on proper nouns. This is based on the least effort principle and the general relaxed spelling standards in CMC. In SMS, switching between capital letters and small letters is considered tedious and an interruption to the messaging process. Ideally, initial capitalisation in SMS is done automatically in T9 but many users have the T9 switched off.

In the results, SMS (33%) and IM (29%) have the highest number of initial absent capital letters as expected based on the least effort and rapidity principles for IM. SNS (23%) is almost average probably because of the balance between the official forums and the informal forums. At (15%) Email makes the least use of absent initial capitalisation.

[52] This was sent as a campaign message against one of the presidential contenders in Kenya's 2008 general elections.

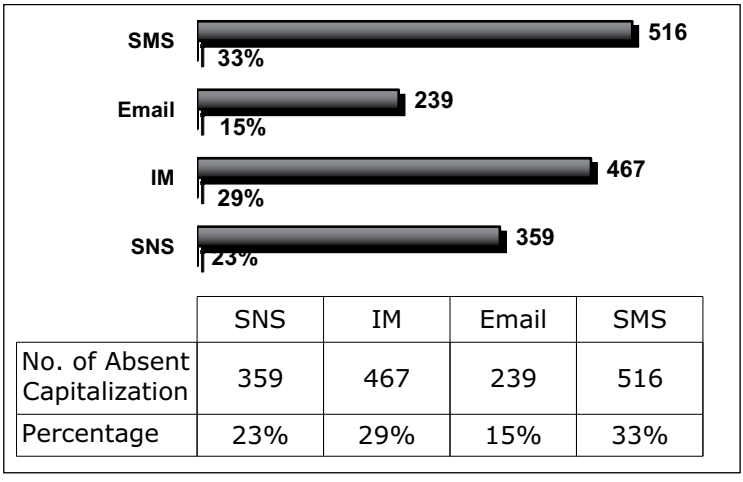
Table 27 Number of Absent Capitalisations per Genre

Genre	Absent Capitalisation
SMS	516
Email	239
IM	467
SNS	359
Total	**1581**

The results of the absence of the initial capital letter are similar to Bodomo's (2009:77) findings where capitalisation e.g. of the first letter of country names, of the first singular person pronoun and capitalisation of the first letter at the beginning of a sentence in Standard English are not strictly followed in CMC.

Although it was expected that this happens in all genres, the findings are that Email has the least missing capitalisations. In fact, 75% of Emails make use of capital letters at the beginning of the message and sentence. This is in relation to their more formal nature like in standard letter writing and the absence of modality and time restrictions.

Chart 9: % of Absent Capitalisations per genre



	SNS	IM	Email	SMS
No. of Absent Capitalization	359	467	239	516
Percentage	23%	29%	15%	33%

Capitalisation in Email is almost done automatically by users after a final punctuation mark. The complete results of the genres with missing capitalisation are shown in table 27 and chart 9.

Illustrations of messages with missing capital letters at the beginning of sentences include

(395) ***u****naduu nini leo afte?*
What are you doing this afternoon? [SMS]

(396) ***n****imefika :-)*
I've arrived. [SMS]

(397) *can **W**e **C**hat now???*
Can we chat now? [IM]

Interestingly like in (397), some of these messages without an initial capital letter may include other (un)necessary capitals in the rest of the message. It could possibly be typographical errors. It could be said to be deliberate in SMS since it takes effort and some concentration to compose this structure.

4.6.5. Variants of English

The different vernacular language backgrounds of Kenyans gives rise to variants of English in form of phonologically spelt texts. These variants are at a personal level and in many cases the user may not be conscious of the practice. This is caused by the fact that some Kenyan vernacular languages do not have particular speech sounds. This makes their speakers likely to mispronounce words containing these sounds and consequently type them that way.

Cheshire (1991) captures this accurately in her hypothesis that some Kenyan communities have phonetic features that distinguish their pronunciation of English. For example the central Bantu e.g. Kikuyu language lacks the alveolar lateral approximant /l/, and pronounces it as [r] while the Nilotic Luo lacks the post alveolar fricative /sh/ and pronounces it as [s]. Other occurrences include the insertion of the nasal /n/ before alveolar plosives by the Ki-

kuyu and the Meru and the unnecessary insertion and dropping the initial /h/ by the Kamba and also the pronouncing [d] instead of /t/ and [b] instead of /p/ and vice versa among the Luhyia. Therefore when speakers of these languages are confronted with challenging words incorporating these sounds, they easily write them in a phonetic way based on their native tongue. One can then decipher the user's vernacular language based on some unconventional spellings as illustrated below.

(398) *the mom should face neglegency* ***judges.***
The mom should face negligent **charges**. (Luhyia) [SNS]

(399) *anyway boyz r just full of* ***clap***
Anyway boys are just full of **crap.** (Kikuyu) [SNS]

(400) ***nisike*** *tafadhali!!! niokote tafadhali* ***nisa*** *anguka kwa* ***mapensi***
nishike tafadhali! Niokote tafadhali **nisha** anguka kwa **mapenzi**
Please catch me, pick me up, I've fallen in love. (Luo) [SNS]

(401) *I know wat ur* ***cabable*** *of*
I know what you are **capable** of. (Luhyia) [Email]

(402) ***himajine halitaka*** *kutake* ***iyo*** *risk*
Imagine alitaka ku-take hiyo risk
Imagine he wanted to take the risk. (Kamba) [IM]

4.6.6. Conclusion

IM has the most cases of misspelling based on its synchronous and rapid nature. SMS has relatively few cases since the keypad size only allows typing with fewer fingers compared to the keyboard which allows typing with more fingers and this raises the chances of errors.

In conclusion, the misspelling and exclusion of space variable clearly shows that in many cases, people do not revise text messages/posts before sending them. The sender simply clicks the button 'ok'. Of course this does not rule out exceptions, for example in some very long messages which go over the number of allowed characters per message e.g., some SMS are usually revised to check if there are unnecessary, inaccurate or too long words which could be discarded, replaced or shortened to release more

space and make the message more accurate, but this could also lead to intentional misspelling in order to save space.

One of the most striking observations on the findings of this variable is that for social CMC messages, the receivers of messages containing misspelled words perfectly understand them from the context. In fact the homophone confusion can be very amusing. Crystal (2001) and Frehner (2008) are in accord that in CMC, misspellings are excusable and are always attributed to the swiftness (rapidity) with which a message is composed. In this way, the receiver does not make a social judgment about the sender's intelligence on the basis of such data. This is a contrast with what would happen if someone wrote a traditional letter containing misspellings. Another observation is that for such misspelling, users are free with CMC and do not feel the pressure to be official or correct. In addition, the users are free with the receivers in that they are friends who 'know' each other (Lewis 1997). Therefore there would be less chance of paying much attention or even critiquing such errors in a small group of 'friends'. Its members might not feel so exposed to potential ridicule (Mason & Bacsich 1998). In fact members of such groups of friends may feel more pressure to become involved in the messaging, since hiding any lack of input would be harder. On the other hand, in SNS, total strangers may also not feel the pressure to have correct spellings because the relationship between the senders and the receivers is 'faceless' and they owe each other nothing. So it does not matter what they think about one's spelling.

The findings show that letter repetition and capitalisation are used to replace emotions that are usually conveyed through vocal stress and intonation. These are most prevalent in IM and SNS. It has been noted that most of the capitalisation cases in the corpus data are followed by (excessive) punctuation in order to show the accompanying emphasis. The trend of missing capitalisations can be attributed to the high degree of informality of the CMC communication.

4.7. Punctuation

ati?????????????????!!!!!!!!!!!!
lol.....watu wanatafuta mavitu kweli!!!!!!!!!
What? (Laughter)...people really search for things surely!

Punctuation marks are symbols that are used in texts to separate words into sentences, clauses, and phrases in order to clarify meaning. According to Frehner (2008:60), the (mis)use of punctuation in CMC may take the following forms besides the standard use:

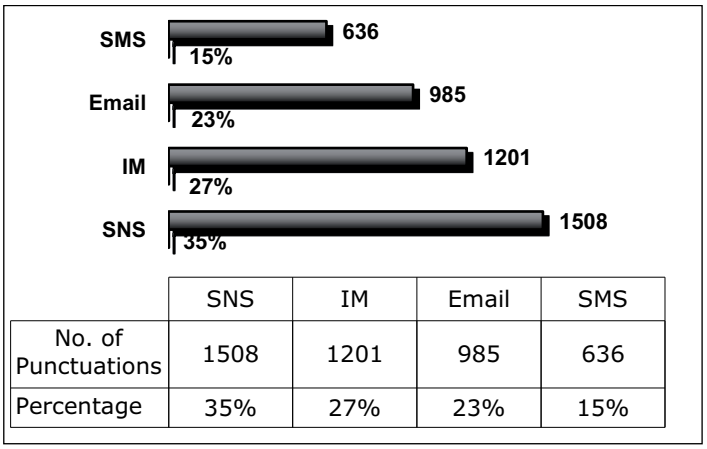
- omission of essential punctuation due to mode limitation and rapidity;
- overuse of punctuation in order to create a dramatic effect which arises from mode limitation;
- use of punctuation in inappropriate contexts.

It is clear that all these forms of usage are heavily influenced by the informal nature of CMC. In fact, punctuation in CMC may be easily disregarded by many as another form of relaxing spelling standards. This is certainly true for cases where punctuations are omitted. In any case, it is remarkable that new ways of using punctuation marks have been developed by users of CMC. These mainly involve the persistent and overuse of the punctuations as illustrated in this section. This kind of use is popular in IM and SNS. On the other hand, the absence of termination marks is prevalent in SMS as will be discussed in section 4.7.10. Hård af Segerstad (2002:145) explains that extensive use of punctuations in CMC is used to express attitude, ask questions or generally to make oneself heard. Bodomo (2009:52) advances further that there seems to be no 'rules' for using punctuation in informal CMC. The used punctuations serve to create more ease in communication, to indicate hesitations and thoughts, to express emotion and mood of the author, to indicate incompleteness of sentences, and to show informality and familiarity in informal situations.

Table 28 presents the general distribution of punctuations in the data. Its summary per genre is presented in chart 10. SMS at

(15%) makes the least use of punctuations while SNS (35%) has the most followed by IM (27%) and Email (23%).

Chart 10: % of Punctuations per Genre



	SNS	IM	Email	SMS
No. of Punctuations	1508	1201	985	636
Percentage	35%	27%	23%	15%

Similar to the use of expressive capitalisation, punctuation in CMC is used expressively as a way to show emotions. It is clear that the least use of punctuations in SMS is influenced by both mode limitation and the least effort principle. Insertion of punctuation marks in SMS takes up space and also involves effort in scrolling through the list to select the required symbol. For example in SMS, one needs 4 key taps for a question mark symbol and 5 key taps for an exclamation mark. On the other hand, the use of punctuation marks conventionally in Email is almost automatic. It is influenced by the standard writing conventions. Punctuation is also used in SNS and IM as an indicator for transitions in chats, public discussions and other forums.

Table 28 and chart 11 give the results and summary of the percentage use of each punctuation mark in comparison with the others. According to these results, the full stop, the exclamation mark, and the question mark are the most used punctuations in CMC. The full stop is used to mark the end of a frame construction in CMC. The exclamation mark is used to mark high emotions e.g. shock, surprise, disbelief etc. The question mark is used to mark questions and also disbelief and shock in some cases. The least

used punctuation marks used are the colon, and the quotation marks. The main deviation of the use of punctuations in CMC from the conventional use is that they are used freely without restraint and occur successively most of the time especially in IM and SNS.

Table 28: Distribution of Punctuations per Genre

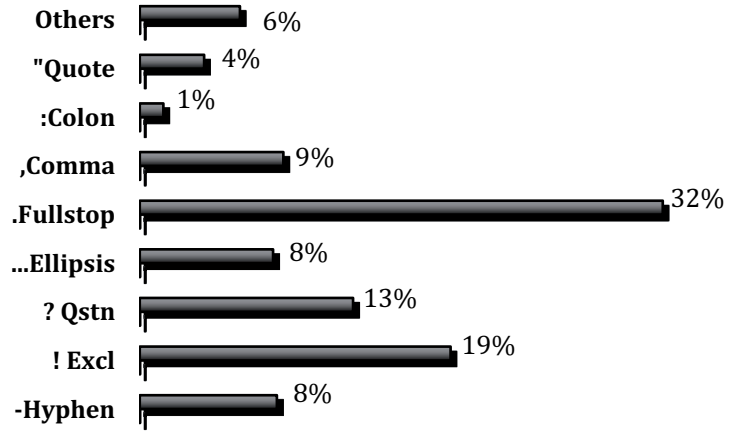
Genre	-	!	?	...	.	,	"	:	Others	**Total**
IM	94	156	204	132	406	80	46	12	71	**1201**
Email	88	111	94	78	404	128	57	13	12	**985**
SMS	75	124	154	37	119	59	9	13	46	**636**
SNS	103	425	108	103	444	110	57	24	134	**1508**
Total	**360**	**816**	**560**	**350**	**1373**	**377**	**169**	**62**	**263**	**4330**

Colon marks at 1% are hardly used as punctuation marks in the studied CMC genres. Most of them appear in form of Smileys (cf. 4.8.1). Quotation marks at 4% are also not popular compared to the rest of the punctuations. Nonetheless, they are more present in the genres with the exception of SMS which only has 5% while IM has 27%, Email and SNS both have 34%.

SMS's low count in the use of quotation marks is explained by the least effort and mode limitation principles in that locating and entering the quotation mark key requires extra effort which has to be done twice for the opening and the closing quotation marks.

Additionally, the quotation marks take up unnecessary space twice. Ellipsis marks are also not very popular in SMS because of the extra space and effort. The ensuing section will discuss the prevalent punctuation marks found in the data.

Chart 11: % Use of Punctuation Marks



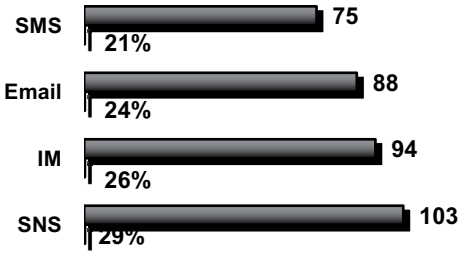
	-Hyphen	! Excl	? Qstn	...Ellipsis	.Fullstop	,Comma	:Colon	"Quote	Others
No. of Punctuation Marks	360	816	560	350	1373	377	62	169	263
Percentage	8%	19%	13%	8%	32%	9%	1%	4%	6%

4.7.1. Hyphens (-)

Hyphens/dashes are a compelling occurrence in the CMC. Not only are the hyphens used in the standard way within compound words and expressions, they are also used to separate thought contexts or transition of ideas in the data. The results of the use of the hyphen are summarised in chart 12.

Notably from the figures in the chart, (SMS 21%, Email 24%, IM 26% and SNS 29%), all the genres register a similar average count in the use of hyphens.

Chart 12: % of Hyphens per Genre



	SNS	IM	Email	SMS
No. of Hyphen Marks	103	94	88	75
Percentage	29%	26%	24%	21%

The results of the use of hyphens in SMS are slightly lower. A reason for this is that in order to type a hyphen in an SMS, one need a minimum of 3 taps. This is considered unnecessary effort especially where the message can still be comprehensible without it. Generally, the results indicate that their use in CMC is approved through all the genres. Although as already indicated, this use is more creative and not entirely identical to the standard use of hyphens.

Hyphens in the data are used to indicate ellipsis as the following illustrations show. Note that the use of the multiple hyphens in example (403) could be explained as a typographical error of the ellipsis marks. Yet, this is unlikely because it is from SNS and the hyphen and dot keys are distinct on the computer keyboard.

(403) *wow! you have given my de a smile---a true shiakago boys composition!!!! you are great---write on...boy*
(de is a misspelling of dame (Sheng dem)while Shiakago is a village high school).
Wow! You have given me a smile...a true Shiakago boy's composition! You are great...write on...boy [SNS]

(404) *gal, we r leaving now, hop 2 b there by 5-cant wait, c u soon*
Girl, we are leaving now. We hope to be there by 5. I can't wait. See you soon [SMS]

(405) *u missed out in western union driftwodz in mombasa sports, Quinz was in the hse. Mwamba rubbed Strath sumthin' crazy 46 - bolxz,in the Finalz - u shod hv, charted a plane from malindi galfrend*

You missed out in western union driftwoods tournament in Mombasa sports. Quins was around. Mwamba beat Strathmore something crazy like 46 goals in the finals. You should have charted a plane from Malindi girlfriend. (bolxz is a misspelling of bolz-balls). [SNS]

(406) *Hi-may all the good things happen to you and may all blessings touch your life.* [SMS]

Hyphens are also used to separate or divide different parts of words in which one part has been compressed.

(407) ***n-way*** *m cool. n u?*
Anyway, Im cool (ok) and you? [Email]

(408) *Gal i miss you, tafuta time* ***u-come*** *home, hata kama ni one week or few days i will appreciate.*
Girl I miss you, find some time and come home even if it is only for a week or a few days. I will appreciate. [SMS]

(409) *glad u r* ***n-joying****.*
I'm glad you are enjoying. [IM]

They are also used to separate some codemixed words as shown.

(410) *Uhoro waku mami,,* ***ngu-op*** *muriokuo wega.utuko mwega*
How are you girl, I hope you're all well. Good night. (This message is in Kikuyu: (*ngu* is a Kikuyu 1st person pronoun) [SMS]

(411) ***Singeku-sms*** *hii story coz niliogopa dre might c it n chukia me.*
I couldn't have SMSd you this story because I feared that Dre might see it and hate me. [SMS]

(412) *Was thinkin that we should form some* ***ka-debate*** *in our campo juu yah ii story ya Waki.*
I was thinking that we should form a small debate in our campus on the Waki commission. [SNS]

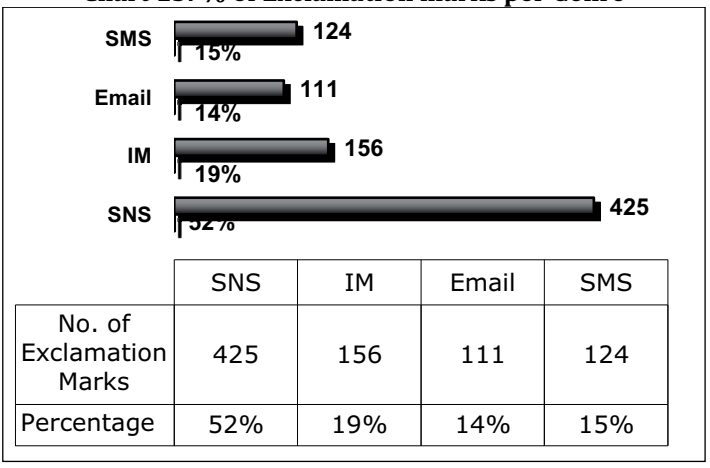
(413) *Nimebonga na huyo mshii na* ***ame-decide*** *kukupay hiyo book coz i told her it was not urs.*
I've spoken to the girl and she has decided to pay you the book because I told her that it wasn't yours. [Email]

It is surprising that in (413) the user has a dash to separate the codeswitch *ame-decide* but does not use it to separate the next codeswitch *kukupay*. It seems that *ame-decide* is considered by the user as more difficult to read without the hyphen.

4.7.2. Exclamation Marks (!)

The use of exclamation marks in CMC is worth considering as they are very popular in all CMC genres. They are especially associated with SNS and IM as explained in 4.7. They are commonly used to assist the user to show their mood e.g. excitement, surprise, disbelief, shock etc.

Chart 13: % of Exclamation marks per Genre



	SNS	IM	Email	SMS
No. of Exclamation Marks	425	156	111	124
Percentage	52%	19%	14%	15%

It is clear from the results in chart 13 that SNS at 52% makes the most use of exclamation marks. This is explained by the fact that most SNS discussion topics are sensitive and ignite animation and high emotions e.g. disagreement, shock and disbelief which are then expressed through exclamation marks. As already suggested, exclamation marks are common in SNS and IM, it is then clear from the results that the rate of occurrence between them is different.

The conspicuous range between SNS (52%) and IM (19%) and the other genres is caused by the nature of the discussions. Unlike in most SNS forums, where opinions are aired, some dialogue in IM

and the other genres may not require high emotions.

As expected, SMS and Email make average use of exclamations marks. The fact that it requires more effort to type exclamation marks in succession on the keypad explains their lower presence in SMS. This is in addition to the need for space limitation. They are only used where it is absolutely necessary. The higher rate of formality restricts their excessive use in Emails in comparison to IM and SNS.

The following are illustrations of the use of exclamation marks in the data.

(414) *Cant wait 2 c al of u again afta so long**!!!***
Cant wait to see all of you again after so long. [SMS]

(415) *ni poa wacha 2**!!!** unakam wen? Cheerz**!!!!!!!!!!!!!!!!!!!!!!!!!!!!***
It is really good! When are you coming? Cheers! [SMS]

(416) *I cudnt afford to liv you out of this latest developemnt.**!!!!** To me charges of 3k to prepare the goat is on the higher side and I doubt if guys will be willing to part with this amount**!!***
I couldnt afford to leave you out of this laterst development!!!! To me charges of 3k to prepare the goat is on the higher side and I doubt if guys will be willing to part with this amount!! [Email]

(417) *haiyaaaaa**!!!!** Waacha.....number one.*
Really? What? Number one? [Email]

(418) *tan of de gas pls**!!!!***
Turn off the cooking gas please! [SMS]

(419) *Once again, congrats**!!!!!!!!*** [Email]

In a number of cases in the data, the exclamation mark is used in place of the letters *i* or *l* in order to mask offensive obscenities (Cf. 4.3.2). Illustrations of this use include

(420) *mav**!** ya kuku* (mavi)
Chicken shit. [SNS]

(421) *sh**!**t*

Shit. [SNS]

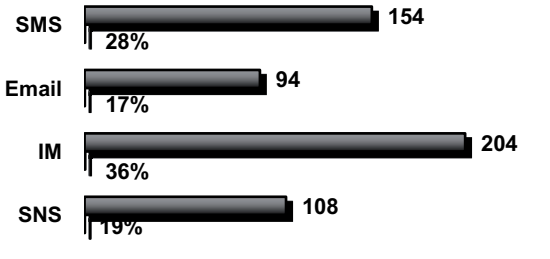
(422) *maɬaya* (malaya)
Prostitute [SNS]

It is notable that all the exclamation marks in the data still conventionally occur at the end of a word, phrase or message.

4.7.3. Question Marks (?)

Besides exclamation marks, CMC also make use of the question mark which is a symbol that is placed at the end of a phrase, sentence or message to indicate a direct question. The CMC data shows that question marks are also very frequent and are also associated with IM and SNS (cf. 4.7). In addition to the standard use of indicating direct questions in CMC, the question mark is used to indicate confusion and puzzlement.

Chart 14: % of Question marks per Genre



	SNS	IM	Email	SMS
No. of Question Marks	108	204	94	154
Percentage	19%	36%	17%	28%

The distribution of the question mark among the genres in the chart has IM at the highest followed by SMS. This is related to the nature of their structure. The synchronous nature of IM enables immediate responses and thus encourages the use of marked questions by both participants. SMS is short and fast and can be used anywhere with network as long as the phone is on. This makes it preferred over Email in making informal inquiries which are marked by question marks. It takes a longer time to get an-

swers via Email especially when the recipient is not in position to check mail. This makes users to prefer IM and SMS for quick and informal enquiries. The comparatively low counts in SNS forums is because forums involve sharing opinions and discussing a topic of shared interest. In some cases a question may be asked at the beginning and the rest will be replies. Some messages may also include rhetorical questions with question marks. The following are illustrations of the use of question marks in the data.

Question marks are excessively used to express curiosity as shown in the ensuing illustrations.

(423) *Who else read the True Love with her on the cover**???** First off, kudos to whoever takes the pictures or photoshops them for True Love. But seriously…. That article**???????????*** [SNS]

(424) *OH MY…. Yes... How did you know**????*** [IM]

(425) *We ile doe ni vipi**??***
You, what's up with the money? (*we* is short for *wewe-you* and *doe* is Sheng for dough / money). [SMS]

(426) *Morning!! Is it 4am**???*** [SMS]

(427) ***??*** [SMS]

Illustration (427) is an example of shock and puzzlement.

In a similar way to standard language, question marks in CMC are also used to mark rhetorical questions e.g.

(428) *What exactly would constitute a good economy for kenya? Little debt with resulting high 'unemployment' or high debt with resulting low 'unemployment. How does one measure unemployment? Are the family living in ushago and living off their land and livestock unemployed?* (Ushago is Sheng for rural home). [SNS]

There are also cases which involve the use of a combination of both the question mark and exclamation mark to show doubt, surprise and sometimes disappointment at the same time:

(429) *Did Mercy show you the sms fathee sent asking me to send him 10k coz mothee's*

*rent has not been paid for 3 months**?!***
Did Mercy show you the SMS that dad sent asking me to send him 10 thousand Kshs because mom's rent has not been paid for 3 months? [Email]

(430) *hi baby gal **!?** wake up its a beautiful day!!* [SMS]

(431) *Niaje kujifeel hivyo**???!!!***
Why all that pride? [SMS]

(432) *Arrest this man straight away !!Is he still a minister**???!!!*** [SNS]

Notably just like with exclamation marks, question marks still conventionally occur at the end of a word, phrase or message. Similar to these findings, Lee (2002), research on ICQ messages found out that exclamation marks and question marks are used in the following ways:

- Repetition of question mark: to emphasise that the message writer is curious and eager to know the truth;
- Repetition of exclamation mark: to emphasise the tone of utterance;
- Combination of *?* and *!* (*?!* or *!?*) to show surprise and doubts at the same time.

These findings are similar to the current findings of this research regarding the exclamation mark, the question mark and a combination of both.

4.7.4. Ellipsis (...)

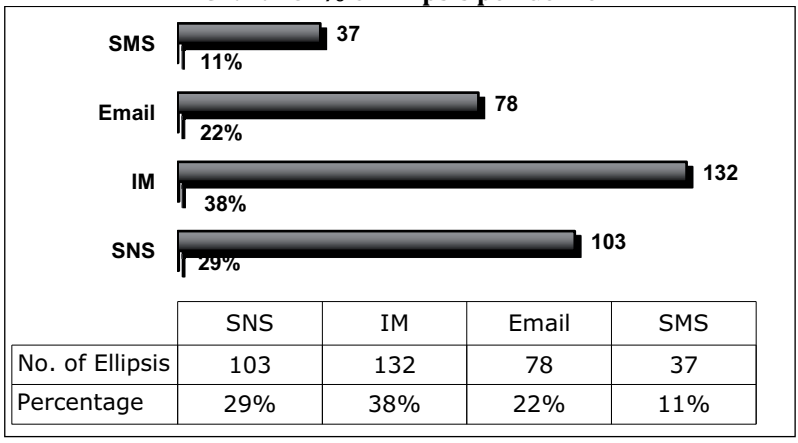
Ellipsis is defined here as a mark or series of marks used to indicate an intentional omission of a word or a phrase from the original text. Ellipsis marks can also be used to indicate a pause in speech, an unfinished thought or, at the end of a sentence, a trailing off into silence. The most common form of an ellipsis is a row of three periods (...). Other encountered forms are: three asterisks (***), one *em* dash (—), multiple *en* dashes (––), and the Unicode ellipsis symbol (…). The triple dot punctuation mark is also called a *suspension point, points of ellipsis, periods of ellipsis,* or colloquially, *dot-dot-dot*.

Maness (2007) extensive work on the *power of dots* in IM rooms reveals that ellipsis is one of the favourite constructions of Internet IM rooms, and has evolved over the past ten years into a staple of text-messaging. He continues that though an ellipsis is technically complete with three periods (...), its rise in popularity as a *trailing-off* or *silence* indicator, has led to expanded uses online. Today, extended ellipses of seven, ten, or even dozens of periods (.....) have become common constructions in Internet IM rooms and text messages. Use of these extended ellipses to end sentences, indicate a pause in speech (a role typically assigned to the comma), stand in for "that was not worthy of a reply", or "bleep out" offensive words has made this one of the most versatile grammatical structures.

In CMC there tends to be a persistent use of dots. In many cases, the use of the dots does not have any relationship to the omission of some text discourse. Additionally, there does not seem to be rules to guide the ellipsis usage. For example in some cases, only two dots are used while in others, more than three dots are used.

The results in chart 15 show the counts of ellipsis marks in the data.

Chart 15: % of Ellipsis per Genre



	SNS	IM	Email	SMS
No. of Ellipsis	103	132	78	37
Percentage	29%	38%	22%	11%

In order to distinguish the ellipsis count from the full stop, only successions of dots were registered as ellipsis unlike single dots for the full stop. SMS (11%) has the least count of ellipsis marks thus showing the influence of a combination of least space and effort as already discussed under punctuations. Each ellipsis mark for SMS takes over space and at the same time, pressing the dot key too long results in a different symbol. Therefore one has to know the exact moment to press again in order to get the next dot in succession to mark the ellipsis. This requires extra effort for SMS users and is thus not commonly used. IM (38%) has the highest count owing to its synchronous nature. The parties involved use ellipsis marks as a form of trailing off to give a chance to the other person to respond. These trail offs mark that their chat is still in progress. Additionally, they signal a change in the subject. Email (22%) and SNS (29%) have an average count.

The next illustrations involve the use of more than three dots for ellipsis marks. This indicates that it is used casually.

(433) *smh**.....**at u thinking u were listening to a common song**.....**@carni**....**that song is an all time LOW**.....*
Msanii..... I was listening to that song..... WHAT?
(Shaking my head) at you thinking you were listening to a common song…at the Carnival (Carnival is a discotheque in Nairobi)…that song is an all time low
Msanii (name) …I was listening to that song….What? [SNS]

(434) *Even money thrives on the fact that it is scarce and it follows that everyone is looking for it**....** great advice**.....**tuuge atia*
what can we say? (Kikuyu) [SNS]

This is the persistent use of ellipsis marks throughout the message:

(435) *Wacha those ones**...**there are you know...those 13 inch wheel caps with spinners**...**joke of the century**..**haki i almost ingiad a ditch**..**hilarious! hehehehehehhe!!*
Leave those ones alone...there are those 13 inch wheel caps with spinners**...**joke of the century**..** i almost ingiad a ditch**..**hilarious! (laughter] [SNS]

(436) *take a look guyz and let me know...am not for or against...but assuming luos are nilotes and most nilotes are from the north...seems like luos were egyptian...enjoy the link* [SNS]

(437) *In mats i read najivunia kuwa mkenya stickers**...**in others navumilia kuwa mkenya & naumia kuwa mkenya**...**hehehe**...***
In some public transport vehicles I read stickers saying I'm proud to be Kenyan...In others, the stickers read I persevere to be Kenyan and I suffer to be kenyan...(laughter).
[SNS]

Some cases like (438) and (439) make use of 2 dots instead of the standard one dot (fullstop) or the standard three dots (ellipsis). These can be interpreted as either typos or the use of least effort. Others like in (433), (434) and (441) make use of four dots. This use of four dots gives different impressions. It could be a typographical error. It is also possible that the sender intends to achieve both ellipsis and termination at the same time or it could be that it is meant to achieve continuity.

(438) *i say pple have gone amok in relations**..**but women lead the game from the front**..**infidelity among the urban women has gone crescendo, while the guys enjoy their pints women both single and married unashamedly gone indulge**..**WHERE ARE OUR MORALS! its both alarming and shameful!**..*** [SNS]

(439) *Jos watakam nxt month**..**I'll tell u so that we can kam and visit u guys**..**gudnytini*
Jos will come next month, I'll tell you so that we can come and visit you guys. Goodnight.
[SMS]

An important trend noted in the use of ellipsis is the fact that in many occurrences, it is used at the end of the message. This is possibly to show that there is still lingering thoughts or that there is room for more discussion both ways.

(440) *anyway im happy so long as alls'well...love you lots...byee**.....*** [Email]

(441) *wooooooiee hebu weka telly chnel 2 kuna movy poa...im watching it!! ina soundtrack poa pia**....***
Please put your TV on channel 2. There is a nice movie that I am watching. It has a nice soundtrack too. [SMS]

(442) *umeziget? hop so oderwys, i fotocopy myn 4 u coz im now in tao**...***
Have you received them? i hope so otherwise, I can photocopy mine for you since Im in town. [SMS]

(443) *sawa, lets mt @ de gate den we enda pa1**...***
Ok, lets meet at the gate then we go together. [SMS]

Some messages display the use of commas in a similar way as the dots. It is unclear whether this is intentional or whether it arises from typos involving the comma which is next to the full stop on the keyboard. The following are illustrations of this.

(444) *stop arguin u people this is a mbeere song,,, u can even hear ther accent,,,so not kao,, R.I.P the black guy (nimesahau jina lake)*
You people stop arguing. This is a Mbeere song...You can even hear their accent...so it is not Kamba...Rest in Peace, the black guy who (I've forgotten his name*).* Mbeere is a Kenyan language similar to the Embu language in Eastern Kenya. Kao is the Sheng word for the Kamba language. [SNS]

(445) *ok ok ok sawa sawa hands down „umekanyaga all mistress's kwa industry*
Ok, you have won by defeating all women in the industry. [IM]

(446) *They should atleast have spared the ambulances to be used in public,,the police got no cars.....why cant they get the good cream and rebuild em to police cars....most of this cars are 1000s times better than the ones running the roads,,,,,...crazy* [SNS]

These illustrations indicate that ellipsis marks are used without any rules or conventions. Instead of only using the standard three dots, users can type two or more than three dots as seen in example (433) and (434). This creates a possibility that although the dots appear like ellipsis and are even derived from it, they could be complete *neologisms* basing on their structure and different functions or the lack of them.

4.7.5. Comma (,)

The comma is used to indicate pauses and the separation of ideas in a text construction. Nonetheless, in the data illustrations, it is not clear whether the use of the comma is deliberate or it is a typo since the comma and the full stop are adjacent on the computer keyboard. An example is

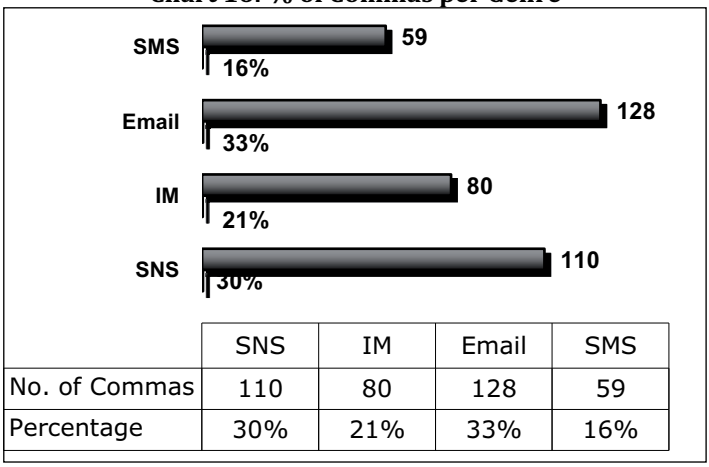
(447) *we're ok tu,*
We are just ok. [SMS]

This illustration has two possible explanations. Either the comma at the end of the construction may have been a typo in place of a

full stop or the message may have had a continuation which was not included.

The results show that the comma in CMC is more prevalent in Email (33%) and SNS (30%). This is caused by the narrative nature of the genres which requires pauses and separation of ideas (and therefore longer sentences).

Chart 16: % of Commas per Genre



	SNS	IM	Email	SMS
No. of Commas	110	80	128	59
Percentage	30%	21%	33%	16%

IM's lower count at 21% is because it mostly uses short constructions for rapidity. The short constructions already mark the separation of ideas.

The lower count by SMS (16%) is because of least effort and mode limitation like in the use of the rest of the punctuations. In the same way as the full stop, the comma does not seem to add any valuable information to the message and is therefore unnecessary for SMS.

The following are illustrations of the use of the comma in the CMC data. One of the observations about the use of the comma is that it is made to function as a fullstop within some messages. This creates run on sentences.

(448) bure kabisa, I'll update you kesho, gudnyt 4 now

Completely useless (is a Sw. swear expression), I'll update you tomorrow, goodnight for now. [SMS]

(449) *Imagn, kina Davy's hao was ingiliwad!! luckily they were not hurt, jst locked in the store& thngs stolen..wameshtuka!!*
Imagine Davys' house was broken into! Luckily theywere not hurt. They were just locked in the store and things stolen. They have panicked! [SMS]

It is used to indicate pauses.

(450) *hi, I've bambad sm nice clad n shoes, si u pitia lunch tym 2 check dem out, I wil tek dem 2 tao jioni n op 2 uza dem-kuja b4 n mek ua pik, C u*
Hi, I've gotten some nice clothes and shoes. Why don't you pass by and check them out during lunch time ? I will take them to town in the evening hoping to sell them. Come before them and make your pick. See you. [SMS]

4.7.6. Quotation Marks (")

Quotation marks are used to mark direct citations. In comparison to the discussed punctuation marks, quotation marks are not commonly used in CMC as the data shows. Their general low occurrence is for the reason that Email, IM and SNS have alternatives; for example italicisation, capitalisation, colour or the use of other symbols that are used to replace quotation marks. The quotation marks only form 4% of the total punctuation marks as shown in chart 11. Their occurrence in SNS, Email and IM is average. They are least used in SMS because of both the least effort and mode limitation principles. These marks are used in a standard way in the data besides some cases of lack of consistency e.g. using a single quotation mark for opening and double quotation marks for closing the quote, or having opening quote marks but no closing marks as the examples (452) and (453) show.

(451) *Why do you have to end your contribution with "Thank you"as if it's a speech you were giving and we were sitted somewhere in a hall listening?* [SNS]

(452) *kweli, "kikulacho ki nguoni mwako...*
Indeed, " what eats you is in your attire" . (This is a proverb in Kiswahili) [SNS]

(453) *Ati the AC is digital, and the garage door opens when you press the button on the visor and sema 'OPEN SESAME!",.....*
That the AC is digital, and the garage door opens when you press the button on the visor and say " OPEN SESAME!" [SNS]

4.7.7. Colon (:) and Semicolon (;)

The colon and semicolon are generally used in text to introduce something that follows e.g. an example, a list etc. As shown in chart 11, they are the least used in CMC at 1% in comparison to other punctuation marks. The reason for this is similar to quotation marks where users have other easier and clearer features like italics, different font or colour to mark the new information. The use of colons and semicolons in Smileys is presented separately in section 4.8.1. Apart from the construction of Smileys, the available cases on the use of colons and semicolons in the data are used in a standard way as the next example shows.

(454) *Stop derailing the thread man, just say mnaumia au la? ANSWER: with the amount of wrath yr responding to this thread tells it all.*
Stop derailing the thread. Just confirm whether you are suffering of no. The amount of wrath that you are responding with tells it all. [SNS]

4.7.8. Apostrophe (')

The apostrophe is used to indicate the possessive case, the omission of characters from a word, or the plurals of numbers, letters, and abbreviations. For example in (455) (456) and (457), the apostrophe has been used in a standard way to mark the omission of *e* in SMSed, to mark the plural of the abbreviation *SMS* and letter *p* respectively.

(455) *Sori thot u **sms'd** me bt it was a frnd of mine.*
SMSed
Sorry I thought you send me an SMS but it wasn't you, it was a friend of mine. [SMS]

(456) *mis u 2 n thnx 4 da **sms's**.*
I miss you too and thanks for the **SMS**. [SMS]

(457) *imagin **P's** zangu zimeeanza improm2*
Ps(Sh: period)
Imagine my **period(s)** has begun impromptu. [SMS]

Past researches like Hård af Segerstad (2002) and Frehner (2008) claim that there is a general tendency of omitting the apostrophe in CMC. In fact Thurlow (2001:289) captures this very well in his statement that CMC is to be blamed for the death of the apostro-

phe. He later discovers that it is not dead yet (Thurlow 2003). My observation is that the apostrophe is currently used in creative contexts. It is mostly omitted or scarcely used to mark regular contractions and possessive contexts. The summary of this usage is provided in table 20 which shows that 64% of contractions omit the apostrophe. The innovative contexts in which the apostrophe is used in the data are as follows:

It is used to mark orthographic change, for example in some cases like in (458) and (459) the apostrophe is used for two purposes. One is to mark the contraction and the other is to indicate the different pronunciation of the preceding word.

(458) ***Lee'me*** *no ow r doin* (omission of *t* and lengthening *e* in *let*)
Let me know how you are doing. [SMS]

cf. (376) ***Y'DO*** U ACT LYK A KID???GROW UP!!!
Why do you act lyk a kid? Grow up! [SMS]

Another example where it is used to mark orthographic change is in illustration (459). Interestingly in this example, the apostrophe is not used on the contraction *can't*, yet it is used on *see you (c'ya)*.

(459) *i cnt wait 2* ***c'ya****.*
I can't wait to **see you**. [SMS]

In the next cases, the apostrophe has been used to mark the past tense. In illustrations (460)-(462) it is also used to mark some internal past tense codeswitches, e.g.

(460) *They* ***tokea'd*** *2 defend themselves..*
Sw Verb *tokea(show up)* + Eng suffix *-d*
They came out to defend themselves. [Email]

(461) *n'liskia ati u* ***lea'd*** *dreadz*
Sw Verb *lea (tend)*+ Eng suffix *–d*
I heard that you are keeping dreadlocks. [IM]

(462) Have u **sikia'd**
Sw Verb *sikia (heard)*+ Eng suffix *–d*
Have you **heard?** [IM]

It is also used to mark the clipping of different parts of words for example, the final part.

(463) *gdnyt & swt drms **darl'***
Goodnight and sweet dreams **darling**. [SMS]

(464) *now dey can relax a **lil'** bit*
Now they can relax a **little bit**. [Email]

(465) *I'l send u **thro'** Mpesa*
I'll send you **through** Mpesa. (money transfer via mobile phone) [SMS]

(466) *u **rem'** her?*
Do you **remember** her? [IM]

(467) *Nilienda **hos'** nikatestiwa*
Nilienda (Sw: I went) hospital nikapimwa (Sw: I was tested)
I went to the **hospital** and was tested. [SNS]

(468) *Tom 2day is lukin fly,hakunishow **w'aving** a date,niko confused.*
Tom is looking smart today, he did not mention that **we are having** a date, I'm confused. [SMS]

(469) ***I'kol** kesho*
I will call tomorrow. [SMS]

In (468) the apostrophe is used to clip the middle part of compound words. The words *some, of* and *good* as shown in (476) and in the following examples recur in the data many times with apostrophe marking their clipping.

(470) *V rili tryd kabisa 2 gt **s'one** 2 haus me invain. 2dy's my last dy in otel.*
I've really tried a lot to get **someone** to house me in vain. Today is my last day in the hotel. [Email]

(471) *yani u wa kidding or **s'tng**?*
yaani(Sw: you mean)
You mean you were kidding or **something** ? [SMS]

(472) ***g'tym***
Good time. [SMS, Email, IM]

(473) ***g'd9t***
Good night. [SMS, Email, IM]

Another recurrence is the use of the apostrophe to clip prepositions and the conjunction *and*. Examples include

(474) *lets mt @ 5* ***n'*** *go pa1.*
(Sw: pamoja)
Let's meet at 5 **and** go together. [SMS]

(475) *Sasa?clozn* ***n'*** *6th,c u den!*
Hi, we are closing **on** 6th see you then! [SMS]

(476) *stl* ***n'*** *bd actually,g'dy!*
I'm still **in** bed actually, good day! [SMS]

(477) *Imagine I kant av jst run out* ***o'*** *ideas!!*
Imagine I cant have just run out **of** ideas! [IM]

(478) *jst a matter* ***o'*** *3 deys*
Just a matter **of** 3 days. [Email]

(479) *out* ***o'*** *gud faith!*
Out **of** good faith. [SNS]

(480) ***t'*** *raind lyk hell*
It rained like hell. [SMS]

(481) *kip* ***t'*** *up 2*
Keep **it** up too. [SMS]

In some cases like (482) the apostrophe is used creatively to separate the two identical numerals. This not only retains the intended emphasis but also their double syllabic pronunciation. In (483) the apostrophe is used to contract the words into one phrase.

(482) *Mazeh! Tulichotwa na makarau jana 4 bein* ***2'2*** *rude*
Gosh! We were arrested by the police yesterday for being too rude. [IM]

(483) ***us'*** *we're well.* (It is common to begin with an object pronoun before the subject pronoun in Kenyan English). [SMS]

(484) *Nilikuwa club jana with* ***ol'*** *frenz.*
nilikuwa (Sw: I was) jana (Sw: yesterday)
I was in a disco yesterday with **old** friends. [SMS]

(485) *Heard u r doing* ***ev'rything***
I heard you are doing **everything.** [IM]

(486) *sm js dn gt wt ts all* ***'bout****.*

Some just don't get what it is all **about**. [SNS]

(487) *Thiz thread seems tyt...& has attracted very **balanc'd** opinions*
This thread seems tight and has attracted very **balanced** opinions. [SNS]

In illustrations (484)-(486) the apostrophe has been used to replace an ellipted letter. Therefore it is neither a matter of least effort nor mode limitation because the effort and the space used is the same had the whole word had been typed.

Apostrophes in CMC are also used to create a distinction between words, for example:

u'r

It is remarkable that the contraction *u'r (you are)* mostly occurs with an apostrophe. This enables it to be distinguished from the possessive form *your* which is commonly presented as *ur*.

(488) *eeh **u'r** LUCKY!!*
Hey you are lucky! [SMS]

(489) *so long as **u'r** a jeng..* (Jeng is Sh: Luo)
So long as you are a Luo. [SNS]

(490) *glad **u'r** progresin*
Im glad you are progressing. [Email]

we're

This also mostly occurs in its standard form with the apostrophe to avoid confusion with *were*.

(491) *its just that **we're** kept too busy (**we are**)* [Email]

Absent Apostrophe

It is worth mentioning that some cases that naturally require the apostrophe do not have it. The following are such examples.

It's

Unlike *u'r* the contraction of *it is (it's)* occurs a lot without the apostrophe, thus resembling the possessive 'its' (ownership).

(492) ***Its*** *his birthday this weekend.* [Email]

(493) ***Its*** *cold hukus!!*
It's cold here. [IM]

(494) *hi. sori* ***its*** *takin too long to load.*
Hi, sorry it is taking too long to load. [IM]

Other contractions without the apostrophe are

(495) *he* ***shunt*** *be mxd in it!!!*
He should'nt be mixed in it! [SMS]

(496) ***Dint*** *manage. sorry;-((*
I didn't manage. Sorry. [SMS]

(497) *C* ***ul*** *come we kunywa n have fun.*
So **you'll** come then we drink and have fun. [Email]

(498) *i rimeba its kitu* ^{4th} *but I* ***aint*** *sure.*
I remember it is on 4th but I **ain't** sure. [SMS]

(499) ***Uv*** *hat my filings*
You've hurt my feelings. [SMS]

Example (496) displays a contraction that has been shortened further by the omission of *d* in *didn't*.

4.7.9. Other Symbols

The findings reveal the use of other symbols. The main interest is that these symbols are used to communicate feelings and emotions. They are used to indicate high emotion like swearing, anger, over-excitement etc. These frequently occur in all the CMC genres where users need to expose opinions and feelings with which they are uncomfortable to use words to express them directly. They therefore resort to the use of symbols. Examples from the data include

(500) *What is sweeter than bitinng n swaaloooooooing a chunk of irigu****#@!***
banana (Kikuyu)
What is sweeter than biting and swallowing a chunk of banana? [SNS]

The phrase *swaaloooooooing a chunk of irigu#@!* is used idiomati-

cally to mean 'a bitter pill to swallow'. Illustrations (501)-(503) show swear words similarly masked with symbols.

(501) *wooooiii uko na pad? imagin P's zangu zimeeanza improm2 sh**%#t!!!* (shit)
Oh dear, do you have a pad? Imagine my period has come unexpectedly! [SMS]

(502) *f**k dis intro n lets get dwn 2 business*
Fuck this introduction and let's get down to business. [SNS]

(503) *where did u get all the money to start these businesses bet u were the one who was washing those ******s* (asses) [SNS]

Illustration (504) exemplifies the indiscriminate use of punctuation in an SNS.

(504) *t's unfortunate that the legistrators, judges and who knows who can illegally be exonerated from paying taxes!!!!! R Ordinary Kenyans following this???.....I too, have exonerated myself from paying taxes..I have been paying taxes for years.. ...I need a break now! Tax break!. I expected a nasty move with this tax issue just like any other corrupt government.... Does it now ring a bell why Kimunya was shown the door???. Disgusting huh?...I hate all selfish MPs. When I get a chance, if only they let me..huh!!, I will collect all the unpaid tax plus interest accrued!! WASHINDWE[53]#!*%...* [SNS]

The use of the * in (529) and (531) and *$* in (535) also illustrate the other contexts of these symbols. All these are products of creativity by the CMC users.

4.7.10. Termination Marks

In this section, I begin by discussing the absence of termination marks at the end of CMC messages. I then describe the occurrence of full stops in the data. The distinction between these two is that full stops occur within the message to mark sentences, while termination marks occur at the end of the entire message. Termination marks is a general term for all kinds of marks that can be used to mark the end of a construction, in this case, an entire message.

[53] *washindwe* is used as a swear word although it literally means 'be defeated' in Kiswahili.

It is of note that although most punctuation marks in CMC are used persistently and in excess within the message, the lack of termination marks in many messages is also very obvious. This omission of termination marks in essential contexts is common in CMC due to the least effort and the fact that in many cases, not much information is carried in the termination mark especially at the end of the message. CMC messages end 'abruptly' without any visual sign. This also points to the relaxed spelling standards in CMC. The results show that 54% of the total analysed messages do not end in a termination mark.

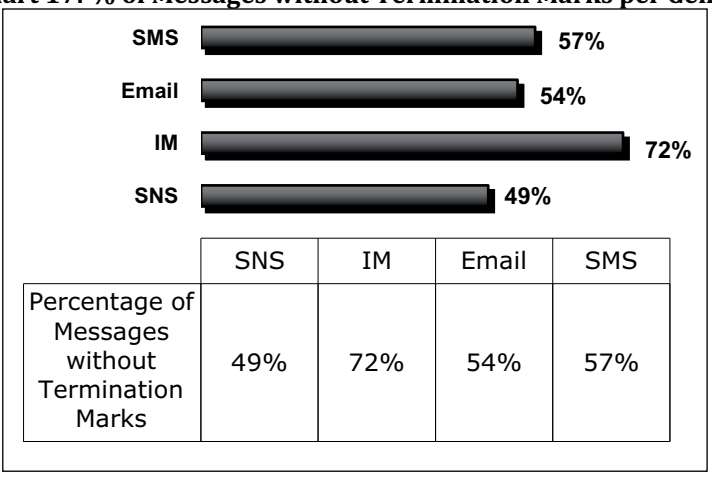
Table 29 and chart 17 show the counts of the messages without termination marks in each genre. IM leads in missing termination marks. In fact, since the calculation is based on 300 SMS, 212 SNS, 35 Email and 32 IM, it means that 72% of IM do not end in termination marks. This is explained by the fact that IM is a form of synchronic conversation and it is left open for a later opportunity.

Table 29: Messages without Termination marks

Genre	Total No. of Messages	No. of Messages without Termination marks	% per Genre
SMS	300	170	57%
Email	35	19	54%
IM	32	23	72%
SNS	212	103	49%
Total	**579**	**315**	**54%**

Additionally as explained in section 5.1.2, the connection may fail during the chat thus not enabling the participants to sign out formally. More than half the messages in the SMS genre i.e. 57% are without termination marks. Email has 54% messages without termination marks while SNS has 49% messages without termination marks. All these are explained by the least effort principle.

Chart 17: % of Messages without Termination Marks per Genre



	SNS	IM	Email	SMS
Percentage of Messages without Termination Marks	49%	72%	54%	57%

Some examples of the messages without termination marks include,

(505) *hii video ni kali...saana tu..saana*
This video is very good. [SNS]

(506) *Hi hpe u'r fyn. hv sm news 4 u..c u g9t*
Hi, hope you are fine (I) have some news for you see you goodnight. [SMS]

(507) *sina credo..nitakukol later*
I don't have credit I will call you later. [SMS]

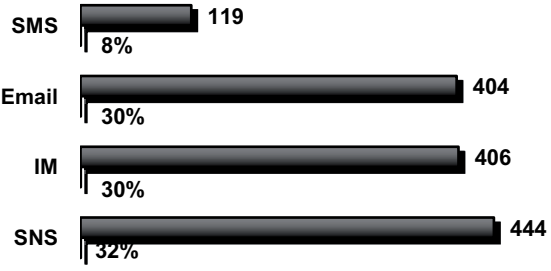
In these examples, one can clearly see that no termination mark is inserted at the end of each message despite the use of other punctuation marks in the main body of each of the messages. Another interesting observation which is similar to Frehner (2008:62) is that messages that end with Smileys or Emoticons do not make use of termination marks. Illustrations where Smileys and Emoticons are used to replace termination marks include (387), (396), (536), (496), (513) and (529).

Full stop (.)

A full stop is a dot mark placed at the end of a series of words. In this research, there is a distinction between the full stop and termination marks. A full stop can occur within the message, while

termination marks include any form of punctuation mark that occurs at the end of the entire message. Table 28 shows the full stop at 32% as the most commonly occurring punctuation mark in the CMC data.

Chart 18: % of Full stops per Genre



	SNS	IM	Email	SMS
No. of Fullstops	444	406	404	119
Percentage	32%	30%	30%	8%

Chart **18** shows SNS, Email and IM with comparably high counts of 32%, 30% and 30% respectively. The explanation to this is that messages in these genres are lengthier and need shorter sentences to communicate successfully. This requires the use of more full stops within the message. SMS has a noticeably low count of 8%. This is caused by the need for least effort coupled with the general brevity of the messages.

The insertion of a full stop requires an extra key tap and yet it does not add any valuable information to the message. It is therefore viewed as superfluous in SMS. This concurs with Feuba's (2009:8) claim that punctuation, such as the full stop, is often unnecessary to most SMS texters as the end of a line will signify the end of the message.

Seemingly, the full stop at the end of the message is regarded as predictable by many readers such that there is no need to see it visually. This is proven by the relatively high count of the occur-

rences of full stops within the message at 32% in comparison to all punctuation marks. In contrast, the scenario is opposite when compared to the counts of termination marks at the end of the message. Illustrations of messages with full stops in the body but none at the end of the message include

(508) *Thanx 4 tha kol. Jus no dat i stil lov u n no 1 wil eva change dat. If u think dat m changing u r rong. Thanx 4 evrthin.Jus kip ur prmise n b strong 4 me i lv u*
Thanks for the call. Just know that I still love you and no one will ever change that. If you think that I'm changing, you are wrong. Thanks for everything. Just keep your promise and be strong for me. [SMS]

(509) *I don't know wether u can call this crazy or madness but this GUY is Hilarious. Even Mathare*[54] *can't accomodate this guy. Thank God he lost in last year's elections!! i just cant imagine how he would behave in parliament.luv it* [SNS]

(510) *Morning!! Is it 4am??? The pics were sent jana. Check mail. We r well, abt 2 go bck 2 slp- gnite*
Morning, is it 4 am? the pictures were sent yesterday. Check mail. We are well and about to go back to sleep. Good night. [SMS]

Some examples are in the data where full stops are used to mask offensive obscenities, for example:

(511) *sh.i.t* [SNS]

(512) *shut de f.ck up ma.t.a.ko wewe*
Shut the fuck up, you ass [SNS]

The full stop within the word is also used to 'shout' or achieve stress like in illustration (390) where it is combined with the use of capitalisation to produce a 'screaming' effect in the word *N.E.V.E.R.*

4.7.11. Conclusion

The use of punctuation marks in CMC is spontaneous and informal. The marks are used repeatedly and persistently, and in most cases overused in both valid and ungrounded contexts. Contradictorily, the punctuation marks are omitted in some contexts where

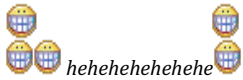
[54] Mathare is the main mental hospital in Nairobi, Kenya.

they should naturally occur for example as termination marks at the end of the message.

SNS makes the most use of punctuation marks followed by IM and Email. This is in order to communicate emotions or transition from one idea to another. SMS has the least punctuations because of its mode limitation, rapidity and least effort. The dot is the most used as a way to separate ideas. The colon punctuation mark is the least used.

The use of punctuation marks in creative ways in CMC can clearly be seen, for example the use of excessive marks to communicate deep reactions or stress e.g. the use of a series of exclamation marks to communicate deep shock or question mark to show confusion or curiosity. Other creative uses of punctuations include the use of apostrophe to mark word clippings, codeswitches.

4.8. Graphics



Smileys and Emoticons derived from the terms *emoti(on) icons* are graphics that attempt to fill the gap of psychophysical features. They are tantamount to CMC and have facilitated it in a major way in terms of visual non-verbal communication in the late 20th and early 21st century (Horn 1998). They are widely used especially in casual communication. Besides making it more interesting, they make it more real in that they present emotions, actions, non verbal communication and even voice cues. They are also very important in CMC because they are used to mark different moods and this helps to avoid misunderstanding. For example jests are clearly marked and cannot be taken literally. Their main drawback is that the receiver is not able to decipher the genuineness of the representativity of the Smileys and Emoticons through the network. For example it is easy to fake well being by sending a happy Smiley or Emoticon even if one is physically obviously covered in injuries. But such factors do not disqualify the usefulness

of these graphics in aiding the presentation of both non-verbal communication and voice cues.

Emoticons are also referred to by some as Smileys. In CMC, the thin distinction between Smileys and Emoticons is that Smileys are mostly composed of punctuation marks while Emoticons are more of graphics and pictures. Both Smileys and Emoticons are associated with IM and SNS. They are easily available in these genres and are used to show feelings, emotions, actions or movements and at times new forms are used for show off amongst peers.

The distribution of the use of Emoticons and Smileys is shown in table 30 and chart 19. The findings are exactly as predicted such that Smileys and Emoticons are mainly used in IM (48%), SNS (41%) and to a lower extent in SMS (8%) and Emails (3%). Emoticons are very popular in IM and SNS which have an open allowance of a very rich database that allows one to pick out useful Emoticons and store for future use. Once the Emoticon is stored under a selected codeword, it is automatically inserted into the message when one types in this code.

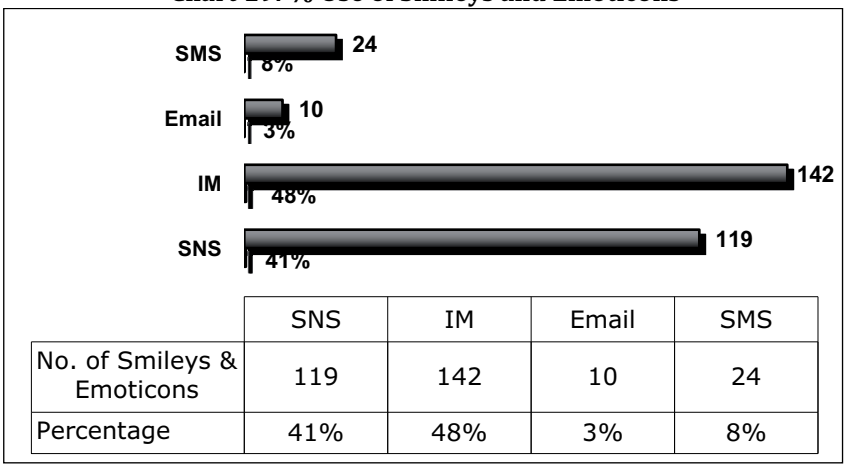
Table 30: Distribution of Smileys and Emoticons

Genre	Smileys and Emoticons
SMS	24
Email	10
IM	142
SNS	119
Total	**295**

The reason for the popularity of Smileys and Emoticons in IM and SNS is that IM users feel obligated to post their emotional reactions graphically just as if they were physically present. Similarly, SNS involve some kind of debate or discussion going on and users feel obligated to add quality by using Smileys and Emoticons. The use of Emoticons and Smileys in SMS is relatively low mainly be-

cause it not only requires space but also requires effort. For example one needs a minimum of 18 phone key taps in order to construct the basic smile *:-)* or *:-(*. Additionally, it requires 3 spaces

Chart 19: % Use of Smileys and Emoticons



	SNS	IM	Email	SMS
No. of Smileys & Emoticons	119	142	10	24
Percentage	41%	48%	3%	8%

In addition, Emoticons are not used because the average phone does not have a database for Emoticons and neither can it display them if received.

Hence it is cumbersome to produce a Smiley and impossible to produce an Emoticon on the phone. The results show that Emails make the least use of Emoticons and Smileys. Email does not have a provision for motion which some Emoticons require. In addition, there are only a limited number of available Emoticons and Smileys in most Emails' database and it would take extra effort and time to copy and paste Emoticons from other sources. This finding corresponds to Frehner (2008:80) whose results contradicted the general assumption of the popularity of Smileys in Email. Frehner claims that, "the reason for the relatively rare use (of Smileys in Emails) is difficult to elicit and to a certain extent user-idiosyncratic".

Table 31 shows the frequency and use of individual Smileys in the data corpus. Note that some Smileys like those denoting attention,

nervousness, disagreement and agreement are not listed because they did not appear probably unintentionally so.

In consonance with my data, Baron (2004:28) and Frehner (2008:122) concur that the most commonly used Smiley is the smile in its various forms (cf. table 3). In my results the *smile's* popularity is followed by the *wink* and the *laugh*. This smile and the wink are available in all the genres, a fact which boost their counts. It points out to happy and jocular discussions amongst peers in CMC genres. The least used Smiley is the one denoting *illness*, and *smugness*. Apart from the fact that these are not available on the average phone, the reason for this is not very clear.

All the same, my assumption is that some information may be considered too crucial and humourless to communicate lightly using Smileys. The following is a list of examples of some Smileys encountered in the data.

Table 31: Frequency use of different Smileys and Emoticons

	SMS	Email	IM	SNS	Total
☺ Smile	8	3	15	11	37
Wink	6	3	11	8	28
LOL	0	0	16	10	26
☹ Sad	9	2	6	9	26
Love	0	0	13	4	17
Shock	0	0	5	12	17
Joking	0	0	9	8	17
Kiss	0	2	9	5	16
Cool	0	0	7	6	13
Yes (nod)	0	0	10	2	12
	0	0	7	4	11
No	0	0	7	4	11
Evil	0	0	4	6	10
Embarrassed	0	0	1	8	9
Play	0	0	6	3	9
Sleepy	0	0	3	4	7
Tears	0	0	2	5	7
Other	1	0	2	3	6
Bye	0	0	5	0	5
Thinking	0	0	1	3	4
Smug	0	0	2	2	4
Ill	0	0	1	2	3
Total	**24**	**10**	**142**	**119**	**295**

4.8.1. Smileys

Smileys are mostly created by combining a number of punctuations as illustrated in the following examples,

(513) cheki kwa shoe hapo chini ya bed ***;-)***

Check in the shoe under the bed (wink). [SMS]

(514) *wacha wachome* **:-(**
Let them burn (sadness). [SMS]

(515) Sisi 2ko tunameela **:-)**
We are here getting drunk (smile). [SMS]

Note that when the Smiley :-) is typed on some keypads and keyboards, the end result automatically turns out to be ☺ on many phones and :- (becomes ☹.

Other creative Smileys include,

(516) ***O_o*** *LMFAO wouldnt that impail her?*
(surprise) Laughing my fat ass off! Wouldn't that impale her? [SNS]

(517) *wow dis is a hot song!! thanks!!* **(^_^)**
WOW! This is a hot song. Thanks! [SNS]

(518) *Theyd chanukad 2day!!!!!* **^_~**
chanuka-d
They had wisened up today! (wink). 'chanuka' is Kiswahili for 'be smoothed' or 'detached'. It has been adapted in Sheng to mean 'be wise'.) [IM]

Smileys are popular in all the CMC genres including the SMS. Interestingly, despite the mode limitation, SMS data showed creativity in the composition of Smileys as presented in illustrations (519) and (520). These illustrations still fit into the 160 SMS character limit.

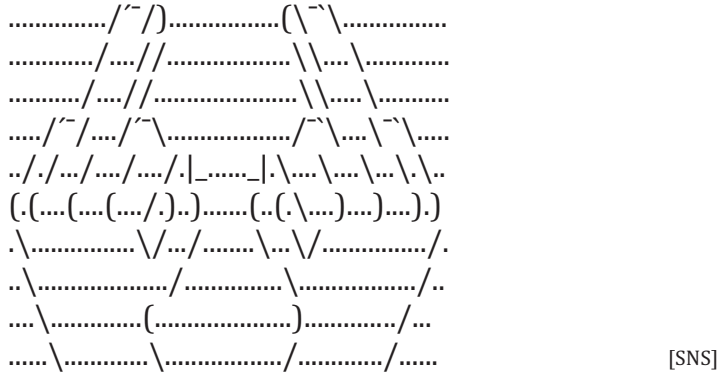
(519) **__/”””””””””””i_**
(__@___@__!---
This car is loaded wit blesings:Lov,joy,peace,hapines,sukses&gud health,Offload them bt return my car.HAPPY NEW YEAR [SMS]

(520)

This cross is sent to wish u gud luck & blessings…pass it on 2 others [SMS]

Emails also make creative use of Smileys for example,

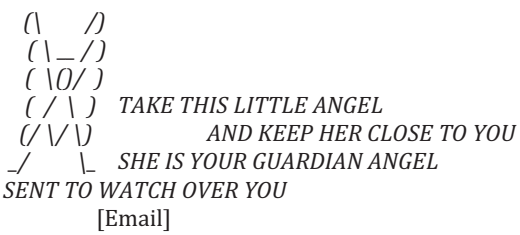
(521)



[SNS]

Illustration (521) is meant as a 'comment' in reaction to a hit song on YouTube. These graphics are included in messages for fun and entertainment.

(522)



[Email]

(523)

[SNS]

Illustration (524) denotes walking footsteps. In (519) there is an angel with flapping wings. The graphic in (523) shows a reading cat and has been used to refer to *'studying'.*

The data clearly shows that many messages with the creative composition of graphics in SMS and Email are friendly forwards and are mainly used as a way to stay in touch and an assurance of solidarity.

(524)

Some people Oooo
come into our live s ( ) /) /
and quickly go.. (_ /

oooO
( ) Some people
\ (become friends
_) and stay awhile....

leaving beautiful Oooo
footprints on our ( )
hearts.... &) /
(_/

oooO
( ) we are never quite
\ (the same because we have
\ _) made a good friend!!!! [Email]

4.8.2. Emoticons

The ensuing illustrations are of Emoticons which are pictures of different emotions of faces. Note that some Emoticons in SNS and IM are animated[55], but it is not possible to capture them here.

(525) ♥ ´¯`•.*special friend.*•´¯`♥
(Love) special friend (love). [SNS]

(526) *don't u tek my* *away!!! (love heartbeat)*
Don't you take my (love) away [IM]

(527) *No wonder his or her identity is UNKNOWN* *(laughter)* [SNS]

(528) *OMG! love it she so pretty and sings great!!!!* ***:D♥♥♥♥*** *(laughter & love)* [SNS]

(529) *paliiiiiiiiiiiiiiiiiiiiiiiiiiiiiz! get a ****ing grip bwana!*
Please, get a grip Mr. (bafflement) [SNS]

(530) *Thanks gk for changamshaing* ***ur*** *brains.*
Thanks (name) for refreshing **our** brains. [SNS]

---

[55] The animated Emoticons are available on http://www.animatedEmoticons.net/

The Smiley in (530) has been used to replace the letter *o* in the word *our*. This is a form of creativity.

4.8.3. Explained Motions

The data also reveals some motions that are explained in words. These occur frequently in IM and SNS, the genres which are close to real time conversations unlike Email and SMS. They are shown in the following illustrations.

(531) *Ditto.* *I had a break down BTW**.... ****reaches for reports and a cigarette*** [SNS]

(532) ***HAHAHAHAHAHAHAHAHAHAHAHAHAHA***

[Breathe in, Breathe out]

HAHAHAHAHAHAHAHHAHAHAHAHAHAHA [IM]

(533) *Hahahaha i suprise myself sometimes **(swallowing saliva)** anyway najua tu hatanyita oh wait! He is a genius! (**pffffft**)*
Hahahaha i suprise myself sometimes (**swallows saliva**) anyway I just know that he will not understand oh wait! He is a genius!(**sniff**) [IM]

(534) *I felt it when Obama made his victory speech! **Sigh**! that was an emotional moment* [SNS]

(535) *loooveeeeeeeee this song.....i say>>$$swing swing$$ & **I dance hands up,,, down,,,!!! dance dance*** [SNS]

4.8.4. Conclusion

The use of Emoticons and Smileys is one of the most creative practises among CMC users. They were introduced to CMC as a necessity immediately after it became clear that text messages were challenged in communicating emotions and body movements to make the communication as real as possible. This enables the messages to be creatively loaded with all the real time graphic information that is very vivid to the message recipient. These graphics are more prevalent in IM and SNS and fewer in SMS and Email. The reasons behind this are that IM and SNS are more real-time like conversations and the non-verbal communica-

tion through these graphics adds quality to the discussion. SMS on the other hand uses very few of these graphics because of the economy of space and most predominantly, mode limitation which does not give users a choice of the graphics like the other genres.

In regard to content, the findings reveal that unlike in IM and SNS where graphics are used to illustrate real time actions and emotions, SMS and Email use graphics in messages which are forwarded as a form of greeting and a show of solidarity.

Chapter 5. Socially Motivated Features of CMC

Bradner (2001:1) claims that there is a strong relationship between CMC design and social interaction by explaining that "technology researchers argue that CMC adoption fails when it interferes with subtle and complex social dynamics of groups." It should be able to complement the social lives of the users. This chapter presents CMC as a socially driven phenomenon. This includes discussions on salutations, codeswitching (CS) and Sheng and Engsh in CMC. Spoken language is the basic form of communication. It is therefore expected that CMC has adapted to it in some ways given its informal nature (Crystal 200:43). In contrast, this observation does not augur well with some scholars. Spitzmüller (2006:33) criticises the likening of CMC to spoken language by observing that the "prima facie similarities between [CMC] and face to face communication are functionally not similar at all". Hinrichs (2006:20) also takes a similar stance and claims that although CMC initially appears as conceptually oral, it is still written text.

My stand is that CMC language belongs to its own distinct language category. Although it involves the likening of CMC and speech in terms of salutation, codeswitching and general informality, this likening does not entail that these elements are analogous with speech but that some of their aspects have been adapted from speech and incorporated into CMC language.

I begin by describing salutation similarities between spoken language and CMC in terms of structure and the language used. I then discuss the featured languages and the general language choice in Kenyan CMC. Next I discuss the subject of codeswitching both at the interword and intraword level. It is a recurrent feature in the data and is closely related to spoken language which is imported into CMC. I finally give an explanation of several motivations of codeswitching in CMC.

5.1. Salutations in Kenyan CMC

sa, uko poa?? im ok. just saying hi. gdnite n take care!
Hi, are you fine? I'm okay and just saying hi. Goodnight and take care.

In the current study, salutation is the general term given to both greetings and valedictions. Salutation is a way of signing in or out of a communication scenario. In CMC, greetings signal the sender's presence or *sign in* while in contrast, valedictions signal the leave or sign out of the participant. All cultures have their own distinct ways of salutation. My salutation approach is similar to Frehner (2008) who analysed and discussed the salutation and farewell variable in her UK, German and German-Swiss CMC corpora. Frehner uses the term salutations as synonymous with greetings whereas here, *salutation* is the general term given to both greetings and farewells. I use the term valedictions to refer to farewells. Her findings are mostly corroborated by mine; the main difference from the current research is in the terminology.

This section presents a discussion of the CMC findings in relation to salutations. It begins with a general introduction of the Kenyan salutation culture. Next I look into greetings and then valedictions as forms of salutations found in the data. This is followed by a discussion of the data in each individual genre. Finally I present a comparison of the salutation findings among the genres.

According to Omar (1993:41), salutations are very important in the Kiswahili culture. She explains that exclusion of salutations or disregard of the cultural norms for performing them may lead to social consequences. Similarly, the Kenyan culture puts a lot of importance on salutations. It is considered very un-courteous and anti-social not to greet or bid a valediction to other people irrespective of whether they are strangers or acquaintances. The importance put on salutations by the communities is such that each must be attended to comprehensively. Greetings in Kenya are generally referred to as *salaams* (the word means *peace*) or *salamu* in Kiswahili. Traditionally, verbal greetings in some Kenyan cultures e.g. the Luhyia were in form of an interview session.

They involved a series of questions and answers. The first statement would announce 'peace', and then subsequent questions would ask about the general health, situation, family, possessions etc. For example in Lubukusu, a sub-group of the Luhyia, the greeting begins with a word meaning 'peace' (I come in peace), *mulembe,* which is then followed by *oriena (how are you?)* then followed by all sorts of questions regarding the situation at home, work, relatives etc. Interestingly the Luhyia in Kenya are known as the people of *mulembe* because of their fascination with greetings on FM stations, where they send the greeting message so that it is read to their loved ones. Some radio stations have designated a whole program for the *salaams*. Among the Kisii (another community in Nyanza, Kenya) the greeting begins with *bwakire* which literary means *it is a new day*. Similarly, the valediction section would involve bidding blessings for example, *go well, God bless* and *barikiwa* (be blessed). This is followed by instructions to greet those that one is going to and finally a promise to meet again.

Nationally, the personal salutation culture involves an active (warm) handshake using the right hands. Additionally, when saluting a senior, then the handshaking should be done while grasping one's wrist using the left hand coupled with a slight bow. The hand shaking salutation goes on for almost as long as the verbal salutation lasts. Another familiar form of salutation for women is to hug two times side to side and then end with a handshake, which implies *we have peace between us.* The verbal greeting goes beyond just saying *hello* to also asking about health, family and work.

Except in SNS, the collected corpus involves data between friends, family and acquaintances. These groups are related to each other informally and are expected to use salutations just as they would do in face to face communication. The greetings are expected to be informal in line with Hudson's (1996:133) claim that, salutations often show authority or power and solidarity relations for example *good morning* (power) vs. *hi* (solidarity). He adds that salutations also indicate the time e.g. *Good afternoon* and *Good night* or season e.g. *Happy Easter, Happy new year* etc. The main

difference from Hudson's findings is in his conclusion that longer greetings are used where people are uncertain of their relations, and thus need reassurance. This practice is the opposite in Kenya where longer greetings are among those who are certain of their close relationships.

Illustration (536) is an example of a vivid graphical greeting from an IM in the data. The words in brackets are translations of the Emoticons. It clearly appears as an adjacency pair with a greeting followed by a greeting.

(536) *Y)* *(handshake)*
X) *(salute)*
Y) Sasa!!!
X) Poa kabisa! Mambo? (very fine, how are things?)
Y) *(I'm cool)*

[IM]

Besides the fun and creativity offered to make such a chat interesting, its presence indicates that the interlocutors are aware of the social emphasis put into greetings as a conversation opener. Given these, there is no doubt that this salutation culture is a likely influence on CMC use in Kenya. Coupled with the fact that there are no laid down stylistic conventions in CMC, it means that a great variety of creativity is possible.

Unlike the findings of Frehner (2008:46-91) and Schmidt & Androutsopoulos (2004:63), who claim that there are very few salutations in (European and American) CMC, my data presents quite a good number of salutations in IM and Email. The data suggests that CMC salutations also follow these conversation conventions.

The ensuing section discusses the findings related to salutations from the corpus of data. I begin with a discussion of greetings and then move on to valedictions.

5.1.1. Greetings

Prior to the description of the qualitative data, I will first present

the quantitative data in form of the distribution of greetings per genre and the number of messages containing the greetings in each of the genres. These will be compared to the number of messages analysed as shown in table 32 and chart 20. The raw results indicate that out of 300 SMS, a total of 94 greetings were present in 81 messages, 35 Emails yielded a total of 86 greetings from 34 messages. Similarly, 32 IMs resulted to 103 greetings from all the 32 messages while 212 SNS resulted to 31 greetings from 30 messages. This demonstrates that the 579 messages yielded a total of 314 greetings from 175 messages. These results are summarised in chart 20. Notably, some messages in the genres contain more than one greeting i.e. a pre-greeting, greeting and a post greeting or follow-up as will be discussed in hereafter.

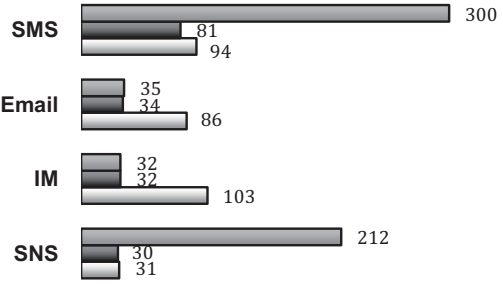
Table 32: Greetings per Genre

Genre	Total No. of Messages	Total No. of Greetings per Genre	No. of messages with Greetings per Genre
SMS	300	94	81 (27%)
Email	35	86	34 (98%)
IM	32	103	32 (100%)
SNS	212	31	30 (14%)
Total	**579**	**314**	**175 (31%)**

IM has the highest number of greetings. This is the expected norm because IM is real time, most similar to face to face communication which conventionally begins with greetings in the Kenyan culture.

For SMS, Frehner (2008:91) showed that 72% of the SMS in her data lacked the greeting. This is similar to the current finding with 73% of the SMS lacking greetings. Schmidt & Androutsopoulos (2004:63) also state that the striking feature about greetings in SMS is that they are most of the time missing. This is true for most SMS but a fact that cannot be overlooked is that close to a third of my SMS data has greetings, showing that greetings in SMS are not totally omitted.

Chart 20: % Distribution of Greetings in Messages per CMC genre



	SNS	IM	Email	SMS
Total No. of Messages	212	32	35	300
No. of Messages with Greetings Per Genre	30	32	34	81
Total No. of Greetings Per Genre	31	103	86	94

However, it is worth explaining that in addition to mode limitation and least effort, this low presence of greetings in SMS is influenced by other factors. First of all, I point out that most SMS messages without greetings are likely to be quick requests or instant replies to a message for example in (538) which is a quick response without the greeting to (537) which is the initial SMS with a request. Secondly, most SMS are mainly focused on the central message for example in (538) where the message is focused on the information requested for. The interesting point here is that although message (537) has a greeting, the reply message in (538) neither has a response to the greeting nor a reciprocal greeting. The natural turn-taking practice is to respond to a greeting with an answer first before reciprocating one's own greeting. Message (538) does not perform any of this since it focuses on the central message. Thirdly, except in SMS messages sent for greetings sake, in most cases, SMS is a temporary or interim form of communication with the participants having met shortly before or likely to meet shortly after. This narrows the need for exchanging and responding to greetings.

(537) hi pls assist me with name's email [SMS]

(538) xxxyyy@yahoo.com [SMS]

The least number of greetings are found in SNS. This is expected because the majority of SNS messages are characterised by comments based on a discussion of a particular theme. Interrupting the flow for greetings is not the norm. In addition, most cases of SNS are public comments that are made anonymously, to an unknown audience, therefore, there is no obligation for politeness. For Emails, these findings are not in line with Frehner's (2008:46) claim that Emails do not necessarily include a greeting or valediction. Her data showed that 30% of greetings were omitted in Emails. This is the opposite of the current findings where the Email genre has a high count of greetings at 98%. The data reveals that just like in conventional letter writing, greetings are used as an opener to the Email. This is the norm of the Email genre. I attribute this to the discussed Kenyan greeting culture. Another reason is that since there is a possible time lapse between receiving and replying of Emails as opposed to SMS and IM, the sender may feel obliged to greet the receiver. Additionally, the presence of greetings in Emails is also influenced by the fact that like letters, Email communication (excluding forwards and other material to be passed on) is mostly used by people without physical proximity. In such cases, the sender is likely to include a greeting.

An interesting point to note is that out of the 35 Emails in the data, only one lacks a greeting. Similar to the SMS example, the reasons for this are that this Email is an immediate reply to a recent Email sent less than 15 minutes earlier. The sender may have therefore assumed the well being of the receiver. Secondly, the content of the Email shows that it is centred on addressing the main issue of discussion which had been raised in the first Email. Both Emails and their translations are presented in illustrations (539) which is the initial Email and (540) which is the reply without a greeting.

(539) Sema bwana! mambo?
Umeskia ati list ya tp imetoka! hmmmm Imagine umepostiwa (name high) Nimeaambiwa its really in shagz full of night runners!!! hahaha u hav 2 chungana with thoz thugs watakunyanyasa mpaka uingie kwa laini;-))) so no hustlin usiku! hehehehe[...]
chungana alwayz!

mimi
Tell me man,
Have you heard that the list for allocation of schools for teaching practice is out? (hesitation *hmmmm*) Imagine you have been posted to Name high school. I've been told that it is really in the village full of night runners! (laughter *hahaha)*. You have to be careful with those thugs, they will harass you until you get aligned (wink). So no hustling[56] at night (laughter) […]
take care always,
me. [Email]

(540) Nini mbaya na wewe bwana :-)
Mambo poa. sijali wea I land lakini pls usilete ile rumours ya waLuhyia huku…..ati nyt runners??
heheheh nedamaka mono! Woooi nitahama niende kuishi tao kwenye hawatathubutu ku nyt run coz of the mingi watchiez!!! -wil kp u posted…hop yenyewe I don't bump into thoz thugz otherwyz nitaishia asap!!! […]
nitakumiss bt u can visit whenever,
byeeee,
NAME

What is wrong with you man (Smile)
Things are fine. I don't care where I end up but please don't bring those Luhyia[57] rumours here…night runners[58]?
(laughter *heheheh*) I'm really scared! (Kikuyu *nedamaka mono*). (Expression of worry *Woooi*) I will relocate to live in town where they wont dare to night run because of the many watchmen. I will keep you posted…I really hope that I don't bump into those thugs, otherwise, I will leave as soon as possible […]
I will miss you but you can visit whenever possible,
bye,
Name [Email]

Notwithstanding, a point worth considering is that the sender may have intended the first line of the Email (540) *Nini mbaya na wewe bwana :-)* to serve as both an attention catcher and greeting even though it is not well-defined as a greeting. It still is in a way a question about one's well being as a greeting. However, I restrict my counts to the obvious greeting styles and phrases in order to avoid bias.

[56] *hustling* is used with an Engsh meaning referring to *hanging out* or staying out late for some fun.
[57] Luhyia here refers to the Luhyia people.
[58] Night running is a queer practice by some people in some communities who prowl around in the night deriving pleasure from scaring people. More about this can be found on http://hubpages.com/hub/The-night-runner-syndrome

The findings also show that the most commonly used greeting expressions are informal in the Kenyan context. These are *sasa 36%, hi 34%,* and *sema 30%*. The common use of *hi* was also registered by Frehner (2008:46) who puts it in what she calls the *h-salutations* that include variants of *hello* such as *hi, hey, heya/hiya*.

Similar to oral communication, one major observation in CMC greetings is that there appears to be a pre-greeting that acts as an attention catcher before the main greeting and at times followed by a post greeting. For example,

(541) Pre-greeting: ***Hi***
Greeting: ***How are you****?*
Post-greeting or follow-up: ***Sema,*** *(Tell me)*
Maisha? *(How is life?)* [SNS]

Notably, the pre-greeting and post-greeting forms are interchangeable and can fit in any of the positions. For example *sema* in illustration (546) is used as a pre greeting whereas it is a main greeting in (545). Examples of the observed pre-greeting forms are *sasa, hi, hey, wsup,* and *sema.* Illustrations of their use as pre-greetings found in the data are:

(542) ***sasa*** *howru? wea can i get de key?*
Hi, how are you? where can I get the key? [SMS]

(543) ***Hi....*** *how are you guyz endelearing? pls call me in the jioz*
endelea (Sw) go on +ring (Eng: continuous tense)
Hi, how are you guys going on? Please call me in the evening. [SMS]

(544) ***Wsup****! How u doing? Js saying hi*
What's up? How are you doing? I am just saying hi. [IM]

(545) ***hey*** *sema? nilitravel niko UG*
Hey, How are you? I travelled I am in Uganda. [Email]

(546) ***sema*** *,habari zenu? op u'l stil kam 4 de wkend*
Hi, how are you? I hope you'll still come over for the weekend. [SMS]

In (542) - (544), the main greeting is *how are you?* which is *habari yenu* in Kiswahili. The other expressions at the beginning of the

message have been used as a way to capture the receiver's attention before the actual greeting.

Sasa

The term *sasa* which is *hi* in Sheng is the most recurrent greeting form in many messages. Besides being used in pre-greetings as seen in (542), it is also used independently for example,

(547) ***sasa*** *mrembo, long time c u kuja online??*
Hi pretty, it's been long since we communicated. Why don't you come online? [IM]

(548) *Sasa cuzo, ati ulikuwa msick?*
Hi cousin, I hear that you were sick? [Email]

Despite knowing the intended receiver, illustrations (547) and (548) also further indicate that it is pleasant to include a name or the title of the receiver as an attention catcher.

As already explained about the Kenyan greeting culture, it is an interesting observation that many messages e.g. (544) and (549)-(558) solely consist of salutations. In effect, by the statement *Js saying hi,* the sender reveals that the message is meant to be exclusively a greeting. This finding appears to be a common practice in African CMC. It has also been asserted by Lamoureaux (2009) in her study of SMS messages among Khartoum university students in Sudan. In some of these CMC messages the salutation is just sent for its on sake, with no real message accompanying it except the greeting and the valediction, or in some cases just a 'shake' to make sure one is fine. Lamoureaux (2009:39) captures this in her explanation that it seems that the actual move of keeping in touch through SMS is the most important unlike the message content. Further examples of such messages which are solely greetings include,

(549) *How r u doin? Hows ur health gal? Hop u r doin gud. Take care n b blessd!*
How are you doing? How's your health girl? I hope you are doing good. Take care and be blessed. [SMS]

(550) *Just saying hello, na salimia kila mtu*
Just saying hello, pass my regards to everyone. [SMS]

(551) *sasa,mambo,vipi kasweetie?*
Hi what's new? How are you my petite sweetheart? (The ka prefix is used before nouns as a diminutive in Kiswahili. Although a diminutive, in some cases such as this, it is used to signify affection and endearment.) [SMS

(552) *sasa??? hop uko sawa, mi ni mpoa, wel jst sayin hi-b blessed*
hi, I hope you are fine. I'm fine too. I was just sayin hi. be blessed. [SMS]

In some instances, the senders of the messages add a sort of reply or assurance in answer to the greeting (or the expected greeting) showing that the senders naturally expect the receiver to be interested in their well being. Illustrations of the pre-conceived answers include,

(553) *hi dear howz the going? huwezi nikumbuka?* ***we r ok*** *jst the usual*
Hi dear, how's the going? Can't you remember me? We are okay just the usual way of life. [SMS]

(554) *Sasa mukhana? umekimya sana.* ***Am gud****, u?*
Hi girl? You are too silent. I am fine. What about you? [IM]

(555) *Hey how're you doin. Its been long.* ***we're cool*** [Email]

In other messages the sender predicts and includes a hopefully positive reply to the greeting e.g.

(556) *Orie mukhana?* ***hop u r ok***
How are you girl? I hope you are ok. [Email]

(557) *Sasa mrembo.* ***mzima****? We r fwain....*
Hi pretty are you well? We are fine. [IM]

(558) *Sasa dear,* ***hop uko poa***
Hi dear. Hope you are well. [Email]

This practise is attributed to the custom greeting culture where the automatic response to a greeting is *fine* even though one is obviously unwell. Therefore as seen in the illustrations above, the message sender assumes that the receiver of the greeting will answer that they are well or fine. Hudson (1996:133) explains that truth and sincerity does not really count in greetings. The following are some examples where the receiver answers that they are

partly fine, and then subsequently explain that they not fine physically, financially, or emotionally.

(559) ***I'm fine****,only a fracture on ma leg*
I'm fine, except for the fracture on my leg. [SMS]

(560) *Hi siz,* ***niko fine*** *only dat nimesota*
Hi sister I am fine, only that I am financially broke. [SMS]

(561) *sasa?* ***im ok*** *though stress mob :-(*
Hi, I'm okay though there is too much stress. [SMS]

5.1.2. Valedictions

Valedictions are used in different ways such as to sign out from a message, to give goodwill wishes to the receiver or to appeal for support, unity and to keep in touch.

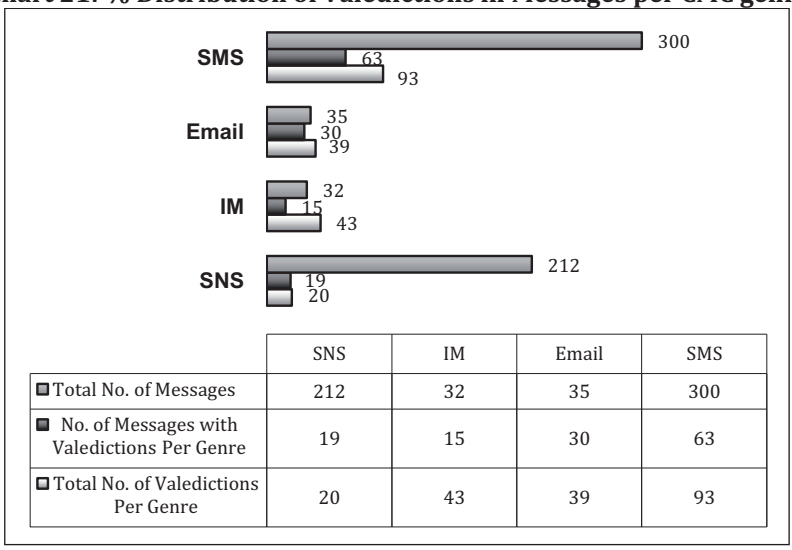
Table 33: Valedictions per Genre

Genre	Total No. of Messages	Total No. of Valedictions per Genre	No. of messages with Valedictions per Genre
SMS	300	93	63 (21%)
Email	35	39	30 (86%)
IM	32	43	15 (47%)
SNS	212	20	19 (9%)
Total	**579**	**195**	**127 (22%)**

The distribution of valedictions in the results shown on table 33 indicate that out of 300 SMS, a total of 93 valedictions were present in 63 messages, 35 Emails yielded a total of 39 valedictions from 30 messages. 32 IMs resulted to 43 valedictions from 15 messages while 212 SNS resulted to 20 valedictions from 19 messages. This demonstrates that the 579 messages contain a total of 195 valedictions from 127 messages. These results are summarised in chart 21. Note that some messages in the genres contain more than one valediction.

As already discussed in the greetings section, IM follows the general conventions of face to face communication.

Chart 21: % Distribution of Valedictions in Messages per CMC genre



	SNS	IM	Email	SMS
Total No. of Messages	212	32	35	300
No. of Messages with Valedictions Per Genre	19	15	30	63
Total No. of Valedictions Per Genre	20	43	39	93

Therefore, the use of valediction as a way of sign-out or leave-taking is prevalent in IM. Users feel compelled to sign out from a chat in a conventional way by giving goodwill wishes or promises to keep in touch. Yet, it is apparent that the number of the IM valediction at 43 is lower than that of greetings at 103 in the same genre. This is explained by the fact that due to frequent poor Internet connections, relatively high connection prices in cyber cafes, and unreliable electricity supply, many IM chats are disconnected prematurely and thus the chatters do not get a chance to sign off conventionally.

The use of valedictions is also common in Email. One of the reasons for this is that the Email is based on the traditional letter which requires a sign out valediction at the end. This practice still holds for most Emails in the data. Another reason for this is the lack of mode limitation for the Email since there is no character limit. Yet in comparison to greetings which are 103, the valedictions in Email at 39 are few. This difference shows that signing-out is not treated as intensely as the greeting. Generally, valedictions are almost brisk and not as multiple as the greetings. Omar (1993:42) who compares greetings and valedictions in Kiswahili

conversations in Zanzibar also asserts this in her conclusion that valedictions are not as heavily constrained as greetings are. This is also the practice in face to face conversations where greetings are more elaborate. Interestingly, both Frehner (2008) and Sproull & Kiesler's (1986) claim that European Email senders use more words for the valediction than for the greeting. Frehner (2008:44) adds that "it seems as if the sender of a message tries to compensate the anonymity and aloofness of the medium with extended leave-taking formulas". In the current study, this observation does not hold true for Emails. Senders are brisk in valedictions and more elaborate in greetings as is the practice in face to face conversations.

Similar to greetings, the number of SMS valedictions is low. One of the reasons for this is the mode limitation. Not much space is left at the end of most messages. Nonetheless, there are short valedictions like *bye* and *c u* which do not take up much space, time or effort. The main explanation for the low use of valedictions therefore is that the inclusion of a valediction is not a convention in SMS. Similar to greetings in SMS, another point to note is that most SMS are quick requests or responses. They are temporary forms of communication and mainly focus on the message. For example, the SMS message in (562) does not necessarily need a valediction. It communicates the message conventionally without sounding impolite. Consequently, the exclusion of valedictions in SMS is a common practice.

(562) *did i leave my keys in your rm?* [SMS]
Did I leave my keys in your room?

As anticipated, SNS has no valediction culture and has the least count of valedictions in the data. One of the reasons for this is that it is more of a continuous comments forum and does not require signing off after posting a comment. This is coupled with the fact that apart from the close knit chat forums that have registered members with a common interest, most of the open forum participants like YouTube are faceless with no social obligations and do not feel compelled to apply social conventions. I observed that most of the SNS messages containing valedictions are from people

who had started threads and now wanted to close them after getting the contributions they needed. This is illustrated in (563). Others are messages in close knit forums where a member is temporarily departing e.g. for holiday or fieldwork as illustrated in (564). Only one SNS contains 2 valedictions. It is from a forum member informing fellow members of his departure from the forum as illustrated in (565).

(563) *Jameni, wacha nikatishe hii thread now that tumeiexhaust kabisa. thnks 4 al ua contributions,* ***laitaz****!*
Friends, let me stop this thread now that we have completely exhausted it. Thanks for all your contributions. **Till later**. [SNS]

(564) **sob* *sob* I was just beginning to enjoy everyone one of you. Just getting to meet new people, touching base and all. But i guess its time for me to shtep[59].*
Yea, am goin to a place bila[60] net (to those who are wondering) probably the most arid place in Kenya. Field work really sux[61] but what to do?!
Anyhow to those i got to know and to those who shared stuff with me, its been real. i enjoyed myself esp. the laughter. we'll do it again when i get back to this life. ***Bye for now****.* [SNS]

(565) *I am off,* ***all the best****, Mashadites. Mungu akipenda* ***"tutaonana" tena****.*
I am off, all the best members of Mashada. God willing, we will see each other (meet) again. [SNS]

Interestingly as shown in the illustrated valedictions, each sender uses the conventional valediction structure analogous to face to face conversation and avoids giving a terminal goodbye by stating that they will be in touch again.

The following are some of the findings and illustrations related to the CMC valedictions in Kenya. Note that the illustrations are only partial excerpts from the message showing the content of the valedictions.

Multiple valedictions in a message and structure of valediction

Some messages contain multiple valedictions in form of pre-

59 *shtep* is derived from the verb step. It means 'take leave' in Engsh.
60 *bila* means 'without' in Kiswahili.
61 *sux* from the word 'sucks'

closing, closing and post-closing elements (Murray 1991). Pre-closing elements like *okay, alright, haya, sawa* etc. are meant to signal a common agreement, or satisfaction prior to the main sign out. In some cases, these elements mark the end of the message like in (566). The following examples are illustrations of messages that contain multiple closing elements.

(566) Pre-closing: ***haya basi*** *(ok then)*
Closing: ***tek care*** *(take care)*
Post-closing or follow-up: *n c u (and see you)* [IM]

(567) *otherwise* ***have a good day****, will keep in touch.****bye regards*** [Email]

(568) ***vulayi mayi****....****take care*** *tooooo much luv from us*
Bukusu (vulayi-bye mayi-lady)
Bye lady, take care -too much love from us. [Email]

Temporary nature of leave taking

Similar to the Kiswahili valediction custom of *kwaheri ya kuonana* translated as *bye see you again*, many CMC valedictions are of temporary nature. It is unconventional to bid someone a permanent valediction. In fact in some cases, the sender is compelled to a promise of more contact soon. Illustrations (563), (564) and (565) are examples of the temporary valedictions. Others include

(569) *bye for now.* [Email]

(570) *tutaonana (Sw)*
We will see each other again. [IM]

(571) *bye* ***Tutabonga****.*
Bye, we will talk. [IM]

(572) *Bye for now.* ***Ntakuhola*** *wen in town.*
nitaku (Sw prefix) + holler
Bye for now. I will call you when I'm in town. [Email]

(573) ***baadaye***
See you later. [Email]

(574) ***Later basi****.*
We will be in touch with you later then. [SMS]

Blessings

Some valedictions in the data are presented as wishes for blessings. These forms are comparable to the English sign out phrase, *all the best* although they are presented with more intensity and not as a cliché.

(575) *barikiwa ushangae (Sw)*
Be blessed until you wonder. [Email]

(576) *b blsd*
Be blessed [SMS]

(577) *baraka tele (Sw)*
Be blessed. [IM]

Prayers

Some valedictions solicit for prayer. The prayer petition can be interpreted as a request to be put in the receiver's thoughts from time to time.

(578) *have a nyce tym n **pray for me**.*
Have a nice time and pray for me. [Email]

(579) *keep me in ua prayerz*
Keep me in your **prayers** [Email]

Proclamations of love

Valedictions are also presented as proclamations of love. This is especially between people in a close relationship e.g. family, close friends or dating or married couples.

(580) *Loads of love* [SMS]

(581) *Miss you and love you lots!* [Email]

Appeals to the receiver to stay in touch

Some valedictions request the receiver to keep in touch for example,

(582) *Stay well and **keep in touch!** Love you and miss you!* [Email]

(583) *take care na* ***usilost***
Take care and don't get lost (disappear). [IM]

(584) ***Hola bak****; n soon.*
Reply soon. [Email]

Greetings to others

Valedictions can also contain requests to greet other people. Although in most cases the people to be greeted are common acquaintances, in some cases they are not. The greeting request is a polite way to acknowledge family or relatives of the receiver.

(585) *mob love,take care.****salimia siz***
Much love, take care and **greet your sister**. [Email]

(586) *bye take care of yourself.* ***say hello to all****.* *[Email]*

Cancelled valedictions

The data also contains some messages with a practice similar to face to face conversations. In this context, the sender gives a valediction and then digresses or goons to surpass it with other information. It is similar to a cancelled valediction for example,

(587) ***have a nice time and be blessed****.hata kaa dint c u in chach leo.*
Have a nice time and be blessed even though i didn't see you in church today. [SMS]

(588) ***gdnite, Salimia watoto****-si sasa ni wakubwa mno?*
Goodnight. Greet the kids- they must be big by now? [IM]

5.1.3. Language Choice for Openings and Closings

Another point worth discussion under salutations is the language choice. It is interesting that in some CMC messages, openings and closings are in a different language than the rest of the message.

(589) ***Hi there****! jana tulikumiss sana! c ukam leo*
(English greeting + Kiswahili and Sheng message)
Hi there, yesterday we missed you very much, why don't you come today. [SMS]

(590) ***sasa, habari ya masiku****??? I am wondering if you have the receipt book coz I*
can't find it...
(Sheng & Kiswahili greeting + Eng message)

Hi, how is life? I am wondering if you have the receipt book because I cant find it. [SMS]

(591) ***Sema;-) efwe balamu ne nywe****? 2takam dis month if all goes well, lakini i wil let* u know exactly nxt wk after my leave days r confirmed. c u
(Kiswahili and Bukusu (vernacular) greeting, while the rest of the message is in English)
Hi, we are well and you? We will come this month if all goes well but I will let you know exactly next week after my leave days are confirmed. See you. [SMS]

(592) *How's that deal of ours?? we'l talk next week when we open colle,* ***lala salama***
(English message + Kiswahili valediction)
How is our deal? we will talk next week when we open college. Good night. [SMS]

(593) ***hi*** *ati mtafika sa?* ***cu***
(English greeting + Kiswahili message + English valediction)
Hi, at what time will you arrive? See you. . [SMS]

(594) ***Hi jim*** *kesho 2naenda nai mpaka sato, c unipitie jioni, nitakuwa 2 kwa hao.* ***L8ter***
(English greeting + Kiswahili message+ English valedictionl)
Hi Jim, tomorrow we will go to Nairobi till Saturday, why don't you pass by my place in the evening? I will just be in the house. See you later. [SMS]

Generally, the data reveals many different languages in salutations. Some forms of openings like greetings and sign-ins found in the corpus include

English

- *Hi, hy*
- *Hello, helo*
- *How r u, ow r u – (how are you?)*
- *Good morning, gud morning, gd morning, morn*

Kiswahili

- *Habari (Hello)*
- *Sema (tell me)*
- *Mambo (how are things?)*
- *Umepotea (You are out of contact)*

Sheng

- *Sasa (hi)*
- *uko poa?(are you well)*
- *gota (a form of hi five using fists)*

Vernacular languages

- *Nono, muriena – (Bukusu - hi, how are you?)*
- *Uga , uhoro, wimwega – (Kikuyu – tell me, your state, are you fine?)*
- *Nang'o – (Luo – how are you?)*
- *Chamge – (Nandi – Hello).*
- *Bwakire – (Kisii – it is light (hello)*
- *Ero – (Masaai – hello)*
- *Uvoo – (Kamba – Hi)*

This establishes that CMC employs many forms of greetings that are in Kenyan languages other than English as summed up in table 34.

Table 34: Use of Languages in Kenyan CMC Greeting

Language	SMS	IM	Email	SNS	Total
English	42	50	39	13	**144**
Kiswahili	21	23	19	5	**68**
Sheng	24	27	25	9	**85**
vernacular	7	3	3	4	**17**
Total	**94**	**103**	**86**	**31**	**314**

Similar to openings, very many forms of closing valedictions are present in the data, including

English

God bless
take care
2moro
ok
2gether
c u - see you
bye
tek kea-take care
Lov u – love you
gudnite – good night

Kiswahili

haya - ok
tuonane - see you
kesho - tomorrow
sawa – ok (Kiswahili)
lala salama – good night
baadaye – later

Sheng

wacha niishie - let me go now
lala poa – sleep well
2ko pa1 (tuko pamoja) – we are together (united)
letaz - later
cheeerz - cheers

Vernacular languages

rucio (Kikuyu) -tomorrow
oriti (Luo) – bye
twigwo na wega (Kikuyu) – bye
bulayi (Bukusu) – bye
kongoi (Nandi) – bye
enda na useo (Kamba) –go well

The findings show that the most commonly used greeting expressions are *hi, sasa,* and *sema* while the most common valedictions are *c u/ see you/ see u* , *goodnight/gudnite/gnite/* and *take care/tek kea*.

Table 35 sums up the raw data of the languages employed in the CMC valediction data.

Table 35: Use of Languages in Kenyan CMC Valediction

Language	SMS	IM	Email	SNS	Total
English	46	19	13	9	**87**
Kiswahili	19	8	10	5	**42**
Sheng	23	14	13	5	**55**
vernacular	5	2	3	1	**11**
Total	**93**	**43**	**39**	**20**	**195**

English is the most used language for salutations. It has a high variety of choices when it comes to lexical compressions, for example the word *good night* (see previous example) can be presented in many ways. This is not readily available in the other languages for example it is not yet conventionalised to use lexical compressions in the other languages including Kiswahili e.g. **kwhr* for *kwaheri* (bye) or *ls* for *lala salama* (good night) in Kiswahili.

The findings in the tables reveal that the use of vernacular languages in salutations is not very popular. One reason is that users could be from different vernacular language groups thus requiring them to use common lingua franca. Another reason as previously explained, vernacular language does not provide choices in terms of language creativity e.g. in lexical compressions.

Sheng greetings on the other hand are more popular due to the students and yuppies data givers. This use of Sheng is favoured as a sign in language because it signals the informal nature of the communication and at the same time acknowledges that the communicators are peers.

Valedictions in Kiswahili are on average. Its use is not as popular as English or Sheng because it does not have a wide variety of choice like English, and that its use would seem formal to the youths and yuppies.

5.1.4. Conclusion

Whether phatic or genuinely informative, greetings and valedictions are a common occurrence in CMC. Their occurrence structure is similar to the salutations in face to face communication. In fact some messages are composed of salu-tations exclusively. This shows the importance of the incorporation of the salutation culture into CMC.

It is conventional for Emails and IM to have salutations unlike in SMS and SNS. This is because the structure of the Email is similar

to the traditional letter which requires a greeting while the structure of the IM is similar to the structure of a face to face chat which requires salutations. SNS has no salutation culture. It has the least count of salutations owing to its nature of continuity which makes it inconvenient for every contributor to sign in and out using salutations. SMS on the other hand do not have many salutations because of their features of quickness, temporariness, and focus on the central message. But interestingly, as earlier mentioned, some SMS messages are coded as a salutation and confirmation of well being e.g. the famous *please call me* which is used to confirm the well being of the sender and as a salutation in general.

Comparably, 195 valedictions in all genres are lower than greetings at 314. This leads me to claim that in CMC people are more compelled to sign in through greetings than to sign off through valedictions. Some of the reasons given for this disparity between greetings and valedictions in IM and Email to some extent include instability in Internet and power connection, and relatively high bills. Additionally valedictions are brisk while greetings are more elaborate. This is comparable to the Kenyan salutation culture.

In regard to language choice, English is the most prevalent language in the CMC data. It is followed by a close ranking between Kiswahili and Sheng. The vernacular languages are the least used for now but this may be a beginning and as a way to familiarise them to CMC.

5.2. Language Choice in Kenyan CMC

To start off, I would like to reiterate that the languages of focus here are English, Kiswahili, Sheng and vernacular languages. The vernacular languages bracket encompasses all the Kenyan indigenous languages. As mentioned previously, Kenya is a multilingual country with over 42 languages. In this section, I present the languages encountered in the general corpus of data. Table 36 shows the general use of the languages in CMC while table 37 splits the language distribution per genre. This is followed by a discussion of this language use in CMC. The data in table 36 on the general

use of the languages is calculated as follows based on the example of the message below.

(595) *hi dear, nang'o? umetuliza videadly aje?? sisi 2ko 2.*
Hi dear how are you? you seem too quiet. we are ok. [SMS]

The general message employs a mix of several languages. I therefore classify it in the codeswitch column. Next, I split the message into phrases/clauses as follows:

hi dear, (English)
nang'o? (Luo vernacular)
umetuliza videadly aje? (Sheng)
sisi tuko tu (Kiswahili)

This results to 4 clauses/sentences with each representing a different language as indicated in the brackets. I further split each phrase/sentence into words as follows:

hi (English)
dear (English)
nang'o? (Luo vernacular)
umetuliza (Kiswahili but used with Sheng meaning-it is classified as Sheng)
videadly (Codeswitch: Swa+Eng)
aje? (Kiswahili)
sisi (Kiswahili)
tuko (Kiswahili)
tu (Kiswahili)

This analysis results in 9 words with the languages indicated in the brackets. Therefore Kiswahili has 4 words, English 2 words, Luo vernacular language and Sheng, 1 word respectively and codeswitch (words involving more than one language) has 1 word.

Table 36: Use of Language in Kenyan CMC data

	English	Kiswahili	Vernacular	Sheng	Codeswitch
Messages 579	161	113	13	_	292
Sentences 2165	856	552	93	_	664
Words 17368	7339	4166	747	1501	3565

Table 37: Distribution of Language use per word across the genres

	English	Kiswahili	Vernacular	Sheng	Codeswitch	Others	Total
SMS	2023	783	277	410	854	0	**4347**
Email	1872	1286	107	329	741	12	**4347**
IM	2021	985	62	305	937	15	**4325**
SNS	1423	1112	301	457	1033	23	**4349**
Total	**7339**	**4166**	**747**	**1501**	**3565**	**50**	**17368**

The category *others* includes words or expressions that could not be classified in any of the listed categories. For example words like *bona notte*, *habibi* and expressions like *hmmmm* and *hehehe.*

The vernacular Languages

Naturally, where CMC interlocutors share the same vernacular language, it is expected that they are likely to use it in their communication. In spite of this, the results indicate that messages in pure vernacular languages are generally used in very few cases in my data. This is in contrast to Hård af Segerstad's (2002:151) finding that her respondents used Swedish (vernacular) most of the times. English was used in few instances and only under 'certain circumstances'. In my data, vernacular languages are mostly used in SNS, then SMS, Email and least used in IM. The most likely reason for their widespread use in SNS is to achieve solidarity through ethnic fora in order to connect and lock out the foreigners who do not understand the vernacular language. For example on YouTube, a video clip performed in a vernacular language or with a certain interest to a given vernacular language is likely to have most comments in the same vernacular language. In some cases, CMC contexts have a vernacular language identity for example

kikuyu.com[62], Bhaluhya Bhefwe forum[63] on Mashada SNS and videos in vernacular languages on YouTube[64] which encourage participants to post their comments in the vernacular language.

In the case of this research, IM has the least use of vernacular languages not only probably because the participants do not share a vernacular language but also because the IM context favours the use of lexical compression that is not available in vernaculars.

According to Milroy and Milroy's (1992) model, standard variants tend to be associated with weak social network ties, while vernacular language variants are associated with strong network ties. Both the standard variants like English and the vernacular language variants in this case are associated with strong network ties. The only distinction is that vernacular languages do not feature as much as the other languages. This has a twofold explanation, the first being that perhaps the interlocutors do not share a vernacular language, and second that vernacular language choice is not popular in CMC communication. As was already explained most vernacular languages are mainly used in speech such that many users are uncomfortable writing them. In addition, CMC keyboards are originally made for international languages. Therefore users are more comfortable using the non vernacular languages and only add in the vernacular languages where necessary for various reasons like identity, accuracy, privacy etc.

However, the explanation that vernacular language is not a popular language in CMC was eroded right from the political unrest that gripped Kenya in form of the post election violence at the end of 2007 and beginning of 2008. Each ethnic group backed one of the two main presidential contenders; Mwai Kibaki of the Party of National Unity (PNU) and Raila Odinga of Orange Democratic Movement (ODM). PNU was mainly backed by ethnic groups from

[62] http://kikuyu.com/eve/forums/a/tpc/f/487108348/m/1231095182

[63] http://www.mashada.com/forums/old-politics-forum-read-only/2756-bhaluhya-befwe-all-luhya-elite-2.html

[64] http://www.youtube.com/watch?v=Qh_FFab-DGk&feature=related

central Kenya with Kikuyu at the forefront. ODM on the other hand was mainly supported by groups in the Nyanza province that is the Luo and Kalenjin from the Rift valley province. Right before the elections, there was a torrent of CMC messages circulating in vernacular language to campaign for the candidates. This went on during the elections with messages going around in vernacular language to monitor and update each other on the counting of the results. Finally after PNU was declared the winner, the inter-ethnic violence erupted. There was a surge of CMC messages, with the SMS in vernacular languages at the forefront in inciting people into violence. This opened an avenue for vernacular languages in CMC which has gained popularity in building strong network ties not only politically but also socially. They are now more prevalent in CMC in Kenya. This shows the important role of language in combination with CMC as a link to generate a socio-political linguistic landscape. Examples of messages in vernacular languages are:

Kikuyu

(596) *Wi mwega? Ndukirathimwo onginya umake.*
Are you fine? Be blessed until you get scared (wonder) [SMS]

Kamba

(597) *Nata?ngoka wa kuambililia,mbai!*
Hi? I will come on Monday. Bye. [SMS]

Kisii

(598) *mbuya mono renta chinge chia Nyakenywa na wamogusii,mwamosioma,masaba etc bamura! zile zama!Gaki*
Hello, bring clips of Kisii nyakenywa, mwanasioma and masaba groups. They are great and remind us of the oldies. Please. [SNS]

Luo

(599) *Wuod nyahera ibiro lakeside city kara 'ngo wamieli live! Yawa uru, ..a good thing is gooduu...Nyaka mama nyuola poka winjo thum ma mit kama!!*
When is the son of Nyahera coming to the lake city we dance to his music live! A good thing is always good...since I was born I have never listened to better music!! [SNS]

Nandi (Kalenjin)

(600) *Nyolu koek tiendab bororiet.....Anyiny nemokimwoee..*
It should be a national anthem, it is so nice. [SNS]

Maasai

(601) *Kemayianisho taa ele singa le Enkai. kimayiani oleng Ole Pakuo.*
This servant of God blesses so much. You are greatly blessed oh, my very dear friend. [SNS]

Much of the other vernacular language data is in form of code-switching as discussed in 5.3. The other languages present in Kenyan CMC are:

Sheng

As explained before, Sheng does not have an independent grammar. The counts are on the basis of words.

(602) *Mabeste walidandwa na karao pale tidys*
These guys were arrested by the cops around Tidys.(Tidys is a restaurant in Nairobi) [SMS]

(603) *vako zako nimezinyita hutoboi na mimi*
I've discovered your nonsense, you won't get away with it. [SMS]

(604) *maze coast nikunare, mafonyifo wananyarwa.*
My dear, there's unrest in Mombasa, prostitutes are being arrested. [SMS]

Kiswahili

A good number of the messages in the corpus are in Kiswahili. Examples include

(605) *Sema! Siku nyingi! umepotelea wapi?? kwani nilikukosea?*
Hi, it's been long. Where have you been? Or did I wrong you? [Email]

(606) *Lala sahii wacha 2ongee kesho*
Go to bed, lets talk in the morning. [IM]

(607) *Eeeeeh wacha maisha ianze sasa;-)*
Yes let life begin now (wink). [SMS]

English

Majority of the messages in the corpus are in English. It is the most commonly used language in CMC. In this case, one of the reasons for this could be that the data givers are university students and yuppies who are comfortable in English. Another reason could

be that the CMC technology is still associated with the western world. Examples of messages in English include

(608) *hi r u stil comin over 2day? If so pls giv me a buz*
Hi are you still coming over today? If so please give me a buzz (call). [SMS]

(609) *just finishd clining wit da last drop of water. headin 2 pik sis. ow waz ur wkend? Travlng 2mr*
I just finished cleaning with the last drop of water. Im heading to pick my sister. How was your weekend? [SMS]

(610) *my cousin was gettin married and her husband hayad them to sing thesong for her, it was soo sweet!*
My cousin was getting married and her husbanc hired them to sing the song for her. It was so sweet. [SNS]

Conclusion

The vernacular languages are not frequently used monolingually in CMC but their popularity is increasing as signalled in code-switching contexts. This is a positive indication that they are getting into the CMC context and that CMC will serve as a way to stimulate their writing.

My findings are similar to Hinrich's (2006:86) and disprove the claim based on Labov's observer's paradox that 'the more consciously a speaker's linguistic resources are deployed, the less likely is the use of the vernacular' (Hinrich 2006:86). My data shows that the use of vernacular languages is by will and as Hinrich explains, 'it is based on the added communicative value'. A good example of this is in SNS fora that employ the wilful use of a vernacular language for identity.

English leads in terms of monolingual messages followed by Kiswahili. These are the lingua francae of Kenya and are the languages that most people are educated in. English also has an advantage because it is the language that is associated with CMC. It is hard to determine Sheng because it has no independent grammar but as proposed earlier, I mainly identify Sheng from its vocabulary.

Another aspect that cannot be ignored is the mixing of languages in the messages. Many messages involve a mix of languages. The presence of language mixing is very common in CMC. This mixing is not only at the message level but also between words in the sentence and within words. This is discussed in the next section of codeswitching in CMC.

5.3. Codeswitching in CMC

Manze Hau ni ma, ebu sema ng'we we show u dust ;-)
Sh vrn Sw Eng
Sure that's it, dare say ng'we and we show you dust [SMS]

Kenya's multilingual nature is concretely captured in the corpus of the CMC data. Not only are there messages in different languages, there are also different languages within messages. This leads to a situation of language contact in CMC which in turn is reflected in codeswitching (CS).

Codeswitching is a field that has fascinated many researchers since the 1950s. Weinreich (1953) defined codeswitching in his description of bilingualism as "the practice of alternately using two languages". Devic (2007) contends that this is what one would call the broad definition, which stops at indicating the juxtaposition of elements from two or more languages or dialects. Devic continues that in the course of the development of CS theories, each researcher added some details to Weinreich's definition, thereby specifying it for their own purposes. This led to disagreements for example on the length of the juxtaposed utterance; whether a whole discourse, a single word, maybe even containing morphemes from two languages and the density of switches in a given spoken or written text, whether the switch is an instance of an individual type or rather of a common type, whether or not it bears social significance, and if so, of what nature (cf. McCormick 2001).

Myers-Scotton (1993a) presents an explanation of different levels of codeswitching which she labelled as *intersentential codeswitch-*

ing and *intrasentential codeswitching*. She defines intersentential CS as occurring at sentence or clause boundaries, while intrasentential CS appears within a clause and may take the form of a phrase, a single word, or a morpheme (cf. Mahootian 2006). She does not approve of the use of the term *intraword codeswitching* and suggests that it should fall under the intrasentential codeswitching. Myers-Scotton's research data was in form of spoken discourse. In contrast my data is in form of textual CMC messages. Most of these messages are composed under the influence of factors like urgency, limited space, least effort and creativity. For example, some of these texts contain only one word while others only contain a graphic, Emoticon or only a single punctuation. For my data, I find it necessary to make a distinction between codeswitching within the word and across words and phrases. I propose to adapt the terms *interword codeswitching* for the latter and *intraword codeswitching* for the former distinction. The term intra-word CS is adapted from McArthur (1998) who describes it as occurring when a change takes place within a word, such as *kuenjoy* (English enjoy with the Kiswahili infinitive prefix *ku*). The word starts off with a Kiswahili infinitive prefix but suddenly changes into an English verb. This is what I refer to as intraword CS. The nature os Sheng forces me to adopt this distinction as I explain below and earlier in 2.1.4.

Illustrations (611) and (612) demonstrate the use of the different languages within the same message (interword codeswitching) while (613) and (614) illustrate the use of different languages in the same word (intraword CS).

(611) *pls kuja speed kabisa*
Eng Sw Eng Sw
Please come quickly. [SMS]

(612) *Haki dont hepa our mpango pls*
Sw Eng Sw Eng Sw Eng
Please don't evade our plan. [Email]

(613) *tulimuexplainia*
tulimu+explain+nia.
Sw Eng Sw
We explained to him/her. [IM]

(614) *they hatad kuhata*
they hata –d ku- hata
Eng Sh Eng Sw Sh
They failed completely. [IM]

One reason why the term intrasentential CS is complicated for the current use is because in CMC, sentential boundaries are not always clear. It is therefore clearer to identify CS at the word level in this research. For example in the message

(615) *yaya sasa umelost sana my dear!!!*
vrn Sh Sw pr+Eng Vrb Sw Eng
Hi brother, how are you? You have disappeared (lost). I don't hear from you my dear.
[SMS]

Basic concepts in CS are the distinction of the Matrix Language (ML) and embedded items. This has been amply discussed by Myers-Scotton (1993) under the Matrix Language Framework (MLF). The generalizations of this approach are that, given a code-switching context, one of the languages involved must play a dominant role. This language is labelled as the Matrix Language (ML). Its grammar sets the morphosyntactic structure of the construction. The other language(s) involved is labelled as the Embedded Language (EL). Therefore in the Matrix Language Hypothesis, the ML sets the morphosyntactic frame for ML+EL constituents. Myers-Scotton continues that this hypothesis is realised as two testable principles[65] namely;

- *Morpheme Order Principle* where the morphemes must not violate ML morpheme order
- *System Morpheme principle* where all syntactically relevant morphemes must come from the ML.

To begin with, some of my data clearly fits the principles in that intraword codeswitching occurs in clauses with English or Kiswahili as the Matrix Language. For example in (616) the whole clause is in a vernacular language (Kikuyu) except for the English root

[65] Myers-Scotton (1993a:83) cf. Overview at http://yayoi.senri.ed.jp/research/re09/namba.pdf

camp which has the Kikuyu prefix *gu*.

(616) *mwadhani ndirafanga **gucamp** kwao*
gu- camp
God, I plan to camp at their place. [SMS]

It is clear that the ML in this message is Kikuyu. Only the word *gu-camp* is presented as a mix of two languages. Therefore according to the definition adapted, the intraword CS is between the Kikuyu (vrn) infinitive *gu* and the English verb *camp*

Nonetheless, these principles present some complexity in analysing the Sheng and Engsh data. In some of these cases, establishing a Matrix Language (ML) becomes impossible. One of the reasons to this is that there are markers of English origin that are not lexical and would count as English grammar making the ML English whereas their function is not always grammatical but rather to render the language use Engsh e.g. note the double plural in

(617) *watejas :* *wa + teja + s*
Sw pl + unreachable + Eng pl
Unreachable by phone. [SMS]

In the case above, the English plural marker *–s* is not used to mark the plural but rather to transform the word into Engsh. It is then futile to categorise it as the pointing to the ML.

Similar to the intraword CS, the interword codeswitching data are interesting because they challenge CS theories in the sense that it is sometimes difficult to determine the matrix language and sometimes contradictory examples are encountered in the grammatical morphemes from two languages in a word.
For example,

(618) *we agreed nitazipik then nikusendie*
ni – ta – zi pik then ni – ku- send- ie
SWA1SG FUT- them ENG Verb root SWA1SG- you-ENG Verb root- Applicative Suffix
We agreed that I will pick and send them to you. [SMS]

In illustration (618), all the grammatical morphemes are from

Kiswahili while all the main vocabulary is in English. Following the MLF, the conclusion is that the ML for this message is Kiswahili.

(619) *ulikam lst wknd ama u rukad?*
u- li- kam lst wknd ama u ruka - d?
SW2SG - PAST- ENG Verb root ShVerb root - ENG PAST

Did you come last weekend? Or you changed you mind? (The basic meaning of *ruka* is 'jump'. In this example, it us used figuratively to mean, 'changed your mind'). [SMS]

Illustration (619) is a bit more complex. The message incorporates grammatical morphemes from both Kiswahili and English. This makes it difficult to identify the ML basing on the MLF profile because the ML in the first clause is Kiswahili but all lexical vocabulary is in English. In the second clause, the verb is Kiswahili but with an English past tense morpheme and a Kiswahili subject pronominal *u*. This mixed construction contains contradicting indications for ML.

Therefore I only discuss CS at the word level. Sheng or Engsh is not analysed at the sentence or message level but only at the word level where the morphological attachments from the other languages attached to the Sheng roots can be identified or where the individual lexeme is clearly Sheng (cf. table 1). Additionally the interest here is not in the ML but in the mixing of languages. I therefore do not analyse the ML in the messages.

Another point worth noting is that the adapted distinction between inter and intra word CS is not relevant for CS studies. In this case, I only use it as an approach to enable the clear identification and counting of the different languages used between and within words. The approach simplifies the complications of Sheng and Engsh being codes based on CS. This means that I will mainly concentrate on the language use at the word level and in turn identify Sheng and Engsh through their exclusive vocabulary. All the other constructions involving more than one language will be classified as cases of CS. (cf. section 2.1.4 for a more detailed explanation on the distinction between Sheng/Engsh and CS).

5.4. Interword Codeswitching

kwani results zitatokea when???
Sw Eng Sw Eng
But when will the results be released? [IM]

As already stated, for the purpose of this research, interword codeswitching is the occurrence of words in different languages embedded within a single CMC message. Table 38 summarises the number of messages that contain codeswitching in each genre.

Table 38: No. of Messages with Codeswitching per Genre

Genre	Total No. of Messages per genre	No. of Messages with CS	% of Messages with CS per genre
IM	32	14	44%
Email	35	31	89%
SMS	300	186	62%
SNS	212	110	52%
Total	579	341	59%

This table has been compiled by singling out only the messages that display an occurrence of interword CS. These occurrences are not counted per sentence but per language combination in the whole message. As a result, if an IM message has 10 occurrences of interword CS, it is only counted once and identified through the CS language combination. Similarly, if an Email message has only one instance of CS, it is also counted only once. This means that having a low count of interword CS messages in a genre does not necessarily mean that the genre has the least occurrences. There could be many occurrences which are registered only once.

It is remarkable that in comparison with the rest of the genres, less than half IM contain CS, which is 14 out of 32 IM. On the other hand, over half of SMS and SNS contain codeswitching (cf. 5.7). Emails have the highest presence of codeswitching. This is the opposite from the findings of Hård af Segerstad (2006:258) whose data registered Emails with no use of CS. The explanation of this was that her (European) Email data were from official Emails and

they were of a formal nature and thus used of only one language. The high count of CS in Emails in my data can be explained by the fact that owing to the unlimited length, Emails are relatively long and this widens the chances of inserting a codeswitch. It appears that the length of the message may be a predictor of the occurrence of a CS especially in genres without mode limitation. The Email data supports this because the Email messages are long and contain a high number of CS unlike the other genres which are relatively shorter. It may be argued that the IM genre has long messages but the structure of IM is like a real time conversation with each turn having a few lines at most. Emails are longer and free flowing. Another explanation for the high CS presence in Emails is related to informal communication among peers.

The data in the corpus reveals 341 (59%) messages with instances of codeswitching. This is over half of the analysed messages although it is not surprising owing to the multi-lingual nature of the participants in addition to the motivations from the CMC (cf. 5.7). Other CMC studies have also reported the frequent use of codeswitching in CMC texts. These studies include Bodomo (2009:203) who reports the codeswitching between Cantonese and English in CMC texts of the bilingual students in Hong Kong. He also reports the codeswitching between English and French among the French youth on Facebook. Hård af Segerstad (2001:151) also reports that in her web chat data, codeswitching was sometimes used with English phrases embedded in Swedish utterances.

Table 39-41 show a more detailed breakdown of language combinations in the codeswitching context. These combinations only count the occurrence of the set of the languages in each message.

Table 39: Multilingual Interword Codeswitching

Language Combinations	SMS	Email	IM	SNS	Total No. of Messages with this combination
Eng + Sw+ Sh + vrn	14	3	3	8	28

The actual language combination structure is not necessarily in the order shown in the table.

The CMC data only has 28 messages (8%) exhibiting interword codeswitching involving all the language under study as shown in table 39. Illustrations (620) and (621) contain examples of such cases. Table 40 shows trilingual codeswitching language combinations in each genre. These results in table 40 reveal that out of the 341 messages, 127 (37%) display instances of trilingual codeswitching. Distinctively, most cases of codeswitching revolve around combinations of English, Kiswahili, and Sheng (cf. illustrations (622) and (623) for trilingual CS examples).

Table 40: Trilingual Interword Codeswitching

Language Combinations	SMS	Email	IM	SNS	Total No. of messages with this combination
Eng + Sw + Sh	30	6	2	22	60 (47%)
Eng + Sh + vrn	14	2	1	9	26 (21%)
Sw + Sh + vrn	13	1	1	8	23 (18%)
Sw + Eng + vrn	9	3	1	5	18 (14%)
Total No. of messages	66	12	5	44	127

Table 41: Bilingual Interword Codeswitching

Language Combinations	SMS	Email	IM	SNS	Total No. of messages with this combination
Eng + Sw	32	5	1	20	58 (31%)
Eng + Sh	28	6	1	13	48 (26%)
Eng + vrn	2	2	1	4	9 (5%)
Sw + Sh	38	2	1	10	51 (27%)
Sw + vrn	7	1	1	6	15 (8%)
Sh + vrn	2	0	1	2	5 (3%)
Total No. of messages	109	16	6	55	186

Table 41 shows bilingual combinations of codeswitching. These combinations are in 186 (55%) messages in the CS data. Just like in trilingual CS, English + Kiswahili or Sheng are widely used. See illustrations (624) and (625).

I will now proceed to present illustrations of the codeswitching from the corpus of data. The involved languages are abbreviated underneath the words in the message.

As already mentioned, illustrations of codeswitching in (620) and (621) show cases where all the four codes are used i.e. English, Kiswahili, Sheng and vernacular languages. The data corpus reveals many examples of these occurrences.

(620) *Hey!!!!!!!! sema! How are u? Niki kura uguo?salamu tu! Nimekuhata mbaya.*
Eng Sw Eng vrn (Kikuyu) Sw Sh
Hey, tell me, how are you? Why have you disappeared? I'm just saying hi. I've missed you badly. [SMS]

(621) *r u free leo? rauka ukam khuche mapema*
Eng Sw Sh Sw+Eng vrn (Luhyia) Sw
Are you free today? Come very early so that we go early. [SMS]

(622) *kumethoka! gun shots kila mahali gai!*
vrn (kikuyu) Eng Sw vrn (kikuyu)
It's total chaos! gunshots are everywhere. [SMS]

(623) *waaaat??????????? wacha wewe! lakini wemwega? tigana nake kwani atado? mwere athie kou. funda eno* ☺
waaaat??????????? wacha wewe! lakini wemwega? tigana nake
Eng Sw vrn (Kikuyu)
kwani atado? mwere athie kou. funda eno ☺
Sw vrn (Kikuyu)
What? You can't be serious! but are you fine? Leave him alone, what is the worst he can do? Tell him to go away, he is such a donkey. [Email]

(624) *Kama this ndiyo lyf ya couple beta kuishi solo*
Sw Eng Sw Eng Sw Eng Sw Eng
If this is the way a couple's life is, its better to stay single. [SNS]

(625) *come twende!!!!*
Eng Swa
Come, let's go. [SMS]

(626) *ati who???*
Sw Eng
That who? [IM]

5.5. Intraword Codeswitching

Enda umshow sitamake kucome ,coz niko na kajob
Go and tell her that I will not make it because I have a small unfinished task (job). [SMS]

In the following discussion, I adapt the term 'intraword codeswitching' to refer to codeswitching data at the word internal level. Based on this, intraword CS defines the occurrence of more than one language within the same word. The corpus shows that 312 (54%) of the messages exhibit words with intraword codeswitching. The number of words and language combinations are shown in table 42 and table 43. These counts are of words which have been extracted from the messages. In some cases, a single message contains several different cases of intraword codeswitching e.g.

(627) Davy ***nunuad*** *mingi buz 4 de* ***jamaaz*** *mpaka* ***wakamsare****!! yani* ***aliwasho*** *adabu....the way* ***niwalianga****!!!*

Davy ***nunuad*** *mingi buz 4 de* ***jamaaz*** *mpaka wakamsare!! yani*
nunua-d *jamaa-s*
Sw vrb-Eng pst *Sh N-Eng Pl*
aliwasho *adabu....the way* ***niwalianga****!!!*
ali- wa-show ni-wa-lianga
Sw pst- pl-Eng vrb Sw Pr-Sw pl-Buk Vrb

Davy bought a lot of alcohol for the guys till they gave up despite their greediness! He taught them a lesson from the way I saw them! [Email]

Each of these is counted separately such that,
nunuad - Sw+Eng (bought)
jamaaz - Sh+Eng (guys)
aliwasho - Sw+Eng (showed them)
niwalianga - Sw+vrn (they are greedy)

Table 42 and 43 indicate that English and Kiswahili are the main combinations used in intraword codeswitching. The next frequent combination is of English and Sheng which is then closely followed by Kiswahili and Sheng.

Table 42: Intraword codeswitching language Combinations

Lang. Combinations	SMS	Email	IM	SNS	Total
Eng + Sw	336	298	313	307	1254
Eng + Sh	392	126	177	280	975
Eng + Vrn	79	97	19	58	253
Sw + Sh	278	117	142	293	830
Sw + Vrn	53	30	12	41	136
Sh + Vrn	14	21	31	39	105
Total No. of Words	**1152**	**689**	**694**	**1018**	**3553**

Table 43: No. of messages with language Combinations

Language Combinations	**Total No. of messages with combination**	**Percentage %**
Eng + Sw	110	35%
Eng + Sh	57	18%
Eng + vrn	29	10%
Sw + Sh	73	24%
Sw + vrn	23	7%
Sh + vrn	20	6%
Total	**312**	**100%**

The vernacular language category is the least represented in the intraword codeswitching process. One of the most likely reasons is that vernacular languages are only getting into the written media and need more time for many people to get used to writing them. Another possible reason is linked to the different vernacular languages that exist in Kenya. It is probable that in some cases, CMC users may not use a vernacular language simply because they do not share the same vernacular language. In such cases, they resort to the other possibilities.

This section presents the different kinds of intraword codeswitching as represented in the data. I will begin with a discussion of inflections on verbs, nouns and finally adjectives and adverbs.

5.5.1. Intraword codeswitching in Verbs

Affixes conjoined to verbal roots are the main foundation of intraword codeswitching. It is therefore vital to dissect them when tackling the issue of codeswitching in CMC.

Table 44: Distribution of affixes in Intraword CS

Genre	Eng Affix	Sw Affix	Vern Affix	Total
IM	380	519	38	937
Email	301	398	42	741
SMS	352	447	55	854
SNS	405	579	49	1033
Total	**1438**	**1943**	**184**	**3565**

The language distribution in terms of affixes is presented in table 44. As the results point to the fact that Kiswahili morphology is the most commonly used in intraword codeswitching. It is not surprising that English affixes are few; this conforms to the English language scarcity in affixes in comparison with Kiswahili.

The grammatical morphology of vernacular languages is scarcely used while Sheng being a language parasitic onto Kiswahili grammar does not have its own grammatical morphology. In regard to affixes from vernacular languages, it is worth noting that some of these vernacular languages e.g. the Bantu languages have a similar grammatical morphology to Kiswahili and can be indistinguishable from it. For example in illustration (628), the word *kumuserve* contains intraword CS where *ku* is an infinitive and *mu* is an object prefix in both Kikuyu and Kiswahili. It may be difficult to categorise the CS in such cases.

(628) *Ngai Akurathime na uthie nabere kumuserve, ogwo*
ku-mu-serve
to-him-serve
God bless you and go ahead and serve him like that. [SNS]

In a bid to manage this challenge, I considered the general linguistic context of the message before categorising the affixes. In this way, I categorised the CS above as between Kikuyu and English because the message context is Kikuyu.

In comparison with the analytic approach of English, Kiswahili agglutinative morphology is more complex. The next section describes the occurrence of Kiswahili affixes in intraword CS. I begin by examining the use of Kiswahili affixation on English verbal roots, and then Kiswahili affixation on English nouns and adjectives. I will make an attempt to detect the Matrix Language where possible.

5.5.2. Kiswahili Inflectional Prefixation

This inflectional prefixation involves the attachment of Kiswahili prefix(es) onto English verbal roots. For example,

*(629) Kwani **ungemadvise** abt clad za **kuchoose**.*

u-nge-m-advise **ku-choose**
you-would-him-advise to-choose
Would you have advised him about the clothes to choose? [SMS]

nge is the tense marker of the present conditional in Kiswahili.

(630) Ukinitrespass nitakudelamere!!!!!!!!

u-ki-ni-trespass ni-ta-ku-delamere
you-if-I-trespass I-will-you -delamere[66]
If you cross my path, I will punish you. [SNS]

Illustrations (629) and (630) show intraword codeswitching between Kiswahili and English. The prefixes are coupled with three English verbs, *advise* and *choose* and *trespass.*

*(631) The drama ws 2 much,cldnt stand it n left them **wajisort**.*
wa-ji-sort
they-themselves-sort
The drama was too much I couldn't stand it and left them to sort themselves out.
[SMS]

[66] 'delamere' originates from a manslaughter case where trespassers/poachers were shot . cf. http://en.wikipedia.org/wiki/Thomas_P._G._Cholmondeley

In illustration (631) the transitive verb *sort* takes on Kiswahili inflections.

*(632) I saw ur Sms late hence **sikurespond.Utakam** lini?*

si-ku-respond u-ta-come .
I did not-neg.pst-respond you-will-come
I saw your SMS late hence I did not respond. When will you come? [SMS]

Message (632) has two parts whereby the first part has English as the matrix language similar to (631) while the second part has Kiswahili.

(633) ***ashasign mtado****?????*
a-sha-sign m-ta-do
he-has-sign you-will-do
He has already signed. What will you do? (sarcastic). [SNS]

In illustration (633), both the verbal roots are English but the prefixes are Kiswahili.

*(634) she says hi n **anakumis videadly***
she says hi n [**anakumis videadly**]
a-na-ku-miss vi - deadly
she-is-you-miss Sw. adv manner
She says hi and that she misses you very much. [SMS]

The matrix language in (634) is English in the first clause and Kiswahili in the second but with both lexical roots from English.

*(635) hi. dint hav much crdt bt op hizo **zitakupush**.*
hi. dint hav much crdt bt op [**hizo zitakupush**].
zi-ta-ku-push
those they-will-you-push
Hi. I didn't have much airtime but I hope that will push you through. [SMS]

The interesting occurrence in message (635) is that although the matrix language of the main clause is English, the Kiswahili plural demonstratives *hizo* has been brought in to support the *zi* plural prefix so that it can refer back to *hizo*. As a consequence, all grammatical morphemes in the embedded clause are from Kiswahili. Note that in Kiswahili such demonstratives act as 'already

mentioned items' which can be referred back to using certain morphemes. This seems to be a characteristic of CS involving Kiswahili demonstratives as seen in further examples like in (642) *hizo,* (301) *zako*, and (688) *zile*. These plural demonstratives in Kiswahili are introduced in order to be referred back to, even if the rest of the sentence is in English.

It is a common phenomenon to have Kiswahili inflectional prefixes added to an English root. This happens not only when the rest of the sentence is in Kiswahili but also in English sentences even if the matrix language in most of the sentence is English. This is demonstrated in (631) and (632). Some clauses show a distribution in which grammatical morphemes and lexical roots are consistently from two different languages i.e. Kiswahili grammar and English lexemes.

5.5.3. Kiswahili Derivational Suffixes

Derivational suffixes in Kiswahili, as in many other Bantu languages, can be extremely productive. The most frequent (and productive) are the causative, passive, stative, applicative, and reciprocal. A verb can carry several derivational suffixes, which in Kiswahili must appear in a specific order after the verb stem. Some suffixes (such as the causative and applicative) raise the valency of the verb, adding an argument. Others absorb an argument (for example the passive, reciprocal and stative), while others, such as the reversive, leave the number of arguments unchanged (Ashton 1944:209ff, Polomé 1967:103ff, Schadeberg 1992:10ff).

In this section I will discuss these derivational suffixes on verbs beginning with the -*i*-*a* applicative, -*iw*-*a* passive, -*ik*-*a* stative, -*an*-*a* reciprocal and -*ish*-*a* causative.

The applicative suffix –*i*-*a*

The first suffix under discussion is the -*i*-*a* suffix which is the Kiswahili applicative marker. This applicative introduces an additional argument into the construction (Deen 2005). The data shows the applicative suffix attaching onto English verbs in the

derivational processes. Note that once derived, the verbs adapt the regular inflectional final vowel *–a*. Examples of this derivational process can be seen in applicative *–i-a* like in the following illustrations (636) to (640).

(636) *Jst come I will* ***buyia*** *u 1*
buy-i-a
Just come, I will buy you one. [SMS]

(637) *Next tym try n* ***smylia*** *camera!!*
smile-i-a
Next time try and smile for the camera. [IM]

(638) *wana****2asesia*** *wapi? nimemaka!*
wa-na-tu-asess-i-a
Where are they assessing us from? I am scared. (Nimemaka is from Kikuyu). [SMS]

(639) *2takukamia saa 3 morn-c u*
tutaku-kam-i-a
We will come for you at 9 am. [SMS]

(640) *relax...nitaku****sendia*** *kesho*
ni-ta-ku-send-i-a
Relax, I will send it to you tomorrow. [Email]

The Kiswahili suffix *–i-e* appears in message (641) and (642). It is used as the applicative. The *e* marks the optative or subjunctive mood.

(641) *ni****kamie*** *na ile cd*
ni-come-i-e
me-come for-optative
Please bring me the cd. [SMS]

(642) *si u2****chekie*** *hizo kago*
u-tu-check-i-e
Won't you check for us that cargo (those things). [SMS]

The passive suffix *–iw-a*

The data presents the Kiswahili passive suffix as very common in the derivation of English verbs. The passive term refers to instances where the subject of the message (part of) is the theme of the action. In Kiswahili, the suffix is marked by *–iw-a*. Here, I

present illustrations of messages showing the attachment of this suffix onto English verbs.

(643) *heeey guess wat! ume**shortlistiwa**!!!*
u-me-**short-list –iw-a**
Hey guess what? You have been short listed. [SMS]

(644) ***aliadmitiwa** lst wk*
a-li-**admit-iw-a**
He was admitted last week. [IM]

(645) *kumbe ati walikopiana exam sa **wamesuspendiwa***
wa-me-suspend-iw-a
Apparently, they copied the exam from each other, so they have been suspended.
[Email]

In (643), the verbal root is actually a compound verb in English *shortlist*. It takes on 2 Kiswahili prefixes *u (Subject prefix) + me* (perfect tense) and the passive suffix *-iwa*. Illustration (644) also follows the same order with the English verb *admit* attached to the Kiswahili prefixes *a-* (Subject prefix) and *–li-* (past tense) and the passive *–iwa.* The matrix language for illustrations (643) and (644) is not clear while it is Kiswahili in (645).

The stative suffix *–ik-a*

This stative suffix can take the form of *–ik-a or –ek-a*. In its canonical use it applies to a transitive change-of-state verb. The agent of the base verb is suppressed, and the object of the base verb becomes the subject. The result generally receives a stative interpretation e.g. *shon-ek-a* from *shona* (to stitch). The meaning of *shoneka* can be represented as *it is able to be stitched.* Some illustrations from the corpus show the attachment of this suffix onto English verbal roots. For example,

(646) *finally ile doc **imeprintika***
i-me-**print-ik-a**
Finally the document has become printed. [SMS]

(647) *Imagine that jeans of mine **ilistrechika***
i-li-**stretch-ik-a**
Imagine that jeans of mine got stretched (expanded).
[Email]

(648) *nime**boeka***
nime-**bore-ek-a**
I am bored. [IM]

The reciprocal suffix *–an-a*

This Kiswahili suffix marks the reciprocity of the action in the verb. The data shows illustrations where it is attached to English verbs. For example,

(649) *wana**stressiana sana***
stress-i-**an-a**
They are stressing each other a lot. [Email]

(650) *c 2kipiane in touch*
Si tu-keep-i-**an-e**
Let us keep in touch (*Note that –e is a variant of –a*). [SMS]

The causative *–ish-a*

The causative derivation in Kiswahili is marked by the addition of the suffix: *-isha* at the end of the verb. It expresses causation in the verb. It attaches to foreign elements and acts as a verbaliser. The data below illustrates some instances of the Kiswahili causative suffix *–ish-a* on English verbs.

(651) *nime**bukisha** nyamchom??*
nime-**book-isha**
I've made a booking for the roast meat. [SMS]

(652) *wame**stopisha** program!!*
wa-me-stop-isha
They have stopped the program. [SMS]

(653) *Hebu **speedisha***
speed-isha
Please, hurry it up. [SMS]

Kiswahili prefixes on foreign adjectives and adverbs

Other parts of speech like adjectives and adverbs are also found in intraword codeswitching although to a lesser degree. This scarcity of adjectives is in agreement with Ogechi's (2005) claim that most

Bantu languages, including Kiswahili have few adjectives. The following are some examples from the data.

(654) *dey r uzaring vicheap jst go now!!;-)*
uza-ring **vi-cheap**
They are selling cheaply just go now. [SMS]

(655) *ana **kafon kapoa***
ka-phone ka-poa
He has a nice little phone. [IM]

In (655) *ka* is used to mark the small size mostly as a diminutive. but in this case, it is rather used as an affective marker: the smaller the cell phone the more attractive it is. The second *ka* leads to adjectival agreement with p*oa* which means *nice* in Sheng while in standard Kiswahili, it translates to *cool*, or *relax*.

5.5.4. English Affixation

The data has many instances of English affixation used for intra-word CS. Besides their various functions in regard to grammar, the affixes also are identified as markers of Engsh as described in the discussion in 5.5.6. I will begin the description of these by first looking at the prefixation which is scarce and then move onto suffixation where the majority of the affixes are to be found.

The English prefixes used are very few.

(656) *They must **rejenga** the floor!*
re-jenga
They must rebuilt the floor. [IM]

(657) *huyo dem alikuwa **muextrasupuu** au ni juu ya ulevi?*
mu-extra-supuu
That girl was extra-attractive, or is it because I was drunk? [SNS]

In (657) the Kiswahili prefix *m(u)* is used to introduce the English intensifier *extra*. This is followed by the Sheng adjective *supuu* 'attractive'.

5.5.6. Affixation in Sheng

As already stated, Sheng/Engsh mainly use Kiswahili or English

affixes. This is illustrated in (658). In illustrations (658)-(662), the initial examples labelled (a) feature Sheng with Kiswahili morphemes. The examples in (b) are the Engsh counterparts made up from equivalent constructions with English morphemes.

(658) a *yangu **ilisanywa** lst wk.*
i-li-sany –wa
Mine was stolen last week. [SMS]

(658) b *mine was **sanywad** last week*
sanywa-d

Sanywa is Kiswahili passive for *collected*. However, in this context, it is used as Sheng for *stolen*. In (658a), the message contains Kiswahili (*yangu*), Sheng (*sanywa*) and English (*last week*). Yet, the matrix language is Kiswahili because of the Kiswahili morphemes on the verbal root. In (b) the same *sanywa* takes on the English past tense suffix *–d* rendering the utterance Engsh.

(659) a *Mabeste **walikinda** nganya majuzi*
wa-li-kinda
Our friends sold their vehicle recently. [SNS]

(659) b *Mabeste **kindad** the nganya majuzi*
kinda-d

The Sheng word *kinda* means *to sell*. In (659a) this Sheng word has Kiswahili prefixes. In (b) it is converted to Engsh by attaching the English suffix *-d* to it.

(660) a *Do u still hold tha plan ama **ulisare**?*
u-li-sare?
Do you still hold the plan or you gave up? [IM]

(660) b *Do u still hold tha plan ama you **sared**?*
sare-d

In illustrations (660) *sare* is a Sheng word that means to *abandon* or *quit,* it could also mean acquiring something free of charge i.e. without paying. It has Kiswahili prefixes in (a) and an English suffix in (b). Similarly in (661a) the Sheng noun *mbuyu (man, father)* takes on a Kiswahili plural prefix *ma* and the verb *ishia (go, leave)*

is also attached to the Kiswahili prefixes *wa-li*. In (b) the noun *mbuyu* is attached to the English plural suffix *-s* and the verb *ishia* is attached to the English past tense suffix *-d*.

(661) *a* ***mambuyu*** *wote* ***waliishia*** *meeting ya security*
ma-mbuyu **wa-li-ishia**
All men (fathers) went to the meeting to plan security. [IM]

(661) b *all* ***mbuyus ishiad*** *kwa meeting ya security*
mbuyu-s **ishia-d**

(662) *a* ***umenichizisha*** *vibaya*
u-me-ni-chizi-sha
You have made me an idiot (fooled me). [SNS]

(662) b you've **chizishad** me badly
chizisha-d

In (662) the Kiswahili prefixes and suffixes attach to the Sheng verbal root *chizi*. The final vowel of *chizi* is *i* while the causative *ish* contains *i* as the initial letter. One of the *i*'s is ellipted to avoid the double vowels as would appear in *chizi+isha* from Sheng.

5.5.6. Affixation in Engsh

In this section, I provide a detailed discussion of the suffixes that get attached to the lexemes to form Engsh as observed in the CMC data (cf. 2.1.4 for details on Engsh). These suffixes appear as a form of grammar that marks Engsh. I therefore present them in the form that I refer to as the grammar of Engsh.

5.5.7. Grammar in Engsh

The corpus shows some verbs from English, Kiswahili, Sheng and even codemixed verbs e.g. (671) and (672) with Engsh grammar. This is reflected in tense and aspect markers *-d* past, *-(r)ing* progressive and *-s/(z)* habitual similar to English. An important function of these English tense and aspect markers is to mark the expression as being Engsh. In the next subsections, I discuss each of the tense and aspect markers.

- **Past Tense Suffix -d**

In the following illustrations, it seems that this suffix *–d* mainly occurs in constructions whose matrix language is English although there are some exceptions like in (667) where the adverb *vikrezy* the phrase that has Kiswahili as the matrix language.

(663) *Lulu was **ambukizwad** that habit of carefreeness*
ambukizwa-d
Lulu was infected with the habit of being care-free. [IM]

(664) *I already know wat **fanyikad** on that day*
fanyika-d
I already know what happened on that day. [Email]

(665) *U even **lengad** the advice we gave you*
lenga-d
You even ignored the advice we gave you. [SNS]

(666) *mum **kaziad** us to go*
kazia-d
Mum refused us to go. [SMS]

(667) *iv bn **changanyikiwad** vikrezy sana wacha 2*
c**hanganyikiwa-d** vi-crazy very wacha just
I have been mixed up like crazy. [SMS]

(668) *the 2 lorries **ingianad** then **lipukad***
ingi-an-a-d lipuka -d
The two lorries rammed into each other and then exploded. [SNS]

However, there are other Kiswahili verbs with foreign origin which have invariant final vowels that the suffix attaches itself to. An example of this is the verb *jaribud* in (669).

(669) *I **jaribud** calling the whole day*
jaribu-d
I tried calling the whole day. [SMS]

(670) *alikuwa **hangid** ka mwenda*
hangi-d
He had a hangover like a mad person. [IM]

Illustrations (669) and (670) prove that the English past tense *–d* does not only attach to verbal roots with the final vowel *–a*. It is a fact that most Kiswahili verbs have a final vowel *–a*, but there are some verbs that are of foreign origin e.g. verbs from Arabic that do not necessarily end in - *a*.

-Sw-Eng (Verb-ishwa-d)

Interestingly, the data also displays some occurrences of verbs with the English past tense suffix *-d* preceded by the Kiswahili derivational passive causative suffix *-ishwa* to result in *ishwa-d*. This displays instances where the suffix is used to mark the word as Engsh. For example,

(671) *Imagin he was **repeatishwad** the DL test!!!!:-((*
repeat-ish-w-a-d
repeat-causative-passive-d
Imagine he was made to repeat the Driving Test. [Email]

(672) *Hahaha Iv bn **hangishwad** mpaka I just laughed in shck*
hang-ish-wa-d
hang-causative-passive-d
I've been ignored until I laughed in shock. [SMS]

It is clear from these examples that the *-d* suffix attaches itself to Sheng and Kiswahili and even to Kiswahili verbs including those of foreign origin.

- **English Progressive suffix -ing**

The suffix *-ing* is also used in a similar way as the following illustrations show.

(673) *i'l be **kujaing** those syds 5th feb*
kuja-ing
I will be coming that way on 5th February. [IM]

(674) *then we can b **ngonyaing**;-)*
ngonya-ing
Then we can be gossiping. [IM]

In example (673), the suffix *-ing* is attached to *kuja*, the Kiswahili verb stem including the inflectional final *-a* while in (674) the suf-

fix is attached to the Sheng verb stem *ngonya*.

(675) *Gai!!! no wonder i hate* ***safiriing*** *usiku*
safiri-ing
My God! no wonder I hate travelling at night. [SMS]

Example (675) presents a case where the Kiswahili verb stem does not end in *-a*. Instead, it ends in *i*. The English progressive suffix attached to it is *-ing*. This leads to a double *i* which means that the final vowel *-i* in the Kiswahili root *safiri* has not been clipped. Both *-i* are maintained.

(676) *im stil* ***malizaring*** *mine*
maliza-ring
I am still finishing mine. [SMS]

(677) *Obash is* ***ingiaring*** *white hse !!!!!*
ingia-ring
Obama is going to the White house. [SNS]

Both messages (676) and (677) are examples whose suffix is the English progressive *-ing* but with *-r* at the beginning which makes up the suffix *-ring*. It is unclear why *-ring* is preferred to *-ing* in such instances.

(678) *they r rungaring now*
runga-ring
They are paying now. [Email]

In example (678), the *-ing* morpheme is attached to the verbal root *runga* which is in a vernacular language (Bukusu-Luhyia).

(679) *woi I was* ***jiburing*** *the fon wen she amkad n anzad* ***liaring***
jibu-ring amka-d anza-d **lia-ring**
Oh dear, I was answering the phone when she woke up and began crying. [SMS]

In (679), the Kiswahili verb stem *jibu* is another example of a verb which does not have the final vowel *–a*.

The findings in this section show that the English progressive suffix attaches itself to verb stems including their final vowel. Such

constructions occur in cases where English is the matrix language. The verb stems are mainly in Kiswahili and Sheng although vernacular languages cannot be ruled out as has been shown in illustration (678).

- **Habitual s / z**

There are also cases of intraword codeswitching where Kiswahili verbs take on the English habitual marker *s*. In some cases, the *-s* is changed to *-z* based on the pronunciation as explained in section 4.1.2.

(680) *Dat kid **chezas** nicely!*
cheza-s
That kid plays nicely. [IM]

(681) *her jamaa **lipias** her kila week!!!*
lipia-s
Her boyfriend pays for her every week. [Email]

(682) *he **malizas** his asgnmnts in rcd tym*
maliza-s
He completes his assignments in record time. [SMS]

Table 45: Engsh Verbs with Engsh Grammar

Verb	-d (Past)	-ing (Progressive)	-z (Habitual)	Root Meaning
683. vibe	*vibed*	*vibing*	*vibe s/ z*	talk
684. heng	*heng(e)d*	*henging*	*hengs*	go to a night-club
685. tweng	*tweng(e)d*	*twenging*	*twengz*	speak with an English accent
686. floss	*flossed*	*flossing*	*flosses*	brag

Interestingly, there are also Engsh verbs that do not occur without the Engsh affixation for tense and aspect. Illustrations of such are in table 45.

5.5.8. Noun plurals as markers of Engsh

As discussed in section 4.1 on phonological spelling, some of the words attached to the /z/ suffix solely use it as an Engsh marker.

In reference to illustration (79) *wagondiz* and (80) *wasupuus*, the second *z/s* plural serves as a marker of Engsh. The following examples illustrate this further.

(687) *Anyway* niliget ***mapplications*** *zako*
ma-application-s
Anyway, I got your **applications**. [Email]

The two interesting things to note about example (687) is that in the noun *ma-application-s*, first of all*, ma* – is the Kiswahili plural marker, but since the noun *applications* begins with the vowel *a*, to avoid the double vowel, the prefix is shortened to *m*. The second thing is that although *ma* is the Kiswahili plural marker, the word *applications* also appears in its plural form- in effect, the word *mapplications* contains double plural marking. Other similar examples with double plurals include

(688) *nimeharvest zile* ***mafruits****!*
I have harvested those fruits. [SMS]

(689) *My* ***vitis*** *r 2 old 2 bebana;-)*
My chairs are too old to carry them along. [IM]

The double plural in illustrations (688) and (689) is convincing evidence that the English suffix –*s* is used as a marker of Engsh.

(*690*) *Hao* ***wamadhes*** *watakam wen??*
Wa-madhe-s
When will those women come? [IM]

Interestingly, the root of the doubly pluralised word *wamadhes* is Sheng. This brings the question as to why the extra plural in English is present since the word is already a Sheng vocabulary in itself. Other illustrations of this kind are

(691) *ma-hater-z (those who disapprove)*
(692) *ma-mbuyu-s (fathers /men)*
(693) *wa-pasu-z (ladies)*
(694) *ma-winche-z (coins)*
(695) *ma-water-s (alcohol)*

(696) miro-z (black people) from the M/MI noun class

Another interesting example is the mass noun *ma-water-s* which also takes on the double plural regardless of the fact that the noun water has limited pluralisation in the language of its origin, English.

5.6. Conclusion

English, Kiswahili, Sheng and vernacular languages are all utilised in CMC although in different proportions. English is the most popular, probably due to the fact that the population group is educated. Vernaculars are the least used but it can be said that they are slowly getting into use unlike before. Some messages in the data make use of all these languages creating an interesting codeswitching scenario. The use of language in CMC as seen in the data is no doubt influenced by the social nature of communication in Kenya. For example the phenomenon of codeswitching in CMC is influenced by the normal use of CS during conversations in informal communication. CS is mainly used orally in informal contexts and CMC users also make use of it in CMC in order to maintain the informality. This is also demonstrated by the fact that unlike previously when Sheng and Engsh were only spoken codes, they are now easily used in CMC texts.

Both Kiswahili and English affixes are utilised with different words thus creating mixed language constructions. Interestingly the communicators are able to decipher the meanings since they are all multilingual and able to understand the involved languages. Engsh makes use of its own marking system to signal its presence in nouns and verbs in form of the tense and aspect markers. The Engsh plural marker *-z* looks like a plural marker but there is more to it because sometimes its presence is not to mark the plural because the word already has a plural or requires no plural or it is not even a noun. These facts suggest that Engsh uses some of the original grammatical morphemes as pure identity markers.

5.7. Motivation of CS in CMC

There are several possible explanations for the occurrence of CS in CMC. These explanations are reminiscent of the general motivations of language use in CMC under discussion in this book; least effort, rapidity, mode limitation and informality including the want to use Sheng and Engsh for various reasons.

(i) Least effort and mode limitation

A major reason for codeswitching in CMC is that users want to spend as little effort and time as possible to compose the message. In some cases, they also need to save space. They therefore use words in different forms so long as the receiver will understand the message.

(697) *hatuioni na 2mefika* ***end***
Sw Eng
We can't see it yet we've reached the end. [SMS]

In illustration (697), which is an SMS message, the lexeme *end* may be more economical in terms of space than its Kiswahili counterpart, *mwisho*.

(698) *I* ***typed*** *haraka kabla stima iende*
Eng Swa
I typed quickly before the electricity went off. [Email]

cf. *(679)* *woi I was* ***jiburing*** *the fon wen she* ***amkad*** *n* ***anzad liaring***
jibu-ring amka-d anza-d **lia-ring**
Oh dear, I was answering the phone when she woke up and began crying. [SMS]

cf. *(633)* ***ashasign mtado*****?????**
asha**sign** mta**do**
He has already signed. What will you do? (sarcastic). [SNS]

When structured fully in Kiswahili, illustration (679) and (633) would have been,

i) Nilipiga chapa haraka kabla nguvu za umeme kutoweka.
ii) Woi, nilikuwa kwenye simu nikijibu ndipo alipoamka na kuanza kulia.
iii) Ameshaweka sahihi. Mtafanyaje?

The Kiswahili translation for *typed* is *kupiga chapa*. It is therefore quicker and it requires less effort to use the English version *typed*. Considering the preceding intraword CS examples, I depict that, for most intraword CS, it is clear that the attached affixes save time and space according to the principles of rapid communication and mode limitation. Nonetheless, there are a few interesting exceptions like in (633). This explanation may work for the first part *asha-sign* 'put a signature' since the equivalent of *sign* in Kiswahili, *weka* or *tia sahihi* does not have a very common occurrence thus may not be immediately accessible. On the other hand, *mta-do* 'what will you do?' whose Kiswahili equivalent is *mtafanyaje* is the complete opposite whereby, the Kiswahili version is common place and the assumption is that it would be the first to come to mind. Yet, the user opts for the English version. Notwithstanding, both of them occur in the same construction. It is important to state that *m-ta-do* is also a rare occurrence that presents a monosyllabic English verb used in Kiswahili grammar. Kiswahili monosyllabic verbs would require the Kiswahili prefix *ku* as in *mtakula?*

(699) a *m - ta - ku - la?*
you.Pl - Fut - Infin - eat
Will you eat. [SMS]

(699) b **m-ta-la*

ii) Rapidity

A major characteristic of codeswitching is that the switched word is the first one that the user thinks of (Myers-Scotton 1997). This is probably true for cases in CMC codeswitching. For example,

(700) *Waaa! I **dozed** sana kwa mat!!!*
Eng Sw Sh
I dozed a lot in the 'matatu' (public transport van). [IM]

Illustration (700) shows a possibility of the sender typing the first form that was thought of. The lexeme *dozed* could have been more quickly recovered than the Kiswahili *sinzia*. The same explanation also applies to *asha-sign* in illustration (633). It is likely that this

form is retrieved quicker than *ameshaweka sahihi*. However, this is only an assumption that would require testing to ascertain the truth.

(iii) Search for accuracy

Some words or phrases are very difficult to translate into other languages. Therefore users find it more accurate to use these words in the original language provided the receiver will understand them. For example the bold word in (701) is difficult to translate with the intended meaning into another language.

(701) *plzz kama kuna mtu anayo plzz* ***upload*** *it.....*
Eng Sw Eng
If anyone has it, please upload it. [SNS]

The term for *upload* in Kiswahili is *pakia* although it is used in the context of loading luggage in to a carriage. In this case, it can be assumed that the user finds it more accurate to use the English version in order to refer to the context of data. This is also an assumption that requires testing in order to get a plausible explanation.

(702) *they called in a* ***mulakusi****!!!!*
Eng Bukusu (vrn)
They called in a psychic/seer. [IM]

mulakusi is a Bukusu word that refers to a person or seer who has some kind of psychic powers to unearth hidden material related to witchcraft. Again, in this case it is assumed that this may be the most accurate form for the intended message.

(iv) Identity, creativity and the use of Sheng and Engsh

Codeswitching can also be used to show some form of identity. For example the use of like *aduaro neni* 'I want to see you' in (704) shows solidarity in the use of the Luo vernacular language. It can also be used to show off one's creativity and prowess for peer identity. For example instead of using a common phrase like 'have *a nice evening' / 'enjoy the evening' / 'good evening,'* the sender in

(703) prefers to use the term *kul (cool)* and *salimia* instead of *greetings to all.*

(703) *kul evening & salimia all.*
Eng /Sh Sw
Have a nice evening and say hello to everyone. [Email]

(704) *Y don't u come ova the weekend aduaro neni.*
Eng vrn (Luo)
Why don't you come over the weekend? I want to see you. [SMS]

Intraword CS is also used because the affixes are used as markers of Sheng. As already discussed in the principle of peer identity, peers tend to use language in CMC in a similar way that they would use in face to face communication. This also justifies the presence of Sheng and Engsh in CMC. For example,

(705) ***nibring'ie*** *imdtly!!!*
ni-bring-ie
Bring it to me immediately. [SMS]

(706) *usimtrust!!!* ***atakuspeedisha***
u-si-m-trust **a-ta-ku-speed-ish-a**
Do not trust him, he will hurry (speed) you up. [SMS]

Similarly, in (705) and (706), the Kiswahili version would have been:

(705) Niletee mara moja.
(706) Usimuamini, atakuharakisha.

In (705), the affixes *ni-* and *-e* are definitely used to make it Sheng. The word seems to consume more space, effort, and time. This leaves the only possible explanation that the message is constructed in this way for different aims. In the same way, in (706), the number of words and characters is the same in the original version and in Sheng translation. Further, both words are relatively common. It shows that there is a different motivation for using the CS version. In this case I also believe that the message has been deliberately 'Shengnised' in order to fit in the Sheng environment of the communicators.

Chapter 6. Codeswitching (CS) and codes

Mi sipickingi phone when tembearing on the street in nai
Sw Sw Eng Sw Eng Eng Sw Eng Eng Eng Eng Eng
I never pick up my phone when walking on the streets in Nairobi
[IM]

This study has a rich accumulation of multilingual data which is reflected through the presence of the use of more than one language or code (e.g. Sheng) within CMC messages. Such multilingual data is not only remarkable but also important for codeswitching studies since some of it provides new dimensions and challenges. This chapter consolidates these new dimensions by highlighting the main challenges that the CMC data poses on code-switching models such as the Matrix Language Framework (MLF) and the Markedness Model. These two codeswitching theories are of particular interest because not only are their features representative of other ML theories, but also that the theories particularly rely on data from Kenyan languages.

It is important to note that some of the information in this chapter appears in other discussions elsewhere in the book but as mentioned, this chapter serves to consolidate it and present it in relation to the two CS theories.

6.1. Properties of Kiswahili – English CS in CMC

This section discusses the general pattern of Kiswahili-English CS based on the CMC data.

Similar to Myers-Scotton (1993) findings, one of the patterns of Kiswahili-English intraword CS is that Kiswahili inflectional prefixes combine with an English root. In other words, the prefix(es) marking the Matrix Language (ML) in an intraword codeswitch is always before the root and the switch occurs at the root and never earlier. Hence if there is a Kiswahili prefix present, then the whole series of all other prefixes until the root would also be in Kiswahili. This is in line with the MLF model as shown in the following illustration.

Eng Root
↑
Sw inflection

The following illustrations demonstrate this.

(707) *tuspeakers*
Small speakers [SMS]

tuspeakers
↑
tu -speakers
Sw INFL Subj Prefix (dimunitive pl) Eng Noun root

It is worth noting that the Kiswahili prefix *tu-* also marks plurality similar to the English suffix *-s*.

(708) *nimewamiss*
I have missed you [SNS]

nimewamiss
↑

nimewa -miss
Sw INFL Prefix Eng verb root
ni me wa miss
Sbj Past perfect Tense Obj (them) Verb

(709) *atakubless*
He will bless you [Email]

atakubless
↑
ataku -bless
Sw INFL Prefix Eng verb root
a ta ku bless
Sbj Future Tense Obj (you) Verb

(710) *ni* ***mkid mbig***
He is a big kid [SMS]

mkid *mbig*
↑ ↑
m -kid m -big
Sw INFL subj Prefix Eng adj root Sw Sg INFL subj Prefix Eng N root

A new pattern that has been discovered in my data is that in a case where the English root 'maintains' its suffix, such a suffix serves to reinforce the Kiswahili prefix. An illustration of this includes *videadly* in (595). Another example is,

(711) *alibrake **vidangerously***
He braked in a dangerous manner [SMS]

vidangerously

vi	-dangerous	-ly
Sw INFL manner Prefix	Eng adj root	Eng manner Suffix

In this case, both the Kiswahili prefix *vi-* and the English suffix *-ly* serve to render the word an adverb of manner. Thus the English suffix *-ly* reinforces the Kiswahili prefix *vi-*. However it is not always that the CS adverbs in the data occur in this manner. In some cases the adverbs occur with only the Kiswahili prefix and not the English suffix e.g. (c.f. 654)

*dey r uzaring **vicheap** jst go now!!;-)*
They are selling **cheaply** just go now. [SMS]

Other illustrations are,

(712) *Hizo fotos zilitokezea **vi-clear***
Those photos came out **clearly**. [SNS]

vi-clear

vi	clear
Sw INFL manner Prefix	Eng adj root

(713) *Ye hubhave **viresponsible***
(S)he behaves **responsibly**. [IM]

vi-responsible

vi	responsible
Sw INFL manner Prefix	Eng adj root

6.2. Properties of English-Kiswahili CS in CMC

The English-Kiswahili intraword CS is the opposite of the CS pattern discussed in the previous section 6.1. This refers to cases where the inflection in form of an English affix(es) is attached to a Kiswahili root. The root can be a noun or verb.

Admittedly, there are not many instances where there is an English prefix preceding a Kiswahili verb. This is related to the fact that English has very few prefixes and they are derivational rather than inflectional. One of them is *rejenga* (cf. 656) and another one is in (714) below.

rejenga
Rebuild [IM]
re-jenga
re- jenga
Eng Prefix (again) Sw Verb root

(714) *antiuchaguzi*
Against elections [SNS]
anti- uchaguzi
anti uchaguzi
Eng Prefix (against) Sw Verb root

A similar interesting occurrence is where the English prefix *anti-* is attached to the Kikuyu noun *mbuku* from English borrowing *(book)* to create;

(715) *antimbuku* (It literally means against books but the interpretation is 'fool' / 'stupid')

[SNS]
anti- mbuku
Eng Prefix (against) vrn (kikuyu)Verb root (books)

A more common structure presented in 5.5.7 is where an English inflectional suffix *-d, -ing,* or *-s* is attached to a Kiswahili verb. The interesting feature of these English inflectional suffixes is that they are added to the complete the Kiswahili stem including derivational suffixes like *-ik, -an, -ish, -iw-* if present but also including the final Kiswahili inflectional *-a*. Thus the English inflectional suffix is added to the following structure Eng/Sw root (+Sw Derivational suffixes) + Sw inflectional final vowel + English inflectional suffix.

Eng/Sw Root - Sw Derivational Stem- Sw Infl final vowel suffix
-ik,-an,-ish,-iw- -a- ↑
Eng Inflectional
suffix -d -ing -s

Most of these English inflections represent tense and number. For example,

Past tense (-d)

(716) *ibiwad*
stolen [SMS]

ibiwa[67] *-d*
ib- -iw- -a -d
Sw Verb root- Sw Derivational Stem- Sw Inflectional suffix - Eng past tense Suffix

(717) *onead*
Biased [Email]

onead
onea -d
Sw Verb root Eng past tense Suffix

Progressive tense (-ing)

(718) *zushianaring*
Quarrelling [IM]

zushianaring
zush(ia)- -an- -a -ring

Sw Verb root- Sw Derivational Stem- Sw Inflectional suffix - Eng progressive tense Suffix

(719) *tembearing*
Walking [IM]

tembearing
tembea ring
Sw Verb root Eng progressive tense Suffix

Habitual tense (-s)

(720) *fanyikas*
Happens [Email]

fanyikas
fany- -ik- -a -s

[67] This verb can also be *ibwa-* in standard Kiswahili.

Sw Verb root- Sw Derivational Stem- Sw Inflectional suffix - Eng habitual tense Suffix

(721) *kunywas*
Drinks [IM]

kunywas
kunywa s
Sw Verb root Eng habitual tense Suffix

One interesting thing to note is that unlike the adverbs of manner in (711) where both affixes can occur simultaneously in the same word, the English tense marker suffixes hardly attach themselves to roots that have Kiswahili prefixes for inflection, otherwise it would lead to an unlikely structure like,

**ali kulad*

*ali kula ~~d~~
Sw infl Sbj Tense prefix Sw Root ~~Eng tense suffix~~

Such a structure does not occur and is impossible in my judgement.

Plural Suffix (-s)

The English plural suffix *-s* is attached to Kiswahili nouns to form the plural. In most cases it is reflected orally and spelt as *–z* in CMC writing.

(722) *simuz*
Telephones [SMS]

simu- *z*
simu z
Sw Noun root Eng Plural Suffix (s)

(723) *nguoz*
Clothes [SMS]

nguo- *z*
nguo z
Sw Noun root Eng Plural Suffix (s)

Apart from the tense and plural suffixes, it is possible for other English suffixes to be attached onto Kiswahili root words to create compound words e.g.,

(724) *afyawise*
Health wise [SNS]

afya	*wise*
afya	wise
Sw Noun root	Eng Infl Suffix

6.3. Challenges to the Matrix Language Framework (MLF) Hypothesis

This section presents some of the complications encountered in the CMC data in a bid to establish the matrix language based on Myers Scotton's Matrix Language Framework. This hypothesis is of particular interest because it was proposed basing on CS data mainly consisting of Kiswahili and English from Nairobi, Kenya.

The MLF hypothesis proposes that in a bilingual CS setting, the languages involved are asymmetrical in such a way that one of them is more dominant and sets the grammar. This dominant language is set as the Matrix Language (ML) while the less dominant language is defined as the Embedded Language (EL) Myers-Scotton (1993:117). The system morpheme is in the ML while the EL has the content morpheme.

Further, Myers-Scotton makes a distinction between content and system morphemes. Content morphemes are presented as carriers of thematic roles. They include nouns, verbs, adjectives and some prepositions which express semantic and pragmatic features. System morphemes on the other hand are function words and inflections whose function is to exhibit a grammatical structure of the construction. This distinction between content and system morphemes is the determining factor in the identification of the ML and EL.

One of the counteracting findings in the CMC data that contradicts the MLF hypothesis is that the ML is not always clearly identifiable and may lead to contradicting conclusions in an attempt to determine it using the MLF hypothesis. An example of this is 614.

they hatad kuhata
They failed completely. [IM]

they hata –d ku- hata
Eng Sh Eng Sw Sh

Similar illustrations could be,

(i) *they chezad kucheza*
They really played.
they cheza –d ku- cheza
Eng Sw Eng Sw Sw

(ii) *they lalad kulala*
They really slept.
they lala –d ku- lala
Eng Sw Eng Sw Sw

The general 'deep structure' of the constructions (i) and (ii) is Kiswahili. Their presentation in standard Kiswahili would be,

(i) *walicheza kucheza*
they really played (lit. they played, to play)

(ii) *walilala kulala*
they really slept (lit. they slept, to sleep)

This reduplication of the verb exemplified above with its first appearance in the stative form and its duplicate affixed to an infinitive is a typical structure in Kiswahili to express emphasis on the verb. It can be illustrated as,

$V_i + [V_i] INFV$

V_i +	$[V_i]$	*INFV*
cheza	[cheza]	ku
play	[play]	to

On the basis of the structure of the above construction, it would be concluded that the ML is Kiswahili. This conclusion is also supported by the presence of the Kiswahili infinitive marker *ku* + verb which in this case is a Kiswahili embedded island and serves as a system morpheme.

However in accordance with the MLF hypothesis, the ML should be based on system morphemes. Therefore on further analysis in regard to the illustrations (i) and (ii) above, going not only by the English pronoun *they* but by mainly the inflection *–d* which is a system morpheme affixed onto the verbs *cheza* and *lala*, the contradictory conclusion should be that the ML is English.

Therefore theoretically, this structure demonstrates the existence of CS constructions where the deep structure is in one language while some of the system morphemes are in another language. This is not accounted for in the MLF theory and presents a challenge in the identification of the ML.

A second challenge that the CMC data poses on the MLF hypothesis is the conflict demonstrated in cases where system morphemes originate from different languages in the CS context. In the MLF hypothesis, it is only content morphemes that can originate from both the ML and EL. System morphemes only originate from the ML.

The abounding cases of this nature are from code switched nouns mainly involving Kiswahili and English. Such a noun in this case simultaneously incorporates plural inflections from both languages thus compounding to a noun root affixed with a double plural form.

For Kiswahili nouns, this CS may involve adding the English plural marker suffix *–s* on the already inflected Kiswahili plural noun e.g. (725)-(727). For their English counterparts, the CS involves adding a Kiswahili plural marker prefix *ma-* on an already plural inflected English noun as illustrated in (728) and (729). Note that

in (728) and (729), the Kiswahili noun class 6 serves as the 'default' plural class so that all nouns that traditionally do not have overt plurals (class 14) and most borrowed nouns especially those in class 9 that should be pluralised in class 10 (in standard Kiswahili), are all pluralised in class 6 (Bosire 2006:190). In principle, nouns of foreign origin in Kiswahili are put in noun class 9/10.

Examples include,

(725)	*wezis (mwizi singular)*			
	Thieves.			[IM]
	wa-	*izi*	*-s*	
	Sw N Pl Infl (class2)	Sw N root	Eng N Pl Infl	
(726)	*mizigos (mzigo Singular)*			
	Luggage.			[Email]
	mi-	*zigo*	*-s*	
	Sw N Pl Infl (class4)	Sw N root	Eng N Pl Infl	
(727)	*vyombos (chombo Singular)*			
	Utensils .			[SMS]
	vy-	*ombo*	*-s*	
	Sw N Pl Infl (class8)	Sw N root	Eng N Pl Infl	
(728)	*mastations (station Singular)*			
	Stations.			[SNS]
	ma-	*station*	*-s*	
	Sw N Pl Infl (class6)	Eng N root	Eng N Pl Infl	
(729)	*mafreshers (fresher Singular)*			
	Freshmen.			[SNS]
	ma-	*fresher*	*-s*	
	Sw N Pl Infl (class6)	Eng N root	Eng N Pl Infl	

The occurrences of these conflicting double plural inflections make it a challenge to identify the matrix language basing on the MLF hypothesis since the system morphemes arise from both languages.

Notably, Myers-Scotton (1993:111) observed a few cases of this nature in her data in words like *ma-watchmen* and *ma-stories*. She concluded that some double forms with class 6 and class 10 prefixes appear in the Nairobi corpus but they are rare with only three occurrences in noun class 6. Notwithstanding, as is clearly seen in the CMC data, double plural occurrences are abundant, common and are spread across different noun classes.

Another challenge to the MLF hypothesis is that it is indeed not unusual to have a 'double switch' of the embedded language (EL) i.e. a Kiswahili inflection can be placed as a prefix, followed by the root in English and then a Kiswahili suffix. In this case, the Kiswahili affixes sandwich the English root,

Eng verb Root(EL)
↑ ↑
Sw prefix *Sw suffix (Der – Infl)*

Illustrations of this are in (730)-(733).

Myers-Scotton (1993:88) claims that "this verb form showing an object prefix, an 'extended suffix', and the final vowel indicating (mainly) indicative mood, is quite unusual among those verbal assemblies including English verb stems in the Nairobi corpus". Some illustrations of this in the CMC data include,

(730) *ametreatiwa*
He has been treated. [SMS]

ametreatiwa

Sw INFL Prefix		Eng verb root	Sw DER Suffix	Sw INFL Suffix
a-	me-	treat-	iw-	a
Sbj,	Pfct tense	Verb	Passive	Indicative

(731) *zishadistributiwa*
The have already been distributed. [Email]

zishadistributiwa

Sw INFL Prefix		Eng verb root	Sw DER Suffix	Sw INFL Suffix
zi-	sha-	distribute-	iw-	a

	Sbj,	completive	Verb	Passive	Indicative

(732) *nitakusendia*
I will send to you. [SMS]

nitakusendia

Sw INFL Prefix			Eng verb root	Sw DER Suffix	Sw INFL Suffix
ni-	ta-	ku-	send	i-	a
Sbj	Fut	Obj	Verb	Applicative	Indicative

(733) *walituspeedishia*
They speeded up things for us [Email]

walituspeedishia

Sw INFL Prefix			Eng verb root	Sw DER Suffixes		Sw INFL Suffix	
wa-	li-	tu-	speed-	ish-	i-	a	
Sbj	Past Tense	obj	Verb	causative	verb derivation	Applicative	Indicative

These 'double switch' illustrations are interesting because they are structurally comparable to the sole illustration *alimbuyia* that came up in Myers-Scotton's (1993:123) corpus but occur frequently in the CMC data.

An additional challenge that arises in an attempt to identify the ML in the CMC data is in CS cases where complete utterances consist of two lexical items of different language origin for example in (625) *come twende*, (626) *ati who* and in (734)-(735). These constructions only consist of content morphemes with a switch between English and Kiswahili.

(734) *twende class*
Sw Eng
Let's go to class [Email]

(735) *jinga type*
Sw Eng
Stupid [SNS]

All these arguments bring out new dimensions in the CS framework by presenting rule governed structures that were excluded but need to be explained by future CS theories.

Code switching is clearly a rule governed demeanour; also in the CMC data. Certain structures like the double past tense markers in **alikulad* do not occur and are felt to be impossible by speakers.

6.4. New Motivations of CS

This section discusses the relationship between CS and socio-psychological factors as its motivations.

Myers-Scotton (1993b) presents various socio-psychological factors as motivations for CS by advancing the markedness model. This markedness model adapts the premise that CS is a reflection of socio-psychological values associated with different linguistic varieties in a specific speech community. CS is used in order to negotiate a change in social distance between the speaker and other participants in the conversation. In this model, the social meanings of code choice, as well as the causes of alternation, are defined entirely in terms of participant rights and obligations.

The markedness model comprises *the negotiation principle*, and three maxims namely; the *unmarked choice maxim, the marked choice maxim* and the *exploratory choice maxim.* The negotiation principle is modelled on Grice's (1975) cooperative principle. According to Myers-Scotton (1993:113), speakers should choose the form of their conversational contribution such that it indexes the set of Rights and Obligations (RO) which the speaker wishes to be in force between him/her and the addressee during the given exchange. Three maxims follow from this principle. The first is the *unmarked choice maxim* which directs the speaker to make his/her code choice the unmarked index of the unmarked RO set in talk exchanges when he/she wishes to establish or affirm that RO set (Myers-Scotton 1993:114). The second is the *marked choice maxim*. This directs the speaker to make a marked code choice which is not the unmarked index of the unmarked RO set in an interaction when you wish to establish a new RO set as unmarked for the current exchange (Myers-Scotton 1993:131). Finally there is the *exploratory choice* maxim which states that when an unmarked choice is not clear, CS should be used to make

alternate exploratory choices as candidates for an unmarked choice and thereby as an index of an RO set in favour (Myers-Scotton 1993:142).

In summary, Myers-Scotton (1993:150) explains that CS occurs due to one of following four motivations. The first motivation is that CS occurs to make a sequence of unmarked choices such that if situational factors change during the conversation, then a new code becomes unmarked. The second motivation is to have CS itself as the unmarked choice with the presumption that a person with the sociolinguistic profile of the speaker will need to index the social identities associated with different codes in the same conversation allowing the speaker to switch between these codes thus realising different identities simultaneously. Further discussion on CS as an unmarked choice is in section 6.5. The third motivation is to have CS as a marked choice in a case where the speaker negotiates a change in the social distance between him/herself and fellow participants with a need to dissociate him/herself. Finally, the fourth motivation is to have CS as an exploratory choice whereby the speaker is uncertain of what language is required either because the interaction is novel or that there is no prior sociolinguistic profile of the fellow participants.

An analysis of the CMC data reveals additional motivations for CS. These motivations have been discussed in detail in section 5.7 and they can be summarised as follows:

- Least effort: CS is used to achieve less effort in the composition and presentation of a message
- Rapidity: CS makes it quicker to use the first word that comes to the mind even if it is in another language.
- Mode limitation: The mode limitation encourages the use brevity in form of shorter words and structural forms using CS.
- Accuracy: A word in one language may appear more accurate than its counterpart thus necessitating the use of CS.

- Identity and informality: CS in CMC presents a form of relaxed communication and identity which enables a user to fit in a given social group.
- Creativity: CS provides a domain for creativity in CMC allowing users to devise new structures.

6.5. CS as an Unmarked Choice

Some new perspectives on motivations of CS in relation to the markedness model are discussed in this section.

One interesting observation relating to the motivation for CS is that it appears that some CS cases in CMC are motivated by the sole fact that one uses two languages.

CS is a manifestation of a code such as Sheng. In this CS manifestation, switching is not generated as a marked choice i.e. not for a particular reason/motivation. Neither is it produced out of habit as an unmarked choice, nor as an exploratory choice. It is produced as a choice to use more than one language. This is to say that it does not matter what language is used in the switch, for example if Lubukusu (see example (736) below) is used as part of the CS, the fact that it is Lubukusu is not significant anymore so long as the end result is the use of more than one language.

In the next illustration, the motivation of the fluid switch from *zimepangwa* to *zikaarangika* does not seem to be anything else but to use English and Kiswahili.

(736) *ngoma zimepangwa mpaka zikaarrangika*
Sw Sw Sw Sw Eng Sw
The music has been arranged until it has become arranged. [IM]

Going by the markedness model, then in fact the use of two languages or the CS itself would be considered as the unmarked code since it is the norm.

A case that cannot be accounted for by the markedness model is in illustration (678) *they are rungaring now*. The verb root *runga* is in Lubukusu (vrn). The motivation of the CS involved from English

to vernacular cannot be explained using the markedness model. There is neither negotiation nor social distance to be created but the use of Lubukusu. Additionally, the involved languages do not have a comparable social status nor do they have associated functions. Another example of this kind is,

(737) *tv yenu ni ya tenee mpaka...*
Eng Sw vrn (kikuyu) Sw
Your television is old till...

6.6. Codes and CS

An important linguistic component that is closely related to CS is the use of codes such as Sheng and Engsh in the Kenyan context

This section presents some ways in which these codes complicate the concept of code-switching.

6.6.1. Sheng

The CMC data demonstrates that Sheng as a code possesses some attributes which challenge CS.

Sheng is a Kiswahili-based code or argot/slang with insertions from other languages. Sheng has spread from being primarily a language of youths, to other social groups and geographically from Nairobi to not only most part of Kenya but also to neighbouring East African countries. It is used for all subjects of conversation and in that sense it is broader than a typical argot or slang. Sheng deviates from Kiswahili in its use of English words, assigning of different meanings to known words, form manipulations and newly coined Sheng words.

Note that Sheng vocabulary keeps changing leading to plenty of synonyms for example, in the exchange below,

In (738a) *vuvuzela*[68] (useless or baseless talk, car horn, loud noise) has synonyms including *pararira, msedes, pang'ang'a* etc. In

[68] This word has its origin from the South African trumpet that gained fame during the 2010 world cup football in South Africa.

b, *ujikate* and *ishia* are also synonyms that mean 'take leave'. The following is an exchange between two participants in SNS.

(738) a. *wacha* ***mavuvuzela*** *zako! all the tym ni* ***pang'ang'a, pang'ang'ana*** *2pu!!!*
tupu
Stop your useless talk. You keep talking all the time but with no substance!

b. *kwani? c* ***ujikate, ishia*** *!!gwedhe wewe*
si
So what? Why don't you leave? You idiot. [SNS]

Note that in spite of the use of synonyms and repetition, the exchange is consistently Sheng.

1. Sheng's unique vocabulary

Sheng is not a mere CS between English and Kiswahili. It also has its unique vocabulary. For example,

(739) *ngonya* (gossip) [SNS]
(740) *ng'am* (see) [SNS]
(741) *noma* (disturbance) [SNS]
(742) *(ma)petho* (boast) [SNS]

The origin of many Sheng words is not traceable to English, Kiswahili or other languages. These words are considered uniquely Sheng.

2. Sheng is codified Codeswitching.

The use of the Sheng code in relation to CS in Kenya has been extensively studied by Mazrui (1995). His study confirms the link between Sheng and CS. Mazrui (1995:171) claims that Sheng is a slang primarily based on Kiswahili-English CS. He describes it as a code characterised by CS. His view is that Sheng is a codified CS. If Sheng is only codified CS, then it means that there is no choice anymore and hence it is actually not CS but borrowing.

However, Sheng goes a step further and is much more than codified CS. Structurally English or non Kiswahili intrusions in Sheng

are part of multilingual or multilectal speech production rather than classical CS.

Bosire (2009:78) makes an interesting observation that CS is not even a necessary condition for a Sheng construction. He uses a Sheng idiomatic expression, *amepigwa pasi* which literally translates to *s/he has been ironed / hit by an iron box* in Kiswahili but the potent meaning in Sheng is *that s/he has been picked*. In my view this illustrates an intricate switch between Sw-Sh. The message is presented in Kiswahili but needs to be interpreted in Sheng.

3. Sheng as a choice for CS

Sheng is a choice for CS and not for the Embedded Language (EL). In terms of motivation to switch, Sheng is different because the choice is for Sheng. The choice is to use Sheng vocabulary to designate Sheng. The marked choice is to speak Sheng and not to use English. For that reason, the CS is between Eng-Sh or Sw-Sh but not Sw-Eng.

Part of this point has been discussed in section 6.5. Embedded words in CS are not consistent. For example a word can be embedded in a clause in one CS context but switched to a standard ML word the next time in the same conversation. It leads to alternation between English and Kiswahili forms. In Sheng there is consistency whereby an exchange in Sheng uses Sheng throughout. Even if there is a change of word, it would be replaced by its Sheng synonym.

4. Sheng Grammar

Sheng has grammatical markers for example clippings and unique endings. This is a counteraction to the previous assumptions that Sheng is marked by an embedment in a construction whose ML is Kiswahili.

Some Sheng vocabulary that is adapted from Kiswahili or English language undergoes some form manipulation in order to 'shengnise' it.

- Sheng's dummy affixation

For example some words are clipped and attached to the dummy suffix *-o* c.f. 4.1.2. Illustrations include words like *market – mako, darasa – daro (class).* In addition to the dummy suffix some vocabulary is attached to a dummy prefix *o-* as shown in table 41.

Table 46: Dummy Suffix in Sheng

Sheng Word	Dummy Prefix	stem	Dummy suffix	English
(743) oticho	o-	-tich-	-o	teacher (SNS)
(744) oruaro	o-	-ruar-	-o	trouser (Sw suruali) (SNS)
(745) okuro	o-	-kur-	-o	Nakuru[69] (IM)
(746) olalo	o-	-lal-	-o	person of Somali origin (SNS)

Another dummy suffix that is used to 'Shenginise' is *–sh*. For example names like Njuguna or Charles are converted to Njugush or Chash (Bosire 2009:82).

- Sheng's grammatical suffix

Another unique ending is Sheng's habitual marker suffix *-ang-*. It clearly defies the MLF framework because it leads to the following structure.

Sw/Eng Root

Sw prefix *Sh suffix* *Sw INFL suffix*

Examples of this nature are presented in table 47.

Using a similar example of *sishtukangi (I never get ruffled),* Bosire (1995:84) argues that such a structure is contrary to the MLF claim which states that the late outsider system morphemes must be supplied by the ML which is the language that is supplying the morpho-syntactic frame in the CS clause.

[69] Nakuru is a Kenyan city in the Rift valley

Note that Myers-Scotton (1993:102) encounters it in the data and explains that it is borrowed from the speaker's first language from the Bantu group as it is categorised as a Bantu suffix.

Table 47: Sheng grammatical suffix *-ang-*

Sheng Word	Prefix	Root	Sheng suffix	SW Infl Suffix	English
*(747) huend**anga***	hu-	-end-	-ang-	-a	I normally go (IM)
*(748) hashughulik**angi***	ha-	-shughulik-	-ang-	- i	s/he doesn't bother (SNS)
(749) wananiudhingi	wa-na-ni-	-udh-	-ing-	- i	they irritate me (IM)
*(750) hustud**yngi***	hu-	-study-	-(i)ng-	- i	usually study (IM)
*(751) Sibuk**ingi***	si-	-book-	-ing-	- i	I don't book / reserve (IM)
(752) nitahepanga	ni-ta-	-hep-	-ang-	- a	I will be sneaking (SMS)

On the contrary, this suffix is used in Sheng by speakers whose first language is not in the Bantu group and even those who only have Sheng as their first language.

6.6.2. Engsh

Engsh on the other hand is described by Kießling & Mous (2004) as a mirror image of Sheng. It draws on English as its base with insertions from Kiswahili and other languages. My point of view is that Engsh is a code within Sheng c.f. 2.1.4. A detailed analysis and description of its linguistic features is presented in sections 5.5.6-5.5.8. This section sums up the characteristics of Engsh that challenge CS.

Apart from the background that Engsh developed from the richer neighbourhoods of Nairobi and that it involves CS with English as the ML, not much is reported on Engsh.

It is true that Engsh is developed from the richer neighbourhoods of Nairobi but it is worth taking note that currently, Engsh is used

generally among those who feel posh[70]. It is a symbol of being polished and fashionable. Some features of Engsh are presented next.

Engsh incorporates a lot of English slang vocabulary like;

(753)	*booze* (alcohol, drink alcohol)	[IM]
(754)	*floss* (boast, brag)	[SNS]
(755)	*dim /daft* (unintelligent)	[SMS]

Engsh also has some English vocabulary with a unique meaning. Some illustrations include 50-52. It also has some vocabulary that is exclusively Engsh like illustrations 53-55.

Table 48: Engsh Vocabulary

Engsh Word	**Sheng Equivalent**	**English**
(756) paged	Shika/pata bol	be pregnant (SNS)
(757) fix	lima	beat (IM)
(758) joint / heng	haree	disco club(SNS)
(759)becks	chaapa	money (SMS)
(760) mots/reck/hummer	moti/dinga/ndai(e)	car(Email)
(761)tweng	bonga kipetho	speak braggingly with an unnatural English/American accent

Structurally, Engsh has a marking system in form of affixes discussed in 5.5.6 and 6.2. These affixes are based on English. They include tense markers like *–d* for past tense, *-(r)ing* for present tense, -z/s for habitual and –z/s for plurals. But notably in Engsh, these affixes do not only serve as inflections in the CS context but also mark words as Engsh.

In essence, Engsh makes use of English grammatical morphemes as identity markers. These markers which are in form of the tense and aspect serve to signal the presence of Engsh in both nouns and verbs.

[70] The people who feel posh are referred to *as mababi* in Sheng.

6.7. CS, CMC and Orality

Both CS and CMC share a form of orality. CS is clearly oral while CMC is for the most part written in oral style. A major part of CMC language is based on the oral language mainly because of the structure, pronounceable letters, numericals and other symbols cf. 4.2 and 4.3. CMC then is a new avenue which facilitates the transfer of CS and orality into written media.

This section discusses the similarities between CS and CMC language.

As already stated, both CS and CMC share a degree of orality. This is of particular interest because in essence CS is oral. However the introduction of CMC has led to the promotion of CS into written texts.

Both CS and CMC are informal. For CS, the informality is mainly based on the fact that the discussions are informal while in addition to this, CMC employs informality encouraged by the involved communication mode.

6.8. Conclusion

This chapter has presented the challenges posed by the CMC data on Myers-Scotton's CS frameworks; that is both the MLF and the markedness model.

In conclusion, it is perceived that although CMC appears to be closely linked to CS, many forms of CMC language display structures that challenge common CS models such as the Matrix Language Frame model.

Nonetheless, it is a major contribution to conclude that linguistic, and structural challenges of ML hypothesis in CS are resolved if we regard the concerned material as manifestation of codes. All the challenges in 6.3 to the MLF model of CS involved instances of the English tense suffixes. If we exclude these tense suffixes as instances of code-switches, the MLF model still stands. In this

view, English suffixes are no longer considered English inflectional material but as Engsh markers.

It is also important to reiterate that new psycho-social motivations for a specific switch are introduced in CMC. These include least effort, brevity, rapidity, mode limitation, identity etc.

In CMC, CS and codes come together. Sheng and Engsh as appropriate styles for CMC render CS written and give it a different function as a marker of modern style.

Finally, it is acknowledged that CS in itself is a wide branch that cannot be exhausted in this book whose main focus is on CMC language. It is therefore recommended that a more detailed research with a deeper analysis be carried out to establish new patterns and motivations of CS in the CMC context. The approach should be wider and involve previous CS studies carried out not only in Africa but also worldwide.

Chapter 7: Grammar in CMC

> *thnx 4 da cash. plan 2 trvl 2mr very ali. Lft cam wit bro nyc fotoz, bn wit her sinz 3*
> Thanks for the cash (money) I plan to travel tomorrow very early. I left the camera with your brother. It has some nice photos. I have been with him since 3 pm. [SMS]

While going through the CMC data, I came across some interesting practises concerning grammar. These are presented in this section.

Generally, CMC grammar has been judged by many as being inconsistent and ungrammatical. Frehner (2008:63) puts it clearly that syntactic reductions are a very central and characteristic feature of CMC. Encouraged by economy and speed with which CMC messages are composed, sentences tend to lack various syntactic features and come to resemble the so called telegram style. Bodomo (2009:114) refers to this practice as the *Economy of Expressions versus Expressivity*. This is a fact when compared to some standard or expected grammar. Yet despite the *non-standard* grammatical scenario, CMC messages are easily comprehended by the participants (sender and receiver). This comprehension points to two possibilities that need further research in order to be substantiated:

- Grammar is not the main determining factor for comprehension in CMC and it is consciously reduced or changed by the participants.

- The comprehension of the message means that the grammar is actually present but cannot be accessed visually.

The task in this section is to illustrate and describe the treatment of grammar in CMC with the main focus being on the omission and in some cases the reduction of grammatical particles.

7.1. Function Words and Grammatical Parts

Although it is an obvious fact that functional words play a very important role in the grammar of any language, it is typical for CMC messages to omit or present them in a reduced form. This section presents data illustrating this in the following subsections. The examples are presented in three lines. The first line presents the original CMC message, and then the second line presents the message again with the omitted particles marked by the symbol ø. In messages that contain reductions, the second line presents the affected particle in bold form (highlighted). The third line presents an interpretation of the message with a highlight on the affected particle. In some illustrations like in (762)-(764), the interpretations are not immediately clear as to which particle (whether definite or indefinite article) has been omitted.

7.2. Articles

Most CMC messages omit articles. The explanation for this is that the articles do not carry a lot of meaning and therefore the message can be understood without them. Ferrara et al. (1991) argue that articles are absent in many human languages and it is therefore not surprising that they disappear in interactive written discourse since the definiteness or indefiniteness can be inferred from the context. The following examples show the omission of the articles *a, an* and *the*.

(762) *plane has crashd n kild MPs!!!*
ø plane has crashd n kild MPs!!!
A plane has crashed and killed Members of Parliament. [SMS]

The message in illustration (762) has omitted the indefinite article *a* as shown.

(763) *She told me she found college and she'l send me the fee structure*
She told me **ø** she found **ø** college and she'l send me the fee structure
She told me that she found **a** college and she'll send me the fee structure. [Email]

Illustration (763) has also omitted *a*. It is interesting that with the exception of this article, this message is written out in a standard

way in terms of structure and spelling, including the definite article *the*.

(764) *want me 2 pay initial instlmnt of 5K*
ø want me 2 pay **ø** initial instlmnt of 5K
They want me 2 pay **an** initial instalment of five thousand Kenya shillings. ('K' is the Sheng for 'thousand.)' [SMS]

Illustration (764) lacks the indefinite article *an* coupled with the pronoun *they*.

(765) *Maze appl a day kps dr away*
Maze **ø** appl a day kps **ø** dr away
Dear, **an** apple a day keeps **the** doctor away. (*maze* is a common Sheng expression used as a form of address.) [SMS]

Illustration (765) lacks the article *an* and the definite article *the* but not *a*.

(766) *rain here is mingi*
ø rain here is mingi
The rain here is a lot. (*mingi* (nyingi) is the Kiswahili word for *a lot*.) [IM]

(767) *I got it in Daily nation*
I got it in **ø** Daily Nation
I got it in the Daily Nation. (The daily nation is one of Kenya's leading Newspapers.) [IM]

Both illustrations (766) and (767) lack the definite article *the*.

7.3. Pronouns

Pronouns are also commonly omitted especially when they refer back to the sender of the message or when the message is a reply or continuation to a prior one with an already set context. This finding confirms findings by Hård af Segerstad (2002), Schmidt & Androutsopoulos (2004) and Döring (2002) whose studies attest the deletion of the subject pronoun, as do the examples below.

(768) *Wl try n hlp bt **ø** wont promise nythng*
ø Wl try n hlp bt **ø** wont promise nythng.
I will try and help but **I** won't promise anything. [SMS]

(769) *Cant get comp bt stil tryin*

ø Cant get ø comp bt ø stil tryin
I can't get a computer but I'm still trying. [SMS]

Illustrations (768) and (769) lack the first person pronoun *I.* Message (769) further lacks the article *a* and the first person pronoun and auxiliary *I am*.

(770) *M ok. Showd bos leta said I tok 2 HOD 2*
M ok. **ø** showd ø bos ø leta ø said ø I tok 2 ø HOD 2
I'm ok. **I** showed the boss the letter. He said that I should talk to the Head of department too. [SMS]

In the initial part of illustration (770), the contraction between the first person pronoun *I* and the copula *am* has been presented economically as a pronounceable *m*. The message lacks the function words *the I, he* and *that*.

(771) *dint mek it 2 wak were toooo tired*
ø dint mek it 2 wak **ø** were toooo tired
We didn't make it to work, **we** were too tired. [SMS]

(772) *jst spok 2 ha...all of us!!!*
ø just spok to her...all of us!
We just spoke to her...all of us. [SMS]

In both (771) and (772), the first person plural pronoun *we* has been omitted.

(773) *hey in de mess, want us 2 bring u sm food???*
hey **ø ø** in de mess, ø ø want us 2 bring u sm food
Hey **we** are in the cafeteria, do you want us to bring you some food. (*mess* is the Sheng word used by Kenyan University students to refer to the *University's cafeteria.*) [SMS]

Additionally in (773), besides the pronoun *we*, the auxiliary *do* and the second person pronoun *you* have been omitted.

(774) *Was nice getin ua mail*
ø Was nice getin ua mail
It was nice getting your mail. [Email]

(775) 3rd *day bila pawa in awa hostl blok bt* ***ø ø*** *bein repaird*
ø ø ø 3rd day bila pawa in awa hostl blok bt **ø** ø bein repaird
It is the 3rd day without power in our hostel block but **it** is being repa [SMS]

The expletive pronoun *it* has been omitted in both (774) and (775). Additionally the third person singular neuter *it* and the third person singular verb *is* have been omitted in (775).

7.4. The Auxiliary Verb *be*

In many cases, auxiliary verbs that combine with present participles are omitted. The examples here involve the omission of the auxiliary verb *are*.

(776) *Hi. Op u beta nw*
Hi. Op u **ø** beta nw
Hi, I hope you **are** better now. [Email]

(777) *hi gal! How u doin? Just wishing u a great day!*
hi gal! How **ø** u doin? ø Just wishing u a great day!
Hi girl! How **are** you doing? I was just wishing you a great day! [SMS]

(778) *How u? We ok*
How **ø** u? We **ø** ok
How **are** u? We **are** ok. [SMS]

(779) *Hope u r well ☺ we r ok.*
ø Hope u **ø** well☺ we r ok.
I hope you **are** well. We **are** ok.
[Email]

In some instances like in (779), the auxiliary verb form *are* is represented by the pronounceable letter *r*.

In relation to the preceding discussion, in many cases, both the first person pronoun and the auxiliary verb are omitted as exemplified in (780) to (782).

(780) *ow waz ur wkend? Travlng 2mr*
ow waz ur wkend? **ø** Travlng 2mr
How was your weekend? **I'm** travelling tomorrow. [SMS]

(781) *hi. in T 4 mtng. Tx*
hi. **ø** in T 4 ø mtng. Tx
Hi. I'm in Thika for a meeting. Thanks. (Thika is a town in the central province of Kenya.)
[SMS]

(782) *Makin lunch 4 us + sun 2 hot 2 tembea out.*
ø Makin lunch 4 us + **ø** sun **ø** 2 hot 2 tembea out.
I'm (indoors) making lunch for us, besides the sun is too hot to walk outdoors. (tembea is the Sw word for 'walk'.) [SMS]

Message (780) lacks *I am* while (781) lacks both *I am* and the article *a*. In (782), *I am, the* and *is,* are missing.

7.5. Conjunctions

The omission of grammatical elements and especially conjunctions is a motivator for parataxis; the juxtaposition of clauses or phrases without the use of coordinating or subordinating conjunctions. According to Frehner (2008), this results to asyndetic sentences. This is illustrated in examples (783)-(784). However it is worth noting that it is not always clear to ascertain the absence of a conjunction.

(783) *Prices zimetrpl I'l fika thea n get intach*
ø Prices zimetrpl I'l fika thea n get intach.
Prices have tripled **but** I will reach there and then get in touch. (fika' is Kiswahili for 'arrive' or 'reach'.) [SMS]

(784) *wat!! hop i didnt get it im now scared*
wat!! ø hop i didnt get it **ø** im now scared
What? I hope I didn't get it **but** I'm now scared. [SNS]

Illustrations (783)-(784) show the omission of subordinate conjunction *but*. Nonethess, conjunctions are not omitted consistently. In some instances they are reduced as shown in illustrations (785)-(789).

(785) *Nd ur advice coz de envrnmentl course has nw requimnts*
ø Nd ur advice **coz** de envrnmentl course has nw requimnts
I need your advice **because** the environmental course has new requirements. [SMS]

(786) *send sm cash if u can **coz** de price is up*
send sm cash if u can **coz** de price is up
Send some money if you can **because** the price is up. [SMS]

In (785) and (786) the conjunction *because* has been reduced to *coz.* This is not unique to CMC but is a common occurrence in

speech. It can be said to have been ready made for CMC because it blends in smoothly.

(787) *i'l try do wat I can..bt b ready 4 anytng*
i'l try ø do wat I can..**bt** b ready 4 anytng
I'll try to do what I can...**but** be ready for anything. [SMS]

(788) *m rily sori bt I cnt put up wit it:-((*
m rily sori **bt** I cnt put up wit it:-((
I'm really sorry **but** I can't put up with it. (sad Smiley) [SMS]

but in (787) and (788) has been reduced by maintaining the initial and final letter to make *bt.* This is a common practise whereby some words are reduced on the condition that they can be understood from the pronunciation of the remaining letters or characters.

(789) *Can I pls come n use ua comp a bit?*
Can I pls come **n** use ua comp a bit?
Can I pls come **and** use your computer a bit? [SMS]

In (789), *and* has been reduced to the letter *n* which can be derived from its pronunciation. This reduction process is similar to the reduction of *be* to *b* in illustration in (787) and *I'm* to *m* in (788). A detailed discussion of this has been presented in section 4.2.

7.6. Conclusion

CMC messages commonly omit or reduce some function words. This omission or reduction is based on the common shared or assumed knowledge between both the message sender and the receiver. These practices lead to brevity and in many cases least effort which are major factors in CMC language.

According to Hård af Segerstad (2002:224), SMS messages lead to grammatical reductions in order to save time effort and space. In this way, senders rely on the receivers' pragmatic inference in order to interpret the message. This is in line with Döring's (2002:14) claim that syntactic reductions in SMS surpass the rate of word deletions in other text communication genres and also in

verbal communication. Frehner (2008:113) also affirms that syntactic reductions in SMS are 1.6 times higher than in Emails. Bodomo (2009:114) finds syntactic reductions both in SMS and MSN chats. In my data, all the genres display instances of grammatical reduction. However, I only did counts of the individual word present and did not count the missing grammatical words. On an interesting note, Bodomo (2009:114) further makes a general assumption that,

> "[..] in shortening a sentence, functional items such as tense and aspect are dropped, while lexical categories such as nouns and verbs are allowed to stay; in fact, this is more of a tendency than absolute adherence to the principle. So while in the morphology lexical items are more susceptible to reduction, in the syntax functional items are more susceptible to reduction or even dropping." (Bodomo 2009:114).

From my observation, the reduction and dropping of functional items is indeed true for all the four CMC genres. Clearly, this practice does not affect the interpretation and understanding of the message. Nevertheless, research should be carried out to ascertain whether the grammar is present but in the background, or whether it is absent and does not determine the interpretation of the message.

Chapter 8. General Conclusion

In conclusion to this study, I would like to revisit the research objectives and hypotheses as presented in 1.4 and 1.5 and then make a general summary of the research. This research had set two main objectives, to analyse the use of language in CMC texts in Kenya in order to reveal the ways in which the CMC text deviates from the standard language use, motivations of these deviations and the end result, which is characteristic of the CMC language. The second objective was to discuss the similarities and differences among the individual CMC genres in terms of their specific registers. Both these objectives have been realised in chapters 4, 5, 6 and 7 where the linguistic variables of CMC have been presented in their occurrence percentages among the genres and described according to their characteristics. I will now present the main conclusions that I deduced from the data analysis.

To begin with, I discovered in the findings that CMC linguistic features loosely fall into two categories i.e. those motivated by the technology in use and those motivated by the social aspects of the users. The features motivated by the technology include the use of phonological spelling, pronounceable letters and numericals, lexical compression, relaxed use of spelling and punctuation, and the use of graphics. The features motivated by the social practices include the use of salutations and the use of different languages and codeswitching in CMC messages.

The main conclusion from the use of phonological spelling, pronounceable letters and numericals is that CMC maintains a close relationship between written and spoken language. Many words are written in a way that reflects their pronunciation rather than their official spelling. In addition to this, an individual letter or numerical can also be used to represent words with which it has a close phonological value. The most commonly used letters are *u (you), n (and)* and *r (are)* while numbers are *2 (to, too)* and *4 (for)*. This is a general agreement by CMC researchers including Hård af Segerstad (2002), Frehner (2008), Crystal (2008) and Bodomo (2009). The new information added to this from my data is that the use of the phonological spelling is closely related to the

Kiswahili alphabet and that the single letters and numbers used to represent words are used basing on the pronunciation either in English or Kiswahili. Another conclusion in my findings that is agreed upon by Hård af Segerstad (2002), Frehner (2008), Crystal (2008) and Bodomo (2009) is that these features i.e. phonological spelling, pronounceable letters and numericals are used in SMS because they involve less effort, are economical in terms of space and make the whole message input process shorter. The new development from the use of numericals shows the emergence of the leetspeak code in the Kenyan CMC. This is an orthographic code that uses numericals to replace orthographically resembling letters e.g. *6* for *b*. It is apparent from the data that this code is mainly used to camouflage obscenities and vulgar words.

Lexical compression is a feature that is frequent in many CMC studies. It appears in research carried out in Europe by Hård af Segerstad (2002), Frehner (2008) (Europe including the UK), Crystal (2001, 2008) in the US and Bodomo (2009) in Asia. It involves the presentation of words in short forms like exclusive consonant writing, abbreviations and acronyms. SMS makes the most frequent use of most forms of lexical compression features except acronyms. Acronyms are most popular in IM as they are useful in rapid typing. An interesting observation on exclusive consonant spelling is that in Hård af Segerstad's (2002) results, exclusive consonant spelling was only manifest in SMS. In contrast, my results have it in all the genres. This proves that it has spread from SMS and has become a general pointer to CMC language.

The CMC data in this book maintains that CMC makes use of a generally relaxed spelling system which develops into new norms and usages. The sender applies the least possible effort, space and time to compose the message in its barest form as long as the receiver is able to comprehend it. This is complemented by the fact that receivers of CMC messages understand and accept spelling errors and typos. In fact according to Crystal (2001) one's literacy level is not judged from this. IM leads in the misspelling and is explained not only by rapidity to keep the chat flowing but also by

casual attitude from the fact that the communicators know each other and are not on the lookout for misspellings. Other forms of relaxed spelling include the unconventional use of capitalisation with the most common being the lack of the initial capitalisation in sentences, proper nouns, names etc.

Punctuation conventions in CMC are also relaxed. The general conclusion in relation to this is that obvious symbols that do not add any valuable information to the message e.g. apostrophes in contractions or full stops at the end of the sentence have been dropped. On the other hand, it is interesting that original punctuation marks are being used in unconventional ways. For example the successive combination of the use of a semi-colon, hyphen and closing bracket to symbolise a wink or successive exclamation marks to communicate shock. This is necessitated by the technological limitations (cf. hypothesis 3). SNS and IM make the most use of punctuation marks. My explanation is that these genres are discursive and need to communicate non verbal feelings. The users resort to the use of different symbols in unconventional ways as way round this.

The use of graphics, Smileys and Emoticons is an interesting feature that is common in CMC. The graphics are used in very creative ways to communicate feelings, emotions, body movements and other actions. They are prevalent in SNS and IM as a way to keep chats as vivid as possible. Smiles, winks and signs of laughter are the most common forms of Smileys. Their frequent use points to the casual and jocular mood of CMC genres. SMS does not make much use of Smileys because inserting them employs extra effort and space.

The second category of the CMC features shows that some CMC features are clearly influenced by the social cultural practices of the users. In this category, I begin by describing salutation culture in Kenyan CMC. I then discuss the featured languages and the general language choice in Kenyan CMC. Next I discuss the subject of codeswitching both at the interword and intraword level. I finally

give an explanation of several motivations of codeswitching in CMC.

The presentation of salutations in Kenyan CMC is clearly influenced by the social practices of the users. Salutations are very important in the Kenyan culture and this is reflected in the use of CMC. My findings are that Kenyan CMC users make an attempt to include some form of salutation in their messages. In fact some messages only contain salutations and an enquiry of well-being. The same is found by Lamoureaux (2009) in her research on SMS among university students in Khartoum, Sudan. This contrasts with Frehner (2008:46-91) and Schmidt & Androutsopoulos (2004:63), who claim that there are very few salutations in (European and American) CMC.

All genres in this study have counts in salutations. Notably in some cases, the messages have a pre-greeting, a greeting and a post-greeting. IM and Email take the lead. IM is synchronous and closely structured as a real time conversation. It therefore begins with greetings by both parties, and ends with valedictions if it comes to a natural end i.e. without being logged out prematurely. Email also makes use of salutations with the general structure of a greeting at the beginning and a valediction at the end. SMS and SNS have a relatively lower count. In addition to the character limitation, my explanation to this is that most of SMS that did not have greetings were influenced by the close proximity of the users for example they had already met face to face and made salutations. Another possible reason is the nature of the message, in that the message may be a quick request, or a short instant reply of some sort. The lower count of salutations in SNS on the other hand is influenced first of all by the commentary nature of the genre. One makes a comment in reaction to a previous line of thought. Additionally, the audience is faceless in that there is no personal acquaintance with each other. This does not give much opportunity for salutations. In fact the data recorded few valedictions in SNS. This is because of its commentary nature where valediction means closing off the thread. For example after getting the needed information, the person who started the thread may

come in and close it by thanking the participants and giving a valediction.

The multilingual nature of Kenya gives CMC users a wide variety of language choice. This is demonstrated in the data which has all the language categories represented. Vernacular languages are the least used. This is related to the users' inexperience in the orthography of vernaculars. Yet, although their count is low, it is impressive that they have been introduced into CMC use. It is my prediction that they will be used more in future and this will both increase their presence in texts and also increase their users' literacy in them. Codeswitching (CS) is another interesting occurrence in the CMC data. It involves the use of different languages in one message (interword CS), or the use of different languages in one word (intraword CS). The data had a few messages that used all the languages in focus. Other studies like Bodomo (2009:203) and Hård af Segerstad (2002:151) also report forms of codeswitching in their CMC data. This use of CS in CMC is encouraged by its use in natural conversations which reflects the closeness of CMC to oral communication. Additionally it is motivated by other factors like least effort, mode limitation, rapidity, search for accuracy and expression of peer identity. More on this is discussed in section 5.7.

Another conclusion from the study concerns grammar. CMC messages commonly omit or reduce function words. This omission or reduction is based on the common shared or assumed knowledge between both the message sender and the receiver. The omitted words are easily predictable. These practices lead to brevity and least effort which are factors in CMC language. This is averred by Hård af Segerstad (2002), Frehner (2008) and Bodomo (2009). In addition to the relaxing spelling standards in CMC, this practice of the omission of some function words is part of what leads to the general agitation that CMC is deteriorating the standard of conventional language.

8.1. Conclusions on CMC Genres

The results of this study lead to various conclusions regarding the genres. First of all, it is notable that the distinctions between these genres have become more fluid. For example, it is currently possible to send Emails and even IM using a mobile phone. Similarly, it is possible to send SMS through some websites. These developments have consequences as they lead to a general form of CMC language without genre specific features. This study was able to collect data before this development from each of the specific genres as discussed in chapter 3 in the methodology. This data was then analysed for features and the occurrences of these features were compared per genre through counts and percentage. This helped reveal some genre specific features that are discussed hereafter.

8.1.1. SMS

One of the features that distinguish SMS data is the general brevity of the messages. They make use of compression methods in message formulation. These compression methods include both the use of pronounceable letters and numbers to replace words, the use of exclusive consonants to form words and avoidance of punctuation marks. In addition to this, SMS is also characterised by other features like the informal writing using a transcription style of spelling and absent capitalisation.

There is a generally lazy and relaxed attitude when dealing with SMS with a focus on brevity. The main effort is applied only to parts that add information or save space in the message. It can be said that these practices were initially triggered by the mode limitation of least space and effort but have gained popularity and become the norm. The general guideline in SMS composition appears to be that 'all that can be left out should be left out and all that can be compressed should be compressed'.

8.1.2. Email

In concurrence with Frehner (2008), the Email genre appears to be a bridge between the traditional letter communication and the CMC. Among the discussed genres, Email is characterised by its

more conventional nature. Its structure in most cases resembles the traditional letter. It begins with a greeting, followed by the message and finally a valediction. The message presentation uses less formal language than the standard. It incorporates some CMC features as discussed in chapters 4, 5, and 6.

Email can be generally described as a blend of writing conventions without extremes on one hand, and its use of CMC features in moderation on the other.

8.1.3. IM

IM is mainly marked by the use of acronyms and graphics. Socially, it is also marked by the use of salutations. Another distinction from the rest is the generally unedited messages with a high frequency of misspellings. These emanate from IM's rapidity and synchronous nature. The synchronicity maps the messages to real-time conversation. This not only encourages the use of graphics to replace real-time emotions and actions but it also leads to rapidity. Rapidity increases the occurrence of misspellings. It also encourages the use of acronyms and other message shortening methods.

It can be summed that IM maintains a carefree attitude in CMC. The general guideline is to keep it real and in close parallel to verbal conversation.

8.1.4. SNS

The use of language in SNS depends on the nature of the forum. It was discovered that official forums which discuss official matters mostly have bureaucratic sanctions and use language in a more formal or conventional way whereas chat and other discussion forums are more casual and informal. It seems that informal discussion forums are accessed as a form of recreation. Therefore users are in a generally relaxed mood which is reflected in their use of language. They do not make effort to keep it formal. Additionally, most participants of the same chat forums have a virtual friendship (buddies) which makes them casual with each other. In contrast official forums maintain a more formal approach to the

discussions as strangers would do. The participants try to keep the formality values of the forum. In fact is some forums, a member is deregistered if he carries on with practises that do not auger well with the forum. The most distinct finding in the SNS data is the use of leetspeak. Participants in forums feel the need to coat their language in this way in order to disguise the use of indecent language. This conforms to social norms where it is considered immoral to be vulgar.

In general the omission of many conventional writing features and adoption of what is literally described as non-standard language (CMC language) affords writing convenience for CMC users. According to Bodomo (2009:307) the youth generally doesn't like lengthy messages, which seems to be a waste of their time (and effort). So they resort to more innovative ways to express themselves in shorter ways. In my view the so called non-standard CMC language is according to the norm in the CMC world.

8.2. Conclusions from Hypotheses

This section presents conclusions from the hypotheses of the research (cf. 1.5.). These are rapid communication, least effort, mode limitation and informal communication.

8.2.1. Principle of Rapid Communication

Quite a large number of CMC messages are sent under urgent conditions which force the user to be brief and fast in typing and sending the message. Some of the CMC communication especially in synchronous communication e.g. Instant Messaging where the communicators are online at the same time require rapidity in typing, sending messages and making replies. In such a scenario, one types swiftly while composing the message and then presses the send button immediately without proofreading the message. This is likely to lead to short and brisk phonologically spelt messages. According to Rheingold (2008), it is faster to type and generate streams of online thought, without having to worry about the niceties of typography. All these lead to rapid communication resulting in grammatical laxity. It seems a common agreement that such briskness in CMC is an excuse not to be too

keen on the grammar of the text. This is generally true for IM. It also holds to some extent for Email and SNS where the message is typed quickly as the composer formulates it. By contrast, it may not absolutely hold for SMS messages where the users invest more time and effort to thoughtfully create and tailor their messages not to surpass the maximum length limit and yet at the same time to communicate successfully using the content. This is vividly clear from the data on spelling errors in section 4.5.1. SMS registers the least number of spelling errors. This means that SMS messages are edited and end up being brisk but with ultimate input by the composer. The main conclusion therefore is that rapidity is more present in IM compared to the rest of the genres. It leads to phonologically spelt texts and a high rate of spelling errors and typos in the genre.

8.2.2. Principle of Least Effort

The principle of human behaviour and least effort was first discussed by Zipf (1940) in *Cognitive linguistics*. Zipf viewed language as a 'tool' that is shaped by its 'jobs' in human society. He introduced the idea that 'useful' behaviours are performed frequently, and frequent behaviours become quicker and easier to perform. The very existence of these quick, easy behaviour patterns then cause individuals to choose them, even when they are not necessarily the best behaviour from a functional point of view. Among these quick useful behaviours is the use of least effort that is recurrent in all the CMC genres. When confronted with text input, the CMC composers obviously choose the most convenient input which requires the least effort even though the end product is not ideal in comparison to the standard. This has been considerably displayed in chapter 4 and 6 where most of the discussed variables at the lexeme and grammar level are in existence because of the employment of the least effort behaviour. This property is mostly exploited in SMS where there is no effort to insert missing grammatical elements like punctuation marks. It originated from the space limitation of the genre, but has now been ingrained into the least effort practice evidenced by variables like the absence of initial capital letters where there is no automatic setting. IM, Email and SNS also exploit this least effort principle. There is hardly any

effort put in to edit or perform a spell-check on IM messages. Email and SNS users put in very little effort in editing the messages such that the Email appears the most formal but its formality is in relation to the other genres. This principle leads to informal texts in CMC genres with use of pronounceable letters and numericals, phonological spelling, missing and irregular use of punctuation, lexical compression, ungrammatical structures etc.

8.2.3. Principle of Mode Limitation

As already pointed out, different CMC genres have their own technical limitations for example the mobile phone keypad is small and the keys are crowded together and some characters are only accessible after multiple key presses. The computer keyboard is more spacious but not only requires one to be stationary and within the range of a particular network but in Kenya, it may also have an almost eroded keyboard or one that is crowded with Chinese characters. Besides the small keypad and screen, of the mobile phone, SMS (in Kenya) are limited to 160 characters, a fact that gives users no choice but to limit their messages using whichever means they can. This leads to omission, compression and reduction as has been illustrated. SNS comments are also limited to 500 characters on some sites like YouTube. In relation to this limitation, users get ingrained to the tradition of shortening their messages even when they would not surpass the character limit. Another point worth noting is that unlike in verbal communication CMC is generally limited in the production of physical and emotional effects like shock, laughter etc. based on this, users become creative and come up with the use of acronyms, punctuation marks, Smileys and Emoticons as substitutes. This is mainly practised in IM and SNS. In the standard style of writing, this limitation is taken for granted and not overcome using the methods seen in CMC.

A close link to the total number of character limitation in SMS, is Hård af Segerstad's (2002) hypothesis that unlimited buffer size would result in more edited or careful language with features characteristic of traditional written language than in genres with limited buffer size. Counter to Hård af Segerstad's hypothesis, the

current data suggests that it is the opposite. A limited buffer size forces the composer to edit their message in order for it to fit, while unlimited buffer size gives more leeway and freedom to use all sorts of features like in IM. This is clearly seen in some sections like on punctuations in section 4.7 where based on its buffer limitation, SMS registers the least use of excessive punctuation unlike in the other acquiescent genres. This mode limitation principle results to the use of pronounceable letters and numericals, lexical compressions, exclusion of punctuation in necessary cases due to space limitation, and yet excessive use in other cases as a substitute to emotions and the frequent use of Smileys and Emoticons.

8.2.4. Principle of Informal Communication

Participants use language differently depending on the formality or informality of the situation. The language is also different depending on the relationship between participants for example; it is more casual between peers, friends and family than say between colleagues during professional communication. The informality in CMC makes it similar to face to face dialogue. This informality is discussed in form of codeswitching, peer communication and identity influence in the subsequent sections.

(i) Codeswitching

The data supports the idea that informal communication and codeswitching is a prevailing practice in Kenyan CMC. This is on both levels of interword and intraword codeswitching. The switching involves all the languages dealt with i.e. English, Kiswahili, Sheng and vernacular languages. The data regarding this has been discussed in chapter 5. English and Kiswahili are the most commonly used languages in codeswitching. The vernacular language category is the least used. The three major reasons for this are that there was a huge possibility that the communicators do not share a vernacular language. Secondly, the literacy rate in vernacular languages is still low, owing to the lack of standardization and this discourages its usage. Thirdly, most of the CMC technologies are not 'vernacular language friendly' in various aspects of text input. This makes users to associate CMC with English and naturally use it for CMC communication in a similar way to the use

of the word *'hallo'* when receiving a phone call. The word is used even by those who do not speak English. Nevertheless it is apparent that the use of vernacular languages in CMC is slowly inching in.

Kiswahili and English are the most common language combinations in both inter- and intraword CS. As discussed in section 5.5, intraword CS reveals different uses of affixation in CS where affixes marking number, tense, and aspect in one language are attached to a lexeme in a different language to achieve the affix representation. Both Kiswahili and English affixes are used with different words thus creating mixed language constructions. Interestingly the communicators are able to decipher the meanings since they are all multilingual and able to understand the involved languages. Findings from the data establish that English affixes on Kiswahili, Sheng or vernacular language lexemes produce Engsh. This is also discussed in section 5.5. Engsh makes use of its own marking system to signal its presence in nouns and verbs in form of the tense and aspect markers. The Engsh plural marker *–z* looks like a plural marker but there is more to it because sometimes its presence is not to mark the plural because the word already has a plural or requires no plural or it is not even a noun. These facts suggest that Engsh has its identity markers resembling a grammar. CS is a sign of informal and speech-like communication in CMC. Most of the genres utilise it to achieve the face to face communication effect although it is additionally exploited creatively in some aspects to save space and to use least effort in constructing the message.

Additionally, it has been established in chapter 6 that a major part of the CS in the CMC language challenges some of the main claims in the MLF and markedness model in CS as proposed by Myers-Scotton (1993a, 1993b). The conclusion drawn from the CMC data in relation to CS is that the codes involved in CMC are open and free to expropriate grammatical elements from other languages which are not normally seen in cases of classic CS

(ii) Peer Communication and Identity

This study has established that some features of informality of CMC language are triggered by the peer relationship among users. A point of interest is that even in some cases like in SNS where the participants are anonymous, the language is informal. It seems that their being on the site creates the assumption that they have similar interests and can therefore communicate freely regardless of the nature of the message. This explains the informal nature of the CMC genres studied. They are mainly casual, between family, friends or even between those who share similar interests.

Another point worth mentioning is that the use of Sheng or Engsh which implies CS in the CMC data depicts peer identity. The use of these codes leads to a better chance of acceptance as a peer. In some cases especially in IM and SNS, language is presented in a creative manner e.g. the use of new acronyms, Emoticons and different affixation patterns in order to impress the peers. The issue of peer communication is also closely tied to the relaxation away from standard language rules. In fact a participant's use of standard language is a way out of the peer group. It would make the rest of the participants insecure in their communicating with him. This establishes that peer influence promotes informal use of language in CMC.

Creativity in the use of CMC language is a factor that has been recurrent as a motivation of some of the CMC features discussed. Bodomo (2009:307) explains that the youth use this in CMC not only to show their innovation but also for keeping abreast with the ever-changing language in the fast-paced world of technology. They also use it as a means to show conformity to the general trends.

8.3. Research Contribution

CMC has taken Kenya and other African countries by storm. Nonetheless, it is still an expanding phenomenon and moving to take root in the rural areas. Cell phones are moving very fast towards achieving this goal. They are driven by their flexibility based on their affordability for both the phone and the airtime credit, capa-

bility for voice-call use even if one is uneducated, their convenient size, and little electric requirement for battery recharge. Emails, IM and SNS for now are a reserve for the urban educated group and notably most of them now access these using their mobile phones. It is generally expected that once basic services like electricity supply get to the rural areas, CMC will be adapted and be enjoyed by all. This makes this book important because it is the first to capture the use of CMC language among the different genres in Kenya. So far, no other research has been carried out giving a detailed and comparative description of CMC genres and language use in Kenya. Secondly, the findings in this book also offer a broader outlook for other CMC studies in Africa and the world.

In terms of the data and discussion, this book provides novel written data not only for codeswitching studies, but also Sheng and Engsh which are codes that are ever-changing. In addition, a detailed description of the features of Engsh is given in sections 5.5.6-0. The corpus also provides data on the presence of the leetspeak phenomenon in Kiswahili. This is discussed in 4.3.2.

In comparison to other CMC studies, this book offers a new approach in the data collection method. Other CMC studies asked respondents to copy out messages for analysis. In addition to taking up time, some features may not be captured accurately. In my data collection approach, I asked my respondents to forward copies of their messages. This is quicker and receives the message as natural and spontaneous as it is. A more detailed explanation of this methodology process is in chapter 3.

This book makes a contribution to CMC language studies because first of all it captures the use of language at a relatively early CMC phase in Kenya and will be a useful resource reference as a basis for the evolution of CMC language in future researches. Notably in a recent wave of innovation, a combination of both a computer and cell phone in one package is available in the current 'smart phones' or 'palm computers' which are increasingly taking over traditional functions of computers such as web surfing, IM, SNS, and Email. This may make SMS, Email, IM and SNS to take the

same format because the typing will be in the same conditions e.g. small screen and keypad. My prediction therefore is that very soon there will be very little distinction of features between the different genres of CMC since all will be originating from a similar source. This book therefore provides a database to trace the distinction.

8.4. Recommendations for further research

In the course of my data analysis, I discovered that SNS has two kinds of forums: the official forums and the informal forums. The official forum focuses on formal discussions and uses language in a more conventional way while the informal forums are more relaxed and apply most of the features that mark CMC language. It would be useful to carry out further research in order to compare the different kinds of SNS forums.

The data that I collected for my research was general because my main point of interest was to identify and describe features and motivations of the youth CMC language. It would be enriching for more specific studies to be carried out on the use of CMC language by different population groups for example from a gender point of view, among adolescents, high school students or even the older generation, rural vs. urban population etc.

It is my view that more research on CMC language use should be carried out in other African countries. This will not only provide deeper insight into the cultural influence on CMC language use but also make a contribution in the description of CMC language in multilingual societies.

In relation to my findings, it appears that during message construction, some lexical forms are mentally retrieved quicker than other synonyms. I made some assumptions of this kind in my discussion in section 5.7. An example is my suggestion that the codeswitched form *asha-sign* is retrieved faster than the Kiswahili *ameshaweka sahihi.* Further research needs to be done in order to substantiate these assumptions. In chapter 6 during my discussion of grammar in CMC, I realised that it may not be the main deter-

mining factor for comprehension in CMC because the intended message could still be comprehended despite the grammar being omitted, reduced or changed by the message sender. It would be interesting to verify this especially by adopting Frehner's (2008:266) recommendation that CMC research should not only focus on the sender but also on the receiver's reading and comprehension of messages. Another area that merits further research is linked to the innovation and peer identity in CMC. It is my assumption that the better the CMC ingenuity, the higher the rank amongst peers. This is a plausible explanation for the new forms of ingenuity that are always appearing in CMC.

It is now a reality that more complex communication technologies like oral video conferencing, computer mediated face to face communication including visual images in real time already exist. It is envisioned that their usage will soon spread owing to the speed with which CMC advances. New research should be conducted not only to describe their use of language but also the challenges of the users in order to offer solutions in updated versions of the technologies.

References

Abdulaziz, M., & Osinde, K. (1997). Sheng and Engsh: development of mixed codes among urban youth in Kenya. *International Journal of the Sociology of Language.* 125, 43-63.

Adams, A. (1996). Language awareness and information technology. *Curriculum and Teaching.* 11 (2), 69-76.

Androutsopoulos, J. (2006). Introduction: Sociolinguistics and computer-mediated communication. *Journal of Sociolinguistics.* 10(4), 419-438.

Androutsopoulos, J., & Georgakopoulou, A. (Eds.). (2003). *Discourse Constructions of Youth Identities (Pragmatics and Beyond New Series)*. Amsterdam: John Benjamins.

Androutsopoulos, J., & Ziegler, E. (2004). Exploring Language variation on the Internet: Regional speech in a chat community. In B. Gunnarsson, L. Bergström, G. Eklund, S. Fridell, L. Hansen, A. Karstadt, B. Nordberg, E. Sundgren & M.Thelander (Eds.) *Language Variation in Europe.* 99-111. Uppsala, Sweden: Uppsala University Press.

Ashton, E. (1944). *Kiswahili grammar: Including intonation*. New York: Longmans, Green.

Baker, A. (1998). CMC Magazine: Cyberspace Couples Finding Romance Online Then Meeting for the First Time in Real Life. *December Communications, Inc. december.com*. Retrieved on January 31st 2006, from: http://www.december.com/cmc/mag/1998/jul/baker.html

Barasa, S. (2007). A Linguistic Analysis of SMS Language amongst Kenyan Urban Youths. Paper presented at the AEGIS European Conference on African Studies. 11 - 14 July 2007 at the African Studies Centre, Leiden. The Netherlands.

Barasa, S., & Mous, M. (2009). The Oral & Written Interface in SMS: Technologically Mediated Communication in Kenya. In I. Van der Craats & J. Kurvers (Eds.), *Low-Educated Adult Second Language and Literacy Acquisition.* 234-243. Utrecht: Lot Publications.

Baron, N. (1984). Computer Mediated Communication as a Force in Language Change. *Visible Language.* 18, 118-41.

Baron, N. (1998). Letters by Phone or Speech by Other Means: The Linguistics of Email. *Language and Communication.* 18, 133-170.

Baron, N. (2004). See You Online: Gender Issues in College Student Use of Instant Messaging. *Journal of Language and Social Psychology.* 23, 397-423.

Baron, N., Squires, L., Tench, S., & Thompson, M. (2005). Tethered or Mobile? Use of Away Messages in Instant Messaging by American College Students. In R. Ling & P. Pederson (Eds.), *Mobile Communications: Re-Negotiation of the Social Sphere.* 293-311. Düsseldorf: Springer-VDI-Verlag. Also available online at:http://www1.american.edu/tesol/Scholarly%20Documents,%20Articles,%20Research/Naomi%20Baron/Tethered%20or%20Mobile.pdf

Bergs, A., & Kesseler, A. (2003). Literacy and the new media: Vita brevis-lingua brevis. In J. Aitchison & D. Lewis (Eds.), *New Media Language.* 75-84. London: Routledge.

Bieswanger, M. (2007). 2 abbrevi8 or not 2 abbrevi8: A contrastive analysis of different space- and time-saving strategies in English and German text messages. *Texas Linguistics Forum.* 50, 1-12. Also available online at: http://studentorgs.utexas.edu/salsa/proceedings/2006/Bieswanger.pdf

Bodomo, A. (2009). *Computer mediated Communication for Linguistics and Literacy: Technology and Natural Language Education.* Hershey: Information Science Publishing.

Bosire, M. (2009). What Makes a Sheng Word Unique? Lexical Manipulation in Mixed Languages. In O. Akinloye, & M. Lioba (Eds.), *Selected Proceedings of the 39th Annual Conference on African Linguistics.* 77-85. Somerville, MA: Cascadilla Proceedings Project. Also available online at: http://www.lingref.com/cpp/acal/39/paper2188.pdf

Boyd, D., & Ellison, N. (2007). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer Mediated Communication.* 13 (1), 210-230.

Bubas, G. (2002). Computer Mediated Communication, Theories and Phenomena: Factors that influence collaboration over the Internet. Retrieved on May 7th 2007, from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.111.822&rep=rep1&type=pdf

Bradner, E. (2001). Social affordances of Computer mediated communication technology: understanding adoption. Retrieved on November 5th 2008, from: http://delivery.acm.org/10.1145/640000/634111/p67-bradner.pdf?key1=634111&key2=9521491721&coll=GUIDE&dl=GUIDE&CFID=87383251&CFTOKEN=18199318

Cameron, D. (1995). *Verbal Hygiene*. London: Routledge

Cheshire, J. (Ed.), (1991). *English around the world: sociolinguistic perspectives*. Cambridge: Cambridge University Press.

Cloud, J. (2006, December 20th). The YouTube Gurus. *Australasian Business Intelligence 1*. Retrieved January 7, 2007, from http://findarticles.com/p/articles/mi_hb4692/is_200612/ai_n17464201/

Crystal, D. (2008). *Txtng: The Gr8 Db8*. New York: Oxford University Press.

Crystal, D. (2001). *Language and the Internet*. Cambridge: Cambridge University Press.

Danet, B., & Herring, S. (2003). Introduction: The Multilingual Internet. *Journal of Computer Mediated Communication.* 9 (1). Retrieved on January 31st 2006, from: http://jcmc.indiana.edu/vol9/issue1/intro.html

Danet, B., & Herring, S. (Eds.), (2007). *The Multilingual Internet: Language, Culture, and Communication Online*. New York: Oxford University Press.

Debski, R. (2004). The Internet in support of community languages: Web sites created by Poles living abroad. *Information Technology, Education and Society.* 5(1), 5-20.

December J. (1997). Notes on Defining of Computer Mediated Communication. In *CMC Magazine.* Retrieved on April 15th 2007, from: http://www.december.com/cmc/mag/1997/jan/decpro.html.

Deen, K. (2005). The Acquisition of Kiswahili. *Language Acquisition and Language Disorders*. Amsterdam: John Benjamins.

Devic, M. (2007). Code-switching in computer mediated communication - a case study of croatian-english discussion forums. M.A Thesis, University of Freiburg. Also available online at: http://www.grin.com/e-book/115170/code-switching-in-computer-mediated-communication-a-case-study-of-croatian-english

Dibbell, J. (Dec. 21st 1993). A rape in cyberspace, or how an evil clown, a Haitian trickster spirit, two wizards, and a cast of dozens turned a database into a society. *Village Voice.* 36-42.

Diki-Kidiri, M. (2001). Les langues africaines dans la société de l'information, in the proceedings of the *international colloquium*

Multilingualism in the Information Society. 9 and 10 March 2001, UNESCO, Paris (not paginated) an article of 12 pages.

Diki-Kidiri, M. (2002). Towards real Linguistic & Cultural Diversity in Cyberspace. On *UNESCO Website*. Retrieved on December 4th 2006, from: http://portal.unesco.org/ci/en/ev.php-URL_ID=6628&URL_DO=DO_TOPIC&URL_SECTION=201.html

Donovan, M. (2009). Writing Forward, a Place for Creative Writers. Retrieved on December 1st 2009, from: http://www.writingforward.com/about-2

Döring, N. (2002). Kurzm. wird gesendet – Abkürzungen und Akronyme in der SMS-Kommunikation. *Muttersprache - Vierteljahresschrift für deutsche Sprache.* 112 (2), 97-114.

Durham, M. (2003). Language choice on a Swiss mailing list. *Journal of Computer mediated Communication. 9: 1.* Retrieved on March 3rd 2007, from: http://jcmc.indiana.edu/vol9/issue1/durham.html

Eckert, P. (2000). *Language Variation as Social Practice: The Linguistic Construction of Identity in Belten High.* Malden, MA: Wiley-Blackwell.

Eldridge, M., & Grinter, D. (2001). y do tngrs luv 2 txt msg? In: W. Prinz, M, Jarke, Y. Rogers, K. Schmidt & V. Wulf (Eds.), *Proceedings of the seventh European Conference on Computer-Supported Cooperative Work.* 219-238. Dodrecht, The Netherlands: Kluwer Academic Publishers.

Erickson, T., & Herring, S. (2001). Persistent Conversation: A Dialog between Research and Design. *Annual Hawaii International Conference on System Sciences (HICSS-34).* 4, 4031.

Ferrara, K., Brunner, H., & Whittemore, G. (1991). Interactive written discourse as an emergent register. *Written Communication.* 8(1), 8-34.

Ferrari, A. (2005). *Description et analyse sociolinguistique du Sheng, langue mixte parlée à Nairobi.* Thèse de doctorat, Paris INALCO.

Feuba, E. (2009). The Sociolinguistics of Mobile Phone SMS Usage in Cameroon and Nigeria. *The International Journal of Language, Society and Culture.* 28(3). Retrieved on March 10th 2010 from:

http://www.educ.utas.edu.au/users/tle/JOURNAL/issues/2009/28-3.pdf

Freud, S. (1930). *Civilization and its discontents*. Translated from the German by J. Riviere. London: Hogarth Press.

Frehner, C. (2008). *Email - SMS- MMS: The Linguistic Creativity of Asynchronous Discourse in the New Media Age (Linguistic Insights. Studies in Language and Communication)*. New York: Peter Lang Publishing.

Gaines, E. (5th March 2008). People Who Write Words in Shorthand by Removing Vowels or Using Acronyms Suck. Retrieved on 12th June 2008 from:
http://randomstuffihate.wordpress.com/2008/03/06/people-who-write-words-in-shorthand-by-re moving-vowels-or-using-acronyms-suck/

Hanson, J. (2007). *24/7: How Cell Phones and the Internet Change the Way We Live, Work, and Play*. Westport, CT: Praeger Publishers.

Hård af Segerstad, Y. (2002). *Use and adaptation of written language to the conditions of computer mediated communication.* PhD dissertation Goteborg: Department of Linguistics, Goteborg University, Sweden.

Hargittai, E. (2007). Whose space? Differences among users and non-users of social network sites. *Journal of Computer Mediated Communication*.13(1), article 14. Retrieved on March 10th 2008, from: http://jcmc.indiana.edu/vol13/issue1/hargittai.html

Herring, S. (2001). Computer mediated discourse. In D. Schiffrin, D. Tannen, & H. Hamilton (Eds.), *The Handbook of Discourse Analysis.* 612-634. Oxford: Blackwell Publishers. Also available online at: http://ella.slis.indiana.edu/~herring/cmd.pdf

Herring, S. (2003). Media and Language change: Introduction*. Journal of Historical Pragmatics.* 4 (1), 1-17. Retrieved on March 10th 2008, from: http://ella.slis.indiana.edu/~herring/jhp.intro.2003.pdf

Herring, S. (2007). A Faceted Classification Scheme for Computer Mediated Discourse. *Language@Internet*. Retrieved on February 4th 2008, from: http://www.languageatInternet.de

Hinrichs, L. (2006). *Codeswitching on the Web: English and Jamaican Patois in E-Mail Communication*. Amsterdam: John Benjamins.

Hijzelendoorn, M. (2009). *The eDatax.mdb Program*. Leiden University.

Höflich, J., & Gebhardt, J. (2005). Changing Cultures of Written Communication: Letter - E-Mail - SMS. In: R. Harper, L. Palen, & A. Taylor, (Eds.), *The Inside Text: Social, Cultural and Design Perspectives on SMS. (Computer Supported Cooperative Work).* 9-31. New York: Springer.

Horn, S. (1998). *Cyberville: Clicks, Culture and the Creation of an Online Town*. New York: Warner.

Hudson, R. (1996). *Sociolinguistics.* Cambridge: Cambridge University press.

Huffaker, D., & Calvert, S. (2005). Gender, identity, and language use in teenage blogs. *Journal of Computer mediated communication.* 10(2). Retrieved on March 10th 2008, from: http://jcmc.indiana.edu/vol10issue2/huffaker.html

Kembo-Sure, S. (1994). *Language attitudes, use and proficiency: A sociolinguistic study of English in Kenya*. Unpublished PhD thesis, Moi University, Kenya.

Kießling, R., & Mous, M. (2004). Urban Youth Languages in Africa. *Anthropological Linguistics.* 46(3), 303-341.

King'ei, G. (1987). Putting the Sheng 'ghost' to rest. *Kenyatta University Bulletin.* 2:22-23.

Kolko, B. & Reid, E. (1998). Dissolution and Fragmentation: Problems in On-Line Communities. In S. Jones (Ed.), *CyberSociety 2.0: Revisiting computer Mediated Communication and Community*. 212-229. Thousand Oaks, CA: Sage.

Kolko, B. (1995). Building a World With Words: The Narrative Reality of Virtual Communities. *Works and Days.* 13(1-2), 105-126.

Kress, G. (1998). Visual and Verbal Modes of Representation in Electronically Mediated Communication: The Potentials of New Forms of Texts. In I. Snyder (Ed.), *Page to Screen: Taking Literacy into the Electronic Era*. 53-79. London: Routledge.

Labov, W. (1994). *Principles of Linguistic Change Volume I: Internal Factors (Language in Society)*. Malden, MA: Wiley-Blackwell.

Lamoureaux, S. (2009). *Message in a Mobile. Mixed-Messages, tales of missing & mobile Communities at the University of Khartoum*. MPhil. Thesis, Africa Study Centre, University of Leiden.

Lawley, E. (1994).The Sociology of Culture in Computer Mediated Communication: An Initial Exploration. Retrieved on March 10th 2008, from: http://www.itcs.com/elawley/bourdieu.html

Lee, K. (2002). Literacy practices in computer mediated communication in Hong Kong. *Reading Matrix*. *2(2)*. Retrieved on August 6th 2008, from: http://www.readingmatrix.com/articles/lee/article.pdf.

Levinson, J. (2004). Music as narrative and music as drama. *Mind & Language.* 19(4), 428–441.

Lewis, P. (Ed.), 2009. *Ethnologue: Languages of the World, Sixteenth Edition*. Dallas, Tex.: SIL International.

Lewis, R. (1997). An Activity Theory framework to explore distributed communities. *Journal of Computer Assisted Learning.* 13, 210–219.

Limo, A. (2008, May 10th). How the Internet may be used to create wealth in a country. *The Daily Nation* 32.

Ling, R. (2004). *The mobile connection: The cell phone's impact on society*. San Francisco: Morgan Kaufmann.

Li-Wei. (ed.) (2000). *The Bilingualism Reader*. London: Routledge.

Luke, C. (2000). Cyberschooling and technological change: Multiliteracies for new times. In B. Cope & M. Kalantzis (Ed.), *Multiliteracies: Literacy learning and the design of social futures.* 69-91. South Yarra, Victoria: Macmillan.

Mahootian, S. (2006). Codeswitching and Codemixing. In K. Brown (Ed.), *International Encyclopaedia of Language and Linguistics*. 2nd Edition, 511-527. Oxford: Elsevier Ltd.

Makoni, S., & Kamwangamalu, N. (2000). *Language and Institutions in Africa*. Cape Town: Witwatersrand University Press.

Maness, J. (2007). The Power of Dots: Using Nonverbal Compensators in Chat Reference. Retrieved on August 3rd 2007, from: http://ucblibraries.colorado.edu/facultyprofiles/files/publications/libmanessj/Maness--The%20Power%20of%20Dots(personal).pdf

Marhoum, A., & Samper, D. (2003). Kenya-Geography, African Studies, University of Pennsylvania, USA. Retrieved on March 20th 2007, from: http://www.sas.upenn.edu/ African_Studies/NEH/k-geog.html

Mason, R., & Bacsich, P. (1998). Embedding Computer Conferencing into University Teaching. *Computers & Education.* 30, 249–258.

Maurais, J. (2003). Towards a New Global Linguistic Order? In J. Maurais & M. Morris (Eds.) *Languages in a Globalising World*. 13-36. Cambridge: University Press.

Mazrui, A. (1995). Slang and code switching: the case of Sheng in Kenya. *Afrikanistiche Arbeitspapiere.* 42, 168-179.

Mbaabu, I. (1996). *Language policy in East Africa: A dependency theory perspective*. Nairobi, Kenya: Educational Research & Publications.

McArthur, T. (1998). Code Mixing & Code Switching. Concise Oxford Companion to the English Language. Oxford: Oxford University Press. Also available online at: http://www.encyclopedia.com/doc/1O29-CODEMIXINGANDCODESWITCHNG.html

McClure, E. (1981). Formal and functional aspects of the code switching discourse of bilingual children. In R.P. Duran (Ed.), *Latino Language and Communicative behaviour*. 69-94. Norwood, NJ: ABLEX Publishing Coorporation.

McCormick, K. (2001). Code-switching: overview. In R. Mesthrie (Ed.), *Concise Encyclopedia of Sociolinguistics.* 447-454. Oxford: Pergamon.

Mi-Kyung, B. (2005). *Insider at border: interactions of technology, language, culture, and gender in computer mediated communication by Korean female learners of English.* DPhil. Dissertation, Ohio State University.

Milroy, L., & Milroy, J. (1992). Social Network and Social Class: Toward an Integrated Sociolinguistic Model. *Language in Society.* 21(1),1-26. Cambridge: Cambridge University Press.

Mitchell, A. (2006 June 12th). A Leet Primer. *Technology News*. ECT News Network, Inc. Retrieved on June 8th 2008, from: http://www.technewsworld.com/story/47607.html.

Murray, D. (1991). The composing process for computer conversation. *Written Communication.* 8: 35-55.

Muthwii, J. (1986*). Language use in Pluri-lingual Societies : the significance of code-switching*. MA. Thesis, University of East Anglia, Norwich.

Muthwii, M., & Kioko, A. (2004). *New Language Bearings in Africa: A Fresh Quest*. Clevedon-Philadelphia: Multilingual Matters Limited.

Mweu, F. (2000). *Overview of Internet in Kenya*. Paper presented at the African, Internet & telecom Summit, Banjul. The Gambia, 5-9 June, 2000.

Myers-Scotton, C. (1993a). *Duelling Languages, Grammatical Structure in Codeswitching*. Oxford: Oxford University Press.

Myers-Scotton, C. (1993b). *Social Motivations for Codeswitching: Evidence from Africa*. Oxford: Clarendon Press.

Nurse, D., & Spear, T. (1985). *The Swahili: Reconstructing the history and Language of an African Society, 800-1500*. Philadelphia: University of Pennsylvania Press.

Ogechi, N. (2005). On Lexicalization in Sheng. *Nordic Journal of African Studies.* 14(3), 334-355.

Omar, A. (1993). *Linking Openings to Closings in Kiswahili Conversations*. Paper distributed by Indiana University Linguistics Club Publications.

Osinde, K. (1986). *Sheng: An Investigation into the Social and Structural aspects of an Evolving Language*. BA. Dissertation, University of Nairobi.

Ouakrime, M. (2001). Promoting the maintenance of endangered Languages through the Internet: The case of Tamazight. In C. Moseley, N. Ostler & H. Ouzzate (Eds.), *Endangered Languages and the Media: Proceedings of the Fifth FEL Conference (Agadir, Morocco, 20-23 September 2001).* Bath, UK. *Foundation for Endangered Languages*. 61-67.

Palfreyman, D., & Khalid, M. (2003). A Funky Language for Teenz to Use: Representing Gulf Arabic in Instant Messaging. *Journal of Computer mediated Communication. 9: 1.* Retrieved on March 4th 2007, from: http://jcmc.indiana.edu/vol9/issue1/palfreyman.html

Paolillo, J. (2001). Language variation on Internet Relay Chat: A social network approach. *Journal of Sociolinguistic.* 5 (2), 180-213.

Parkin, D. (1977). Emergent and stabilized multilingualism: Polyethnic peer groups in urban Kenya. In H. Giles (Ed.), *Language, Ethnicity and Intergroup Relations*. London: Academic Press Inc.

Pinsonneault, A., & Heppel, N. (1998). Anonymity in group support systems research: A new conceptualization, measure, and contingency framework. *Journal of Management Information Systems*. 14, 89-108.

Polomé, C. (1967). *Kiswahili Language Handbook*. Washington, D.C.: Center for Applied Ling.

Prah, K. (1998). *Between Distinction & Extinction: the harmonization & Standardization of African Languages*. Witwatersrand: Witwatersrand University Press.

Rheingold, H. (2008). The Virtual Community: Living online, being here. Retrieved on October 4th 2009, from: http://www.rheingold.com/electricminds/html/vcc_vc_intro.html

Riva, G. (2001). Communicating in CMC: Making Order out of Miscommunication. In L. Anolli, R. Ciceri & G. Riva (Eds.), *Say not to Say: New perspectives on miscommunication.*197-227. Amsterdam: IOS Press.

Romaine, S. (1984). *The Language of Children and Adolescents: Acquisition of Communicative Competence*. Cambridge, MA: Blackwell Publishers.

Rúa, L. (2005). Shortening devices in text messaging: a multilingual approach. *Neuphilologische Mitteilungen.* 16, 139-155.

Samper, H. (2002). *Talking Sheng: The role of a hybrid language in the construction of identity and youth culture in Nairobi Kenya.* Dissertation, University of Pennsylvania.

Schadeberg, T. (1992). *Sketch of Kiswahili Morphology*. Köln: Köppe.

Scherba de Valenzuela, J. (1992). Definition of Communication. Retrieved on June 6th 2006, from: http://www.unm.edu/~devalenz/handouts/defcomm.html

Schler, J., Koppel, M., Argamon, S., & Pennebaker, J. (2006). Effects of age and gender on blogging. Retrieved on March 6th 2009, from: http://u.cs.biu.ac.il/~koppel/papers/springsymp-blogs-07.10.05-final.pdf

Schlobinski, P. et al. (2001). Simsen. Eine Pilotstudie zu sprachlichen und kommunikativen Aspekten der SMS-Kommunikation. *Networx 22*. Retrieved on July 7th 2006, from: http://www.mediensprache.net/networx/networx-22.pdf

Schmidt, G., & Androutsopoulos, J. (2004). löbbe döch. Beziehungskommunikation mit SMS. Gesprächsforschung – Online-*Zeitschrift zur verbalen Interaktion.* 5, 51-76. Retrieved on October 3rd 2006, from: www.gespraechsforschung-ozs.de/heft2004/heft2004

Schneider-Hufschmidt, M. (2005). Usability Issues of Sending Text Messages. In: R. Harper, L. Palen, & A. Taylor (Eds.), *The Inside Text: Social, Cultural and Design Perspectives on SMS. (Computer Supported Cooperative Work).* 223-236. New York: Springer.

Short, J., Christie, B., & Williams, E. (1976). *The Social Psychology of Telecommunications.* New York: John Wiley & Sons Ltd.

Shortis, T. (2001). *The Language of ICT – Information and Communication Technology*. New York: Routledge.

Siebenhaar, B. (2005). Variet¨atenwahl und Code Switching in Deutschschweizer Chatkan¨alen. *Networx* 43. Retrieved on November 20th 2006, from: www.mediensprache.net/de/websprache/networx/docs/index.asp?id=43

Sperlich, W. (2005). Will Cyberforums save endangered languages? A niuean case study. *International Journal of the Sociology of Language.* 172: 51-77.

Spitzmüller, J. (2005). Spricht da jemand? Repräsentation und Konzeption in virtuellen Räumen. Retrieved on June 7th 2007, from: www.ds.uzh.ch/dsDB/download.php?id=45

Sproull, L., & Kiesler, S. (1986). Reducing social context cues: Electronic mail in organizational communication. *Management Science.* 32, 1492-1512.

Sundén, J. (2003). *Material Virtualities*. New York: Peter Lang.

Tagliamonte, S., & Denis, D. (2008). Linguistic Ruin? LOL! Instant Messaging & Teen Language. *American Speech.* 83,3-34. Also available online at: http://technology.newscientist.com/article/mg19826566.600

Thurlow, C. (2001). Talkin' 'bout my communication: Communication awareness in early adolescence, *Language Awareness.* 10, 1-19.

Thurlow, C. (2003). Generation Txt? The sociolinguistics of young people's textmessaging. *Discourse Analysis Online.* 1 (1). Also available online at: http://www.shu.ac.uk/daol/articles/v1/n1/a3/thurlow2002003-paper.html

Thurlow, C. (2007). Fabricating youth: New-media discourse and the technologization of young people. In S. Johnson & A. Ensslin (Eds.), *Language in the Media: Representations, Identities, Ideologies.* 213-233. London: Continuum.

Warschauer, M. (2000). Language, identity, and the Internet. In B. Kolko, L. Nakamura, & G. Rodman (Eds.), *Race in Cyberspace.* 151-70. New York: Routledge.

Webb, V., & Kembo-Sure, S. (2000). *African Voices: An Introduction to the Languages and Linguistics of Africa*. New York: Oxford University Press.

Weinreich, U. (1953). *Languages in Contact: Findings and Problems*. Berlin: Mouton.

Werry, C. (1996). Linguistic and interactional features of Internet Relay Chat. In S. Herring (Ed.), *Computer Mediated Communication.* 47 ff.

Whittaker, S. (2002). Theories and Methods in Mediated Communication. In A. Graesser, M. Gernsbacher, & S. Goldman (Ed.) *The Handbook of Discourse Processes.* 243-286. New Jersey: Erlbaum. Also available online at: http://dis.shef.ac.uk/stevewhittaker/CMC-review02.pdf

Whiteley, W. (1974). Some patterns of language use in the rural areas of Kenya. In W. H. Whiteley (Ed.), *Language in Kenya*. 319-350. Nairobi: Oxford University Press.

Witmer, D., & Katzman, S. (1997). On-line smiles: Does gender make a difference in the use of graphic accents? *Journal of Computer Mediated Communication.* 2(4).

Wright, S. (2004). Multilingualism on the Internet. Theme Issue, *International Journal of Multicultural Societies*. 6 (1). UNESCO Social and human Sciences. Also available online at: http://dis.shef.ac.uk/stevewhittaker/CMC-review02.pdf

Zipf, G. (1940). *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. New York: Addison-Wesley Publishing.

2002 National Centre for Educational Statistics. Retrieved on February 8th 2010, from: http://nces.ed.gov/

2009 Africa Mobile Factobook. Retrieved on March 10th 2010, from: www.africantelecomsnews.com

CyberAtlas. (2003, June 6th). Population explosion! Retrieved on March 10th 2010, from: http://www.clickz.com/showPage.html?page=3603526

Information Society Statistical profiles 2009-Africa. Retrieved on March 10th 2010, from: http://www.itu.int/dms_pub/itu-d/opb/ind/D-IND-RPM.AF-2009-PDF-E.pdf

Internet World Stats usage and population statistics. Retrieved on March 10th 2010, from: http://www.Internetworldstats.com/stats1.htm

Synovate Company (29th January 2010). National survey shows Kenyan Internet market heading towards "critical mass". Balancing Act News *Issue No. 489*. Also available online at: http://www.balancingact-africa.com/news/back/balancing-act_489.html

UNESCO Institute for Statistics (2005).Language diversity on the Internet: an overview. Retrieved on February 8th 2010, from:

http://portal.unesco.org/ci/en/ev.php-URL_ID=20972&URL_DO=DO_TOPIC&URL_SECTION=201.html

Summary in English

The use of Kenyan languages in Computer Mediated Communication (CMC) is the focus of the investigation presented in this book. SMS, Email, Instant messaging (IM) and Social Network Sites (SNS) are the four CMC genres of focus.

The introductory chapter presents a general overview of CMC, its challenges and a presentation of the genres under study. Subsequently, it also presents an overview of CMC in relation to language and finally, the objectives, rationale and hypotheses for the research are presented. The investigation sets out to achieve two main objectives. The first objective is to analyse the use of language in CMC texts as compared to the everyday formal and informal *standard* language use. The second objective is to discuss the similarities and differences among the individual CMC genres in terms of their unique language use for each genre. This chapter also introduces the main principles to be tested as hypotheses in the investigation. These are the principles of rapid communication, least effort, mode limitation, and informal communication that includes codeswitching and peer communication.

The second chapter presents the necessary background to both the language and CMC situation in Kenya. It initially sheds some light on the multilingual set up in Kenya and concludes that that in this complex multilingual environment, the average Kenyan has at least 3 languages, that is, a vernacular language, Kiswahili and English. Besides these three languages, typical Kenyan youths and young adults additionally have Sheng in their linguistic repertoire. It is also pointed out that owing to various reasons, it is now becoming common for Kenyans to possess a combination of Kiswahili, English and Sheng with no vernacular language. This chapter consequently introduces the CMC service providers in Kenya. An analytical discussion of the internet penetration into the region is also provided.

The methodology of the investigation is covered in chapter three. First the procedure of participant selection for the study is described. The participants were selected from Kenyan Universities, middle level colleges and Young Urban Professionals (yuppies). Next, the scope and limitation of the investigation and ethical considerations are provided. A detailed description of data collection per CMC genre is presented. The chapter also gives an in-depth discussion of the variables involved in the data analysis. Finally, the program used and procedure of data analysis is explained.

The fourth chapter initiates the presentation of findings. This chapter is confined to the technologically motivated features of CMC. These features are a result of users' adaptation of the language of their messages to suit the technology involved leading to what many distinguish and refer to as CMC language. The chapter includes the discussion of features like phonological spelling, pronounceable letters and numericals, and different forms of lexical compression like abbreviations, consonant spelling, contractions and clipping. Additional features discussed include relaxing spelling standards, changing use of punctuation and the increasing and innovative use of graphics in CMC texts, all influenced by the technology.

Similar to the previous section, chapter five also deals with the presentation of findings. However, it confines itself to the socially motivated features of CMC. The chapter begins by describing salutation similarities between spoken language and CMC in terms of structure and the language used. Next is a discussion of the featured languages and the general language choice in Kenyan CMC. Subsequently, a discussion of the subject of codeswitching both at the interword and intraword level is presented. It is notable that codeswitching is a recurrent feature in the data and is closely related to spoken language which is imported into CMC. Finally an explanation of several motivations of codeswitching in CMC is advanced.

A considerable amount of the data collected in this investigation was multilingual in nature. This provided a solid base for a deeper

discussion of codeswitching. It was discovered that codeswitching in CMC presents a supplementary dimension to the fundamental theories on codeswitching. The sixth chapter consolidates these new dimensions by highlighting the main challenges that the CMC data poses on code-switching models such as the Matrix Language Framework (MLF) and the Markedness Model. It is concluded in the chapter that linguistic, and structural challenges of ML hypothesis in CS are resolved if we regard the concerned material as manifestation of codes.

The alleged ungrammatical nature of CMC has been a major point of debate in several studies. Chapter seven illustrates and describes the treatment of grammar in CMC with the main focus being on the omission and in some cases the reduction of grammatical particles. It is concluded in the chapter that this omission or reduction of some grammatical particles is based on the common shared or assumed knowledge between both the message sender and the receiver.

The final chapter presents a conclusion of the study. It summarises the book by first revisiting the objectives and the hypotheses of the investigation that were presented in the first chapter and links them to the findings of the study. This chapter consolidates and presents all the conclusions from each section of the book.

Samenvatting in het Nederlands

Het onderzoek in dit boek richt zich op het gebruik van Keniaanse talen in communicatie via de computer (*computer mediated communication*, cmc). Sms, e-mail, chats (instant messaging, im) en socialenetwerksites (sns) zijn de vier genres van cmc waarop we ons concentreren.

Het inleidende hoofdstuk presenteert een algemeen overzicht van eigenschappen en problemen van cmc en de vier gekozen genres. Vervolgens geeft het een overzicht van de verhouding tussen cmc en taal en, tot slot, de doelstellingen, de achterliggende gedachte en de hypothesen van het onderzoek. Het eerste doel is om het taalgebruik in cmc-teksten te vergelijken met alledaags formeel en informeel *standaard* taalgebruik.. Het tweede doel is om de overeenkomsten en verschillen in taalgebruik tussen de verschillende cmc-genres te bespreken. Dit hoofdstuk geeft ook een inleiding in de belangrijkste principes die in de vorm van hypothesen zullen worden getoetst in dit onderzoek. Dit zijn de principes van snelle communicatie, van de minste moeite, van modaliteitsbeperking, en informele communicatie, hetgeen codeswitching en communicatie in *peer groups* omsluit.

Het tweede hoofdstuk geeft de noodzakelijke achtergrondinformatie over zowel de taalsituatie als het gebruik van cmc in Kenia. We bespreken eerst de meertaligheid in Kenia en concluderen dat de gemiddelde Keniaan in deze complexe meertalige omgeving minstens drie talen tot zijn beschikking heeft, te weten, een minderheidstaal, het Kiswahili en het Engels. Naast deze drie talen hebben doorsneel Keniaanse jongeren en jongvolwassenen ook nog *Sheng* in hun taalrepertoire. We wijzen erop dat het om een aantal redenen gewoon wordt dat Kenianen een combinatie van Kiswahili, Engels en Sheng beheersen, zonder minderheidstaal. Vervolgens introduceren we de cms-dienstverleners in Kenia en geven een analyse van het internetbereik in de regio.

De methode van onderzoek wordt besproken in hoofdstuk drie. Allereerst beschrijven we de procedure voor deelnemerselectie. De deelnemers werden geselecteerd van Keniaanse universiteiten, middelbare scholen en onder yuppies. Vervolgens bespreken we het onderzoeksdomein en zijn beperkingen. We geven een gedetailleerde beschrijving van de dataverwerving per genre en bespreken de variabelen voor de analyse in detail. Tot slot leggen we uit welke software we hebben gebruikt en hoe de data geanalyseerd zijn.

Het vierde hoofdstuk maakt een begin met de presentatie van de gegevens. We beperken ons tot de kenmerken van cmc die op de een of andere manier technologisch bepaald zijn. Deze kenmerken zijn het gevolg van het feit dat gebruikers hun taal aanpassen aan de technologie, wat leidt tot wat we cmc-taal zouden kunnen noemen. We bespreken kenmerken als fonologische spelling, uitgesproken lettervormen en cijfers, en verschillende vormen van lexicale compressie, zoals afkortingen, medeklinkerspelling, contracties en *clippings.* Additionele besproken kenmerken zijn minder strenge spellingnormen, een veranderend gebruik van leestekens en het toenemende en vernieuwende gebruik van grafische middelen in cmc – deze worden allen door de technologie beïnvloed.

Net als het voorafgaande hoofdstuk bespreekt hoofdstuk vijf de empirische bevindingen. Het beperkt zich echter tot de sociaal gemotiveerde kenmerken van cmc. Het hoofdstuk begint met een beschrijving van de overeenkomsten in de structuur en de taalkeuze van begroetingen tussen gesproken taal en cmc. Vervolgens bespreken we de betrokken talen en taalkeuze in het algemeen in Keniaanse cmc. Ook gaan we in op codeswitching, zowel op het niveau binnen het woord als over woordgrenzen heen. Codeswitching komt in de data steeds terug en dit houdt verband met het spreektalige karakter van cmc. Tot slot geven we een verklaring voor de verschillende redenen om te codeswitchen in cmc.

Een aanzienlijke hoeveelheid gegevens in ons onderzoek was veeltalig van aard. Dit geeft een goede basis voor een diepgravender

bespreking van codeswitching. We stellen vast dat codeswitching in cmc een nieuwe dimensie toevoegt aan de fundamentele theorieën over codeswitching. Het zesde hoofdstuk geeft een steviger basis aan deze dimensie door de belangrijkste problemen te bespreken die onze cmc-gegevens opleveren voor modellen van codeswitching zoals het Matrix Language Framework (MLF) en het Markedness Model. We trekken de conclusie dat de taalkundige en structurele uitdagingen van het cmc-materiaal voor de MLF-hypothese kunnen worden opgelost als we het materiaal in kwestie beschouwen als een code-manifestatie.

Het zogenaamde ongrammaticale karakter van cmc is een belangrijk discussiepunt geweest in eerder onderzoek. In hoofdstuk zeven illustreren en beschrijven we de manier waarop de grammatica in cmc behandeld wordt, met als belangrijkste punt van aandacht het weglaten en soms reduceren van partikels. We concluderen dat dit weglaten of verwijderen gebaseerd is op een reële of veronderstelde gedeelde kennis tussen de zender en de ontvanger van een boodschap.

Het laatste hoofdstuk is gewijd aan de conclusie van deze studie. Het vat het boek samen door eerst de doelstellingen en hypothesen die in het eerste hoofdstuk aan de orde kwamen opnieuw te bespreken en deze te koppelen aan de bevindingen van ons onderzoek. We brengen bovendien de conclusies van ieder hoofdstuk nogmaals versterkt naar voren.

Sama ya Sheng

Fokas ya hii buku ni use ya malanguajez za Kenya kwa CMC, kaa kwa SMS, Email, chat na Social Network Sites (SNS).

Chapta ya fao inadil na mambo ya kawa ya CMC. Inasho machallenges na zile vindu ziko anda stadi. Pia inaanlish vindu zenye zinarilet na languej. Maobjektivs, rationale na hypotheses ya resach yenyewe pia imeshoiwa. Objektiv ya fao ni kuanalyz vile lugha inatumika kwa CMC text kulingana na venye lugha inatumiwa daili viformal na informally. Ya seko ni kudiskass mifananos na madifarenzes kwa kila CMC kwa lugha yenye uiunda. Hii chapta pia itapiga intro ya maprincipals zenye zitatestiwa kaa hypotheses. Hizi ni kaa vile, principle za kubonga fasta, kutumia nguvu less, mode limitation na informal communication yenye ininclude codeswitching na communication ya wasee matinee.

Chapta ya seko inapresent background yenye inahitajika na situation ya CMC na lugha in Kenya. Inashed light kwa set up ya malanguajez mob za Kenya halafu inakonklud ati hii environment ya lugha mob, imeduu wasee wakue na lugha kaa 3, yaani, lugha ya masa, swa na kingoso/kilami. Kando na hizi lugha 3, mayouth na matinee wana add sheng kwa hiyo list yao. Coz ya mambo kathaa, imekua common kwa wathii wa Kenya kuwa na kombi ya swa, kingoso na sheng bila lugha ya masa. Hii chapta pia inapiga intro ya CMC na network providers wa Kenya. Pia inaanalyze na kudiskass penetration ya mtandao Kenya.

Methodology ya investigation imekuwa covered na chapta 3. Procedure ya selection ya wathii wa hii stadi imeelazewa fao. Inasho venye wathii walisororwa kutoka campo za Kenya, kole na matinee wenye wanatinga waks. Halafu pia inasho scope na maproboz za hii study na pia venye data ilikolectiwa kwa kila aina ya CMC. Kwa hii chapta pia kuna vile inasho kiundani vile data imechambuliwa, halafu finally ile progi imetumiwa kuchambua hiyo data imeshoiwa.

Chapta ya nne ndio inaanzisha kushoo findings. Lakini kaa bado, kwanza inashoo mafichaz ni venye wathii wamezoea na kuadapt

lugha ya messages isuit hiyo technology na ndio wathii mob wanaiita CMC language. Kuna storo ya discussion kaa vile phonological spelling, mapronouncible letters na namba na maabreviation, consonant spelling, contractions na clipping. Pia kuna vindu kama vile kulenga maspellings, kuchenj punctuation na kuongeza uduu na magraphics kwa CMC texts, zote zikiwa zimebringiwa na mambo majuu.

Kaa hiyo section yenye imethela, chapta kobole pia inadil na presentation ya mafindings. Hata hivyo, injitinga kwa socially motivated features za CMC. Hii chapter inaanza na kudeskraib vile kugoteana ni kaa iko sem kwa lugha ya kubonga na CMC kistrakcha na lugha yenye inatumika. Next kuna diskashon ya malanguajez zenye zimetumika na pia lugha ya kawa kwa Kenyan CMC, halafu, pia kuna subject ya codeswitching kwa both interword na intraword level. Ni poa kucheki venye codeswitching inajirudia rudia kwa data, na inajirilet na lugha ya kubonga halafu inaimpotiwa kwa CMC. mambo zenye umotivate codeswitching kwa CMC .

Data kibao yenye ilikolektiwa ilikuwa ya kumix malanguajez mob. Hii ilibring base poa ya kudiskass codeswitching. Ikakuwadiskavad that codeswitching inapresent supplementary daimeshon kwa theories soo za codeswitching. Chapta ya sita nayo inakonsolidet hizi madimensions za wanga kwa kusho maproboz za CMC models ati data ya CMC inapoz kwa zile zingine kaa vile Matrix Language Framework (MLF), na Markedness Model. So inkuwa konkluded ati linguistic na structural challenges za ML hypothesis kwa CS zitaresolviwa kaa tutajipanga na vindu zenye ni concerned kama codes tu.

Hizo mablanda za lugha kwa CMC, imekuwa stori mferu ya debate na machopi kibao. Chapta saba inasho vile grama inatritiwa kwa CMC ikifokas kwa kuchuja au kulenga maneno mengine. So inakonkliudiwa kwa hii chapta ati kulenga ama kuchuja maneno zingine inakuanga tu assumed kwa knowledge ya sender na receiver wa message.

Chapter ya lanyo inapresent conclusion ya hii study. Inasamaraiz hii buku kwanza kwa kugiv objectives na hypotheses zenye zil-ishoiwa kwa chapta ya fao, halafu inazilink na findings za hii stadi. Hii chapta inabring kila kitu pamoja halafu inazipresent zote kama conklushon kutoka kila sekshon ya hii buku.

Curriculum vitae

Sandra Nekesa Barasa was born on the 7th of December 1975 in Bungoma Kenya. She attended Holy Rosary primary school in Turbo and later Milimani primary school in Naivasha. In 1989 she joined Cardinal Otunga high school in Bungoma. Afterwards in 1993 she joined the University of Eastern Africa Baraton in Kapsabet where she obtained a Bachelor of Education (B.Ed) with a major in English and a minor in Kiswahili. In 1998 she joined Moi University in Eldoret for a Master of Philosophy (MPhil) degree in Linguistics. In 2004 to 2006, she worked in the linguistics department at Masinde Muliro University of Science and Technology in Kakamega. She later joined Leiden University as an external PhD researcher in 2007. She is currently attached to the Centre for Language and Intercultural Communication at the University of Technology, Eindhoven.