

**TRANSLATION ERRORS IN CROWDSOURCED TRANSLATIONS: AN  
ANALYSIS OF KISWAHILI FACEBOOK USER INTERFACE CONTENT**

Matthias Kavuttih Kathuke

October 2019

**A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT FOR THE DEGREE  
OF MA IN TRANSLATION, IN THE CENTRE FOR TRANSLATION AND  
INTERPRETATION, UNIVERSITY OF NAIROBI**

**Declaration**

**This dissertation is my original work and has not been presented for examination in any other University.**

**Matthias Kavuttih Kathuke**

**Y61/12963/2018**

**October 2019**

---

**This Dissertation has been Submitted for Examination with our Approval as University Supervisors:**

**Prof. Jayne Mutiga**

-----

**Date**

-----

**Mr. Rufus Karani**

-----

**Date**

## **Dedication**

To Ruby, my daughter, who at two years old loves the alphabet with great curiosity: let education be the golden key that opens great doors ahead of you and usher you into places beyond your wildest imagination

## **Acknowledgement**

First, I would like to thank the Almighty God who has given me the mental, physical and psycho-spiritual health to be able to commence and complete this project and my entire MA.

I appreciate my wife Lilian and daughter Ruby, for being very supportive and patient with “the bull” when I have had a very tight schedule juggling between work, school, ministry and family. You are the best! I love you. And to Cate Waeni for the coffee and timely lunch, *nashkuru sana*.

To my mother, Rahab, who was also my teacher in primary school, and a great guide at home – thanks a lot mum; my late father, Jeremiah Kathuke, for instilling in me the spirit of industry, professionalism and resilience and my late sister, Carol, who believed in me more than I believed in myself, you pushed me towards becoming the best version of myself – I miss you every day; my sisters “Queen” Dorcas, Julie, Minna and brother Victor, for always believing in your big brother – *asanteni sana!*

Many thanks to the University of Nairobi through the Director, Graduate School for awarding me the scholarship that has enabled me to complete my studies.

Many thanks to Prof. Jayne Mutiga, my supervisor and the Director at CTI for all your support right from the scholarship, class instruction and dissertation supervision. My gratitude also to Mr. Rufus Karani, my supervisor, who has been available for priceless guidance and advice throughout my course. My thanks also to Mr. Paul Warambo, who has been of great help and inspiration, who indeed introduced me to translation both as a practice and a scholarly domain through Translators Without Borders, and has been there as my lecturer at CTI and best friend

outside. My appreciation to my lecturers at CTI – Mr Gitonga, Dr. Mureithi, Dr. Wachira, Dr. Kamau – for your great instruction in this noble profession.

My thanks to Simon Andriesen of The Netherlands, who offered me the first opportunity to train and work as a translator though TWB Kenya. My appreciation also goes to Marek Pawelec of Poland, who introduced and taught me MemoQ, my favorite and most efficient CAT Tool. Many thanks to Gary Smith of England and Spain for encouraging me to ask for better fees – that simple trick has worked wonders. I am thankful to Alfred Mtawali, the chair, EAITA, for being a personal friend and mentor in this great profession.

I appreciate my “crew” in Daniel Mwikya, Mike Kimanthi, Emmanuel Edlet, Ali Suleiman and Faith Mwelu, for bearing a portion of my work so that I may study and research. My thanks to Abby Mike for cheering me on especially when at breaking point. Lots of thanks to Frank Murithi for being my research inspiration. I would have been so lost without your insight. My many thanks to my comrades at CTI, translators Sharon Barang’a, Winrose Barasa, Drusillah Lebanon, Abdihakim Ibrahim and Aweys Hussein; interpreters Agnes Mwangangi, James Mutwiri, Mbesa Kalondu, Lena Kostial and Johnstone Katuta – pleasure to have met you all!

My thanks to Prof, Eric Aseka, who is my Pastor at Divine Grace Ecclesia, Pastor Joel Ambundo, Elder Cole Otage, Elder Derrick Murunga and the entire church for being my family. Lots of thanks to my Bible Study Group in Kikuyu and Nu Heritage for understanding when I missed several sessions to be away for classes. And my appreciation to FC Balokole and Jiamini Futsal Planet, though Dennis Kavutti, for giving me a chance to keep fit and fresh in a busy 2019.

## Contents

Declaration .....	ii
Dedication .....	iii
Acknowledgement .....	iv
List of Abbreviations .....	ix
List of Definitions .....	x
Abstract .....	xi
CHAPTER ONE .....	1
1.1 Introduction .....	1
1.2 Background .....	1
1.3 Statement of the Problem .....	6
1.3.1 Objectives .....	8
1.3.1.1 General Objective .....	8
1.3.1.2 Specific Objectives .....	8
1.4 Research Hypotheses .....	8
1.5 Rationale of the Study .....	9
1.6 Scope and Limitations of the Study .....	10
1.7 Theoretical Framework .....	11
1.8 Literature Review .....	12
1.9 Methodology .....	18
1.7.1 Data Presentation .....	18
1.7.2 Data Analysis .....	18
CHAPTER TWO .....	20

2.0	Quality Issues in Crowdsourced Translations.....	20
2.1	A Comparison of Outsourcing and Crowdsourcing.....	21
2.2	The Facebook Crowdsourcing Model and the Trend towards Crowdsourcing .....	24
2.3	The Rationale behind Crowdsourcing Translations.....	27
2.4	Voices of Antagonism towards Crowdsourcing .....	29
2.5	Pro-crowdsourcing Domains .....	31
2.6	Crowdsourcing for Kiswahili Translations.....	32
2.6	Conclusion .....	34
CHAPTER THREE .....		35
3.0	Data Presentation .....	35
3.1	Stylistic Errors .....	40
3.2	Accuracy Errors .....	42
3.3	Grammatical Errors.....	44
3.4	Conclusion .....	46
CHAPTER FOUR.....		47
4.0	Data Analysis, Discussion and Findings.....	47
4.1	Stylistic Errors .....	48
4.2	Accuracy Errors .....	51
4.3	Grammatical Errors.....	54
4.3.1	Wrong Pronouns.....	56
4.3.2	Wrong Prepositions .....	57
4.3.3	Wrong Tense Marker .....	57
4.3.4	Wrong Noun Class .....	57
Conclusion.....		57

4.4 Apparent Translation Strategies Used by Facebook Translators .....	58
4.4.1 Literal or Word-for-word Translation .....	59
4.4.2 Oblique Translation.....	60
4.5 Communicative Implications of the Observed Translation Errors .....	61
4.5 Conclusion .....	61
CHAPTER FIVE .....	63
5.0 Summary, Conclusion and Recommendations .....	63
5.1 Error Types Identified in this Study.....	64
5.2 Apparent Translation Strategies Used by Crowdsourced Translators on Kiswahili Facebook UI .....	65
5.3 The Implications of the Translation Errors on Communication and Usage of Facebook UI .	66
5.4 Recommendations.....	66
Bibliography .....	68
Appendices.....	71

## **List of Abbreviations**

**CTI:** Centre for Translation and Interpretation

**LQA:** Linguistic Quality Assurance, which is the maintenance of a desired level of quality in a translation, especially by paying attention to every stage of the process of translation, including sourcing of qualified translators.

**PM:** Project Manager

**ST:** Source Text

**TM:** Translation Memory

**TT:** Target Text

**TWB:** Translators Without Borders

## List of Definitions

**Backend:** the part of a computer system or application that is not directly accessed by the user, typically responsible for storing and manipulating data.

**Crowdsourcing:** The practice that involves having a large group of people solving a problem or completing a task such as translation, for an individual or, more commonly, for an organisation

**In-house translation:** the process of engaging employed translators to undertake a translation project on employment terms

**Outsourcing:** the process of engaging freelance, unemployed translator to undertake a translation project on prior agreed terms

**Placeholder:** a character, word, or string of characters that may be used to take up space until such a time that the space is needed

## **Abstract**

This study is set out to investigate translation errors in crowdsourced translations for Facebook's Kiswahili user interface. These errors were identified, categorized and their apparent assessed. In order to achieve the study's purpose, data on translation errors was collected from Facebook user interface by taking screenshots of texts that were selected through purposive sampling. Translation errors were then identified based on the basic concepts of the Dynamic Equivalence Theory proposed by Eugene Nida, and the content analysis method. The errors were grouped into three categories; grammar errors, stylistic errors and accuracy errors. The analysis of these errors revealed that they arose mainly from literal translation, lack of provision of a context on the part of Facebook and lack of translator training. The study also assessed apparent translation strategies that were used by translators which may have led to these errors. These were mainly word-for-word translation and oblique translation. Moreover, the study provided a basis for further studies by recommending the use of a hybrid translation approach between crowdsourcing and outsourcing, arriving at a stronger translation approach which leverages on the strengths of both approaches.

## **CHAPTER ONE**

### **1.1 Introduction**

Outsourcing is a major translation approach, where organizations seek an external workforce for the functions for which they lack internal capacity. Amid increased use of outsourcing as a business strategy, another strategy, crowdsourcing, is being widely adopted in the translation world. Translation crowdsourcing entails casting a translation task to a crowd/community of volunteers who translate the content at their convenience at a small fee, or as is often the case, without any form of remuneration.

The recent internet technology has accelerated the use of crowdsourcing a source of translations for organizations advances, particularly the use of online platforms and the increased use of devices like phones, tablets and computers online. There have been concerns, especially among professional translators, that outsourcing translations results in low-quality translations while at the same time interfering with their profession. This study, therefore, aims at exploring the quality of crowdsourced translations and uses qualitative methodology in its data presentation and analysis. Facebook will be used as the case study where the Kiswahili translated User Interface will be analyzed for grammatical, semantic, and pragmatic errors.

### **1.2 Background**

The term crowdsourcing was a 2006 brainchild of Jeff Howe, which is a blend of the words ‘crowd’ and ‘outsourcing’ (European Commission, 2012). In simple terms, it is the practice of obtaining translations by enlisting (sourcing) the services of a large number of people, (the crowd), either paid or unpaid, typically via the Internet. It is also the act of a company or

institution taking a function once performed by employees and outsourcing to the unidentified (and generally large) network of people in the form of an open call (Yifeng Sun, 2015).

In-house translators in companies and institutions traditionally offered translation services. In the event of work overload or other circumstances, these companies or institutions outsource for translators. Outsourcing is the business practice of hiring a party outside a company to perform services that traditionally were performed in-house by the company's employees and staff. The age of the internet has made crowdsourcing possible, thus moving translation provision from in-house translation to outsourcing, and now crowdsourcing. The emergence of crowdsourcing has affected a range of sectors by closing the gap between amateurs and professionals; organizations, companies, and corporations are now embracing 'crowdsourcing' more than even 'outsourcing' as it is a way of taking advantage of the talent held by the public. Crowdsourcing is the new engine driving collaborative processes by exploiting the capabilities of a large group of people who are willing, able, and available to perform a given task, mainly for an outsourcer, by leveraging on the power of the internet.

Anastasiou and Gupta distinguish between crowdsourcing and community-based platforms, two terms often used interchangeably, which is incorrect. In crowdsourcing translation terms, on the one hand, the community translators are connected and have a vested interest in the content, like on Facebook. On the other hand, generally, a crowd is undefined, uncontrolled, and can lead to crowd-slapping, i.e., a retaliation of the crowd against the crowdsourcing process which can destroy a company's hard-earned credibility and humiliate them on a global scale. However, communities are created out of the crowd; so the importance of the crowd is undeniable, but the

crowd has a 'lightweight' attachment to the translation project, whereas the community a 'heavyweight' attachment, which has a huge bearing on the quality of translation.

Although the term crowdsourcing was first coined in 2006, translation of content by their users and not professional translators is not a new phenomenon. Linux is maybe the oldest example, which was translated by its users and still today Ubuntu, a 'distribution' of Linux, is translated by emotionally attached users who want to use the systems in their mother tongue (Anastasiou, D & Gupta R, 2011).

Crowdsourcing is not just used for translation exercises. Ushahidi, an organization formed by a group of bloggers, uses crowdsourced volunteers to help monitor and curb the spread of post-election violence in Kenya. It involves diverse team members, from programmers to geo-mappers as well as translators. It has been a huge success, having begun in Kenya and spreading worldwide and helping in situations like earthquakes in Haiti to tsunamis in Japan.

Translators Without Borders is one of the major organizations that use crowdsourcing in the translation of its content. This is a non-profit organization that provides translation services to other non-profits. As of April 2019, it had translated over 77.5 million words to various non-profits helping alleviate the impact of disasters from the Ebola outbreak in Congo to the Haiti earthquakes.

Facebook is perhaps the most recent example of a major platform which has been translated using crowdsourcing. In March 2008, the entire Facebook site was translated into French in 24 hours by 4,000 dedicated French speakers (HBS Digital Initiative, 2015). Another successful

project of crowdsourcing is the ‘TED Open Translation Project’ which offers subtitles, interactive transcripts, and the ability for any talk to be translated by volunteers worldwide.

A standard crowdsourcing translation workflow includes a website starting with the main page, followed by a log-in screen and registration of willing volunteer translators. According to their credentials, registered people are categorized as translators, reviewers, clients or administrators, and are registered to the corresponding dashboard. After the translations are submitted, they go for approval or rating, and the best translation is accepted and often allotted stars. This is the model that Facebook uses. Translators are ranked according to the number of words translated and/or approved, without necessarily considering the quality of the translated work.

There are four main advantages of crowdsourcing for translations. Crowdsourced translation yields more multilingual content and in less time. A good example is the translation of the entire Facebook UI into French in under 24 hours, a feat performed by about 4,000 crowdsourced translators (HBS Digital Initiative, 2015). Crowdsourced translation is also relatively cheap as translators are not remunerated at market rates. They are never paid at all, so the crowdsourcing entity only needs to fund for the translation tools and maintenance systems. Fourthly, the crowdsourcing entity hopes that having users contribute translation content will offer an insider perspective that will be more receivable by prospective users, although this is not always the case.

The main challenges of crowdsourcing translation are quality concerns from a user's perspective and the lack of monetary motivation for the volunteer translators. The two concerns are interdependent in that unmotivated translators do not commit meeting deadlines, following guidelines, and above all contributing high-quality translations. It is nearly impossible to

determine the aggregate aptitude and competence of the volunteering crowd or community, something that is prerequisite for quality assurance.

Crowdsourcing translation content receives diverse negative reactions, which include the argument that it is unethical to use amateurs where professionals are needed, that it is a threat to the translation profession, and perhaps most importantly, the assertion that the resulting content is of very substandard quality. For example, the American Translators Association criticized LinkedIn and Facebook for crowdsourcing translation content, branding the efforts “disappointing and highly unprofessional” (De Wille 2017, p 25). Crowdsourcing is changing ordinary readers - and in the case of Facebook, users – to translators; from passive spectators to ‘prosumers,’ who are producers as well as consumers of content.

As we have seen above, organizations that do not deal with social media content also crowdsource translation content. However, our interest in exploring Facebook UI content for Kiswahili is informed by the fact that the content is in public domain, unlike other organizations' content that may be protected by non-disclosure agreements and therefore difficult to access and study. Moreover, the trend and impact of social media, and especially Facebook, makes it an interesting singularity to study. Facebook is used by about 2 billion people in the world, most of whom come from outside Canada and the USA. Facebook supports over 100 languages, all of which are translated by crowdsourced translators. Has this translation model been successful so far?

Therefore, against this background, this study aims at investigating the linguistic quality issues emanating from crowdsourced translations of the Facebook user interface content translation from English to Kiswahili from a perception of professional translators.

### **1.3 Statement of the Problem**

Translation crowdsourcing entails casting a translation task to a crowd/community of volunteers who translate the content at their convenience at a small fee, or as is often the case, without any form of remuneration. Crowdsourcing is gaining prominence as a translation approach, especially among digital companies, whose content's level of formality is low, such as social media content. Crowdsourcing is a popular translation approach because of its three main advantages, namely affordability, the ability to produce a huge number of translated words in a short time, and the ability to produce and test translations at the same time by using crowdsourced translators as "prosumers", that is producers and consumers.

Facebook is a major company employing this approach, and its Kiswahili user interface has been translated to a usable extent within a few months. This quick turnaround time has been made possible thanks to the availability of translation crowdsourcing. However, this quick turnaround time, comes with its own challenges, mainly translation errors caused by the lack of Linguistic Quality Assurance, translator vetting, training and preparation as well as lack of context description to aid the translator.

The most important determining factor of translation quality is the expertise of the translators. Lack of requisite translation competence by translators creates the opportunity for translation errors to slip in. This, coupled with the lack of LQA severely affects the quality of the translation produced. The resulting translation is unable to communicate the intended message, therefore failing in the communicative and interactive function that the App was intended for by its creators. Facebook.inc agrees that most of its translated content does not meet professional

quality and that it is only used to depict a gist of the original meaning (Guzman, Pino, Husa, & Ayan, 2018).

Against this backdrop, this study is set out to address translation errors in crowdsourced Facebook user interface content for the Kiswahili App with the aim of assessing the communicative effectiveness of the App. In the review of previous literature on this subject matter, no other involving crowdsourcing for Kiswahili content was discovered, but there are studies in other languages to be compared as we shall see in our discussion.

### **1.3.1 Objectives**

#### **1.3.1.1 General Objective**

This study will aim at assessing the quality of crowdsourced translations for Facebook UI content for Kiswahili. This general objective will be guided by the following specific objectives:

#### **1.3.1.2 Specific Objectives**

1. To identify and categorize translation errors in crowdsourced Kiswahili translations for Facebook user interface content
2. To discuss the apparent translation strategies used by the crowdsourced translators for Kiswahili Facebook UI
3. To assess the implications of the translation errors on communication and usage of the Facebook platform

### **1.4 Research Hypotheses**

Based on the preliminary investigations in preparation for this study, we have made the following hypotheses:

1. The main translation errors in crowdsourced Kiswahili translations for Facebook user interface content are semantic, grammar, and terminology errors
2. The apparent translation strategies used by crowdsourced translators for Kiswahili Facebook UI are literal translation, word-for-word translation and faithful translation
3. The user experience for Kiswahili Facebook UI is poor due to the translation errors caused by crowdsourced translations

## **1.5 Rationale of the Study**

As already seen in earlier sections of this work, crowdsourcing for translations is rapidly becoming a preferred translation method, especially for Web content. Analysing the translation errors emanating from crowdsourced translations for Kiswahili Facebook UI content will provide objective observations on the feasibility of crowdsourcing translations. These observations can be applied to different types of content besides social media user interface content, and therefore serve a larger scope of translation users and practitioners.

The field of translation involves a range of players such as organizations like Facebook and Google.inc, professional translators, amateur translators, aspiring translators, and researchers, among others. Overall, the study will underscore the merits and demerits of crowdsourcing translations in general, and Kiswahili in particular, and therefor serve as a reference resource for interested parties.

For organizations, the findings of this study will act as a guiding principle in differentiating between the quality of standard and crowdsourced translations. For translators and translation agencies, this study will enable them to have an informed opinion on the quality of crowdsourced translation and, based on the conclusions, whether or not it should be used instead of standard translation. Translation agencies can use the findings in this study to advocate for more stringent quality assessment and control mechanisms in crowdsourcing to avoid the errors that are commonly caused by crowdsourced translations. For researchers, the findings of the study will add to the available literature concerning crowdsourcing translations.

It is important to note that the concept of crowdsourcing for English to Kiswahili translations is an unexplored frontier in the international market, and especially so in the Kiswahili speaking world, and therefore has not been studied academically. So this serves as a pioneer study in this interesting field, breaking ground for more studies in this language pair.

### **1.6 Scope and Limitations of the Study**

The focus of this study is Facebook User Interface (UI), which is the means and platform in which a user interacts with the Facebook website through the computer or mobile device. If the is full of errors, obviously user experience is compromised, such that the user will need more time to perform tasks, feel dissatisfied and may opt out of the service. A bad user experience means that the very purpose of creating Facebook is unmet, which is social interaction with friends. Therefore, a high quality User Interface is paramount in the success of Facebook, and any social media platform, and is therefore the sole focus of this study.

The researcher only samples translations from Facebook UI translated into Kiswahili from English. The errors present in the Facebook UI content in Kiswahili may be too numerous to capture and analyse in the scope of this study, therefore the researcher has sampled various errors in three categories namely accuracy, grammar and style. While there may be other minor error categories, this study focuses on the mentioned categories in order to make its objectives attainable.

The errors under investigation are purely linguistic, excluding graphic and design problems such as text truncation or overflow, which is as a screen layout issue. Translation errors resulting from

content posted by users on Facebook will also not be considered, the focus solely being on the user interface content developed by Facebook's backend.

Facebook is in continuous change through daily translations and edits, meaning that some errors are corrected and others introduced over time. Therefore, this study limits itself to the translation errors observed and sampled within the period of the study, which is July and August 2019.

This study only assesses the translation errors that the researcher observes that they affect the UI to any extent, minor or major. So the main question is not on the severity of the error(s) so long as quality is compromised whether slightly or severely.

### **1.7 Theoretical Framework**

This study is founded on Eugene Nida's Dynamic Equivalence theory, which is also referred to as the Functional Equivalence theory. According to Nida (1995) dismisses traditional phrases like "faithful translation," "free translation," and "literal translation" and instead advocates for the use of dynamic equivalence and formal equivalence as the basic guidelines and directions in translation. Formal equivalence gathers attention around the form and content in a specific message by harmonizing the message in the target audience with the various parts of the original language. In a nutshell, the formal equivalence model attempts to equate the original text to the original text such that the translated text is a reflection of the original language's linguistic features like structure, syntax, grammar, and vocabulary.

Conversely, dynamic equivalence emphasizes the importance of the message received by the audience in translation. Compared to the formal equivalence approach which stresses on what

would be called direct translation in the attempt to maintain similarity between the original and translated text, the dynamic equivalence model holds that messages are meant to be understood and appreciated both in their form and content. The approach calls for translators first to understand the original meaning as opposed to merely concentrating on the grammatical form of the original text. The dynamic equivalence approach is adopted in this study because it puts weight on meaning—intelligibility—such that even if a translated text is understandable and grammatically correct but does not elicit immediate meaning, then such translation is deemed ineffective. Instead of concentrating on the form of the message, Nida (1995) proposes the consideration of how the receptor responds to the message. In the context of the current proposed study, translated text may still contain pragmatic errors regardless of its grammatical correctness and understandability—the meaning that an audience derives from a translated text should, therefore, always take center stage in translation.

## **1.8 Literature Review**

This section will present the existing literature about the attractiveness and perceptions of crowdsourced translation; the drivers of motivation for individuals to engage in crowdsourcing translations for Facebook; the experiences and skills necessary for Facebook users to participate in translation; the effect of personal background on translation; and the obstacles faced during translation. Lastly, the section will review the interaction between crowdsourced translation and professional translation for Facebook content.

It is important to recap the meaning of crowdsourcing and crowdsourcing translations to align this section with the previous sections and to appreciate the literature invested in defining these concepts. As aforementioned herein, the term crowdsourcing was conceptualized in 2006 by Jeff

Howe as ‘the practice of obtaining translations by enlisting (sourcing) the services of a large number of people, (the crowd), either paid or unpaid, typically via the Internet.’ Estellés-Arolas & González-Ladrón-De-Guevara (2012) studied a wide range of definitions of crowdsourcing with the aim of harmonizing these definitions to derive a universal definition. They deduced the following definition of crowdsourcing:

*A type of participative online activity in which an individual, an institution, a non-profit organization or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. The undertaking of the task, variable complexity and modularity, and in which the crowd should participate in bringing their work, money, knowledge, and/or experience, always entails mutual benefit. The user will receive the satisfaction of a given type of need, be it economic, social recognition, self-esteem, or the development of individual skills, while the crowdsourcer will obtain and utilize to their advantage what the user has brought to the venture, whose form will depend on the type of activity undertaken. (p.*

*197)*

This definition fronts several attributes of crowdsourcing: (a) an online endeavor pitting a range of tasks provided by a group, (b) voluntary participation, (c) mutual benefits between the outsourcer and the participants. These attributes are directly applicable to crowdsourcing

translations since it brings together a crowdsourcer such as Facebook or Google (online platforms) and volunteer groups (Facebook or Internet users) to tackle a translation problem voluntarily. While crowdsourcing translations is mutually benefiting to the crowdsourcing organization and the platform users, there are usually no monetary remunerations.

O'Hagan (2016) appreciates the impact of globalization and the current technological milieu on the translation frontier, especially crowdsourced translations. There is an increased need for translation given the surge in daily internet users and organizations as they attempt to communicate with their international audience and potential customers. O'Hagan (2016) notes that this need for translation has led to an increase in free online translation solutions that are mostly grounded on automatic translation (machine translation) on platforms like Facebook Translate, Google Translate, and Microsoft Bing Translator. For instance, Google Translate handles more than 143 billion words daily, a reflection of the extensive and pervasive adoption of crowdsourced translations among internet and social media users (Wolverton, 2018). However, crowdsourced translations are widely criticized, particularly by professional translators, for being 'quick-fixes' that prioritize immediacy, convenience, and cost over quality and professional assistance.

Pascoal, Furtado, and Chorão (2017) reiterate that the onset of localized crowd translations particularly on Facebook in 2008 resulted to an acceleration in crowdsourced translations such that the practice is now becoming the next big thing in translation. However, Pascoal, Furtado, and Chorão (2017) do not shy away from highlighting the controversies among professional as a result of using volunteer work and fans to translate content. Professional translators view crowdsourcing as an impediment to their profession and perceive it to violate the integrity of

translated content in terms of quality. In their study, Pascoal, Furtado, and Chorão (2017) found that a majority (60%) of the professional translators who featured in the study viewed crowdsourcing translation as a threat to their profession. They also believed that crowdsourcing hurts the quality of translation. According to professional translators, translation is more than just a combination of language knowledge and the Erasmus of becoming a translator. Professional translators also hold the view that crowdsourced translations should not be taken seriously since volunteer translators are mostly amateurs in the field. These findings underline the dilemma of settling for the immediacy, convenience, and low-cost of crowdsourced translations at the expense of the quality, dependability, and integrity offered by professional translators at a comparably slower rate and higher cost.

If monetary gains are not integral aspects of crowdsourcing translations, then what motivates participants to contribute to the initiator's call for translations? According to Pinto and dos Santos (2018), where monetary benefits are not involved, contributors in crowdsourcing activities derive most of their motivation intrinsically through the fun, learning, and satisfaction as well as extrinsically through acknowledgment by the organization that initiates the crowdsourcing. The findings of this study mean that so long as volunteer translators remain intrinsically motivated as they participate in crowdsourcing translations through various online platforms like Facebook and Google, such crowdsourcing will remain resilient regardless of the perceived quality issues by professional translators.

Kaminskienė and Kavaliauskienė (2012) reviewed the skills and competencies necessary for translation and interpreting. The survey featured employers of the European Union institution, lecturers, alumni, and Vilnius university students. The study unearthed seven core competencies

necessary for one to become a reliable translator. These are technological competencies, language competence, cultural competence, production dimension, interpersonal dimension, thematic competence, and information mining competence. In a similar study, Calvo (2011) attempts to distinguish between translator competence and translation competence or whether the two are synonymous concepts. The study concluded that translation competence is a product of professional translator qualifications.

Against this backdrop, it could be valid to argue that most of the people who volunteer for crowdsourcing have neither translation competence nor translator competence as the two concepts are intertwined.

Apart from the competence needed for one to translate content reliably, Dolmaya (2011) raises a question of ethics, especially because crowdsourcing translations rely on volunteer labor from people that do not necessarily possess the training nor competencies needed in translation. Moreover, Dolmaya (2011) criticizes the use of free labor by giant organizations that are perceptibly capable of remunerating the crowd involved in translation.

While there are several benefits attached to crowdsourcing translations such as enhancing the visibility of translation, helping smaller languages to gain more visibility online, and showcasing the value of crowdsourcing to the society, there is an underlying concern of devaluing the role of professional translation as well as lowering professional translator's occupational status. Dolmaya (2011) adds that before judging crowdsourcing translations as unethical, it is crucial to understand the kind of translated content in question from a pool of three crowdsourcing

models—product-driven, cause-driven, and outsourcing-driven. Product-driven translation is usually done to localize certain freely available documents or open-source software. Cause-driven translation mainly deals with translating content for humanitarian missions that are often carried out by non-profit organizations.

Lastly, outsourcing-driven translation is often an initiative of a for-profit organization that has no specific social or humanitarian endeavors but still wants to rely on the public for translation services. Facebook and Twitter are examples of for-profit organizations that crowdsource their translations from their users through the outsourcing-driven model. Arguably, voluntary outsourcing for product-driven and the cause-driven translation is acceptable, even among professional translations as opposed to outsourcing-driven translation for for-profit organizations.

## **1.9 Methodology**

This study has employed a descriptive qualitative approach to examine and analyze the translated text for Facebook UI content for Kiswahili. The researcher's intention was to examine English-Kiswahili translations used on the Facebook user interface by an analysis of the errors present in this translation.

### **1.7.1 Data Presentation**

Collected data was organized into tables to represent an identified error, the classification of the error and suggested translations. Throughout, the English Facebook User Interface was used to extract the source text while the Kiswahili User Interface was used to extract the target text.

Data was organized into error categories namely Accuracy, Grammar and Stylistic Errors. Each error type has a table with columns which show the source text, target text and suggested translation.

### **1.7.2 Data Analysis**

This study has applied the content analysis method to describe and quantify the collected data. Content analysis is a method that can be used to analyze either qualitative or quantitative data inductively or deductively way. Satu Elo and Helvi Kynga's (2007) say that "qualitative content analysis is commonly used in nursing studies but little has been published on the analysis process and many research books generally only provide a short description of this method." The same authors posit that content analysis is a method of analysing written, verbal or visual communication messages. The latter is the type of data relevant to this study, derived from screenshots extracted from Kiswahili Facebook UI, and therefore is the subject of our analysis.

We have used Content analysis to distil text into fewer content related categories. It is assumed that when classified into the same categories, words, phrases and the like share the same meaning (Cavanagh 1997). We have used content analysis to make replicable and valid inferences from data to their context, with the purpose of providing knowledge, new insights, a representation of facts and a practical guide to action (Krippendorff 1980).

Our aim is to attain a condensed and broad description of the data, and the outcome of the analysis is concepts or categories describing the data. Our purpose is build a group of error categories that help describe the translation errors observed from our data. We have used a deductive approach based on the Gyde Hansen's Classification of Translation Errors, therefore moving from the general to the specific.

Content analysis method involves 3 steps namely preparing, organizing and reporting data (Satu Elo and Helvi Kynga's (2007). The preparation phase starts with selecting the unit of analysis, in this case the Facebook UI content that has been translated from English to Swahili. The sampled errors are a general representation of the translated text from which it is drawn, since it may not be possible to analyze the entire Facebook UI content which has been translated into Swahili.

## CHAPTER TWO

### 2.0 Quality Issues in Crowdsourced Translations

This section provides an extensive discussion of the theoretical issues underpinning translation errors and quality issues in crowdsourced translations. We will look at the question of how crowdsourcing affects quality in translation in general and in Facebook for other languages besides Kiswahili. To achieve this, we will look at the crowdsourcing process independently and how it compares to the standard process of outsourcing, the rationale for crowdsourcing, criticisms against crowdsourcing as well as the domains where translation is permissible, or not permissible and the Facebook crowdsourcing model.

In their study on crowdsourcing translations, Omar Zaidan and Chris Burch of John Hopkins University call it sourcing “professional quality from non-professionals”. This is obviously paradoxical because one would not expect professional quality from non-professional translator. This is further compounded by the fact that crowdsourcing employs a massive crowd, not just a small group of non-professionals. This makes quality control very difficult, if not impossible. These authors concede that “naively collecting translations by crowdsourcing the task to non-professional translators yields disfluent, low-quality results if no quality control is exercised” (Zaidan, O & C. C Burch, 2011). They conclude that it is possible to obtain high-quality translations from non-professional translators but only with the employment of mechanisms which increase the translation quality to near professional levels. These mechanisms involve collecting translations from the crowd indiscriminately, scoring the translations by quality to sort the acceptable and unacceptable, vetting the translations by a select group of professionals or credible non-professionals and finally, editing by yet a smaller group of professionals or credible

non-professionals. The last two steps call for the end client to provide some incentives to the manpower involved. Some clients who may not afford these incentives opt to skip this steps, at the expense of the final quality.

## **2.1 A Comparison of Outsourcing and Crowdsourcing**

Outsourcing, which is the standard process used by translation companies and agencies, is designed to ensure premium quality translations. As a result, the process involves meticulous pre-translation and post-translation steps which collectively make up the LQA process. Before translation, the translation project manager (PM) prepares reference materials such as Style Guides, Translation Memories (TMs) and glossaries which aid the translator in producing high quality translations. A Style Guide provides directions on general principles of translation preferred by the client, such as formality levels, tone, date formats, units of measurements, how to treat brand names and proper nouns etc. TMs are used by the translator to populate previous translations so that they do not have to re-translate previous segments. Glossaries are used to offer definitions and translations of technical terms. The translator also uses spellcheckers within their Computer Assisted Translation (CAT) Tools, available, to make sure there are no sloppy typological errors. These resources enable a pool of translators to maintain consistency and harmony in different files within the same project, without which the end product is compromised.

Another pre-translation step is building the translation team. The most important determining factor of translation quality is the expertise of the translators. Translators must understand the

subject matter they are translating at least as well as the target audience. If the translator gets it wrong, it will be challenging for the other professionals up the line to perfect the translation. The PM selects the most qualified translator(s) for the work at hand. For this they consider the translator's education level, translation experience and specialization among other factors. These qualifications are prerequisite to high quality translations. For instance, a translator who has a 7-year experience in legal content is better positioned to handle a legal translation than a translator with the same level of experience but in a different field, such as medical.

After translation, the PM appoints another professional, preferably of a higher qualification and experience than the translator to review the translation. This is a 2-step process where an editor performs a word-by-word comparison of the source and target text to ensure accuracy, among other qualities, while a proofreader later reads through the target text to ensure that it flows naturally. After this, the proofreader or PM runs a QA check to catch any pending errors such as double spaces, number mismatch, omissions etc.

This methodical LQA process in standard translation process leaves little to no room for errors. While errors may still occur, they are not as critical as when the standard LQA process was not followed. The fact that crowdsourcing does not employ a thorough LQA process as described above results to the errors of the scale witnessed in this study.

Outsourcing entails contracting out internal business processes to identified third parties. Where crowdsourcing leverages on the participation of people with any background, outsourcing employs individuals with specified sets of skills, abilities, and competencies. Outsourcing and crowdsourcing are significantly divergent, each focusing on distinct functions. Crowdsourcing

involves the division of tedious work online. As such, most or all crowdsourcing initiatives are conducted online.

Regarding the labor force, crowdsourcing employs a significantly flexible workforce. All online users can take part in crowdsourcing projects, contribute to the progression of tasks or ideas of interest and earn some kind of remuneration depending on the terms of translation (Yuan, Liang, & Xue, 2014). In contrast, outsourcing processes feature rigid workforces with fixed staffing and distribution models.

The two concepts also differ from each other in their reward schemes. Crowdsourcing providers often pay their workers based on quality output—only the best translated content may be paid for while the rest is discarded without necessarily having to pay (Yuan, Liang, & Xue, 2014). This phenomenon is motivated by the fact that in many cases, there are no legally defined relationships between crowdsourcing clients and the translator. Conversely, outsourcing bases its pricing on headcount and hourly rates—there are well-defined relationships between the companies and their organizational members. Also, crowdsourcing does not rely on fixed costs like outsourcing tasks which usually have some prepayments. Similarly, there are no defined quality bottom lines in crowdsourced labor unlike outsourced labor whose quality is guaranteed under documentations like Service Level Agreement (SLA).

The presence or absence of an LQA process, defined working relationships and incentives are direct and indirect determinants of translation quality. These make outsourcing somewhat superior to crowdsourcing as far as translation quality is concerned. While crowdsourcing greatly saves time and money, this is often at the expense of quality.

## **2.2 The Facebook Crowdsourcing Model and the Trend towards Crowdsourcing**

O'Hagan (2016) appreciates the impact of globalization and the current technological setting on the translation frontier, especially crowdsourced translations. There is an increased need for translation given the surge in daily internet users and organizations as they attempt to communicate with their international audience and potential customers. O'Hagan (2016) notes that this need for translation has led to an increase in free online translation solutions that are mostly grounded on automatic translation (machine translation) on platforms like Facebook Translate, Google Translate, and Microsoft Bing Translator. For instance, Google Translate handles more than 143 billion words daily, a reflection of the extensive and pervasive adoption of crowdsourced translations among internet and social media users (Wolverton, 2018).

The journey of crowdsourcing translation since the inception of the term itself is interesting. In 2008, Facebook challenged its subscribers to engage in a voluntary localization project that aimed at translating the English User Interface into more than 75 languages (Flanagan, 2016). After its inception at Facebook, non-paid crowdsourcing of translations developed into a case study. Today, it is almost impossible to mention crowdsourcing of translations without including Facebook Translate. The translation of the site into many languages was successful and took place in a record time. The trend was rapidly adopted by other sites like Twitter and YouTube. However, when another social networking site, LinkedIn, attempted to use the same strategy, an overwhelming uproar rose up, questioning the essence of crowdsourcing translations for free, especially for a for-profit organization (Newman, 2009). Perhaps this was the trendsetter—today, most professional translators have kept off attempts by for-profit organizations to lure them into crowdsourcing translations for free at the professional scale.

Facebook has displayed notable excellence in motivating unpaid translators to work on the web and undertake large scale translation tasks. Facebook is a pioneer in employing crowdsourcing to translate its User Interface into many languages in a global localization project. Before 2008, Facebook had no localized user interfaces; the site only had the English version. At that time, the company had about 70 million active subscribers, 60% of whom were outside the United States (statista.com). The company was continually seeking ways to serve the international audience more efficiently. A combination of the Facebook buzz, the need to internationalize, and the new concept of crowdsourcing facilitated the localization of Facebook into many languages. The use of professional translators would have been a long, slow, and expensive process—the company opened up its translation platform that prompted users to suggest translations from English to their native languages. The process was fast and inexpensive. Within 24 hours, Facebook was already translated into French by more than 4000 native French speakers (HBS Digital Initiative, 2015). Within two years after the launch of Facebook Translate, the site had been translated into more than 75 languages.

The company has invested time and other resources in the development of its translation tool. Initially, they were criticised for incorrect translations as well as the use of free labor where they would have traditionally invested a lot of money and resources. According to the Facebook, the aim of introducing the translation tool was to allow its users to take a more active role in the growth of the platform. Verification and voting mechanisms were later introduced to improve the crowdsourced translations (HBS Digital Initiative, 2015). Facebook uses the Translate Facebook App, a translation platform which presents users with words that need to be translated into their native languages, and they submit their entries. The system then invites other users to vote on

which translations are best. The winning entries are instantly used in the UI from where the general users access and interact with the Facebook app. The source of Facebook users' motivation in localizing the site remains unknown with arguments that users wanted to realize their potential as translators by attempting to translate accurately while at the same time maintaining a general communication function. Indeed, top contributors for each language are displayed on a weekly leaderboard. While there are style guides for some languages meant to ensure that there is consistency in the translation, they do not seem to have worked based on the frequency and type of errors resulting from the translations.

There is still a notable growth in the amount of digital content that needs translation. This trend is confronted by another trend—there is an insufficient number of professional translators to meet individual, organizational, and business demands as far as translation is concerned. With the trend of an increase in the content that will need translation set to prevail, the place of professional translators is being uptaken by another type of translators that has no interest in monetary remunerations for their translation labor—social translators (Flanagan, 2016). This group of translators derives its motivation to translate content from their interests in a given topic. For this reason, social translators, though not necessarily professional, are considered as a threat to professional translators.

Crowdsourcing of translations, especially by for-profit organizations has spiked the discussions regarding the ethics and fairness of this trend. While some schools of thought have hailed the crowdsourcing of translations for enabling the access of online content by marginalized and disadvantaged people, others have criticized, terming it as a violation of workplace regulations (Schlagwein, Cecez-Kecmanovic, & Hanckel, 2019). Since crowdsourcing is a relatively new

phenomenon in the translation industry, one can only wait to see the direction that it will take in the future. The fact remains that, in spite of the quality issues associated with crowdsourcing, it is still a preferred translation approach especially for social media companies.

### **2.3 The Rationale behind Crowdsourcing Translations**

Various digital companies have thrived through the benefits offered by crowdsourcing translations. It is fast, free or low cost, and the fact that users are also producers gives the advantage of creating and testing a translation product instantly. Through this, some disadvantaged people, subscribers, or users who are sometimes denied the thrill of online presence due to language barriers are given the opportunity to join their more advantaged counterparts on the digital space thanks to crowdsourcing.

Dolmaya (2011) argues that before judging crowdsourcing translations as unethical as is often the case among professional cohorts, it is crucial to understand the kind of translated content in question from a pool of three crowdsourcing models—product-driven, cause-driven, and outsourcing-driven. Cause-driven translation mainly deals with translating content for humanitarian missions that are often carried out by non-profit organizations. Lastly, outsourcing-driven translation is often an initiative of a for-profit organization that has no specific social or humanitarian endeavors but still wants to rely on the public for translation services.

Facebook and Twitter are examples of for-profit organizations that crowdsource their translations from their users through the outsourcing-driven model. Arguably, voluntary outsourcing for product-driven and the cause-driven translation is acceptable, even among professional translations as opposed to outsourcing-driven translation for for-profit

organizations. The translation of the Facebook User Interface into Kiswahili gives ordinary people the opportunity to have a presence on social media, albeit with some level of guidance from those who are bilingual. Similarly, some users simply want to enjoy using online platforms in a language that they can own and identify with. Where the English language is widely deemed a foreign language, people may use the Kiswahili Facebook Interface because of their appreciation for the local language.

Where monetary gains are not integral aspects of crowdsourcing translations, then what motivates participants to contribute to the initiator's call for translations? According to Pinto and dos Santos (2018), where monetary benefits are not involved, contributors in crowdsourcing activities derive most of their motivation intrinsically through the fun, learning, and satisfaction as well as extrinsically through acknowledgment by the organization that initiates the crowdsourcing. The findings of this study mean that so long as volunteer translators remain intrinsically motivated as they participate in crowdsourcing translations through various online platforms like Facebook and Google, such crowdsourcing will remain resilient regardless of the perceived quality issues by professional translators. The emphasis of this premise is that to some extent, commercializing translation is not always right—people, particularly the amateur, should be allowed to explore the fun and satisfaction that can be gained from translating their content of interest.

## 2.4 Voices of Antagonism towards Crowdsourcing

The benefits of crowdsourcing translations are undeniable, especially for digital companies and social media users. However, there are various voices that oppose it, especially among scholars and professional translators, mainly due to the ethical and quality issues surrounding the practice.

Pascoal, Furtado, and Chorão (2017) reiterate that the onset of localized crowd translations particularly on Facebook in 2008 resulted to an acceleration in crowdsourced translations such that the practice is now becoming the next big thing in translation. However, crowdsourced translations are widely criticized, particularly by professional translators, for being quick-fixes that prioritize immediacy, convenience, and cost over quality and professional assistance. Pascoal, Furtado, and Chorão (2017) do not shy away from highlighting the controversies among professional as a result of using volunteer work and fans to translate content. Professional translators view crowdsourcing as an impediment to their profession and perceive it to violate the integrity of translated content in terms of quality.

In their study, Pascoal, Furtado, and Chorão (2017) found that a majority (60%) of the professional translators who featured in the study viewed crowdsourcing translation as a threat to their profession. They also believed that crowdsourcing hurts the quality of translation. According to professional translators, translation is more than just a combination of language knowledge and the drive of becoming a translator. Professional translators also hold the view that crowdsourced translations should not be taken seriously since volunteer translators are mostly amateurs in the field. These findings underline the dilemma of settling for the immediacy, convenience, and low-cost of crowdsourced translations at the expense of the quality,

dependability, and integrity offered by professional translators at a comparably slower rate and higher cost.

Kaminskienė and Kavaliauskienė (2012) reviewed the skills and competencies necessary for translation and interpreting. The survey featured employers of the European Union institution, lecturers, alumni, and Vilnius university students. The study unearthed seven core competencies necessary for one to become a reliable translator. These are technological competencies, language competence, cultural competence, production dimension, interpersonal dimension, thematic competence, and information mining competence. In a similar study, Calvo (2011) attempts to distinguish between translator competence and translation competence or whether the two are synonymous concepts. The study concluded that translation competence is a product of professional translator qualifications.

Dolmayer (2011) raises a question of ethics that was mentioned briefly herein, especially because crowdsourcing translations rely on volunteer labor from people that do not necessarily possess the training nor competencies needed in translation. Moreover, Dolmayer (2011) criticizes the use of free labor by giant organizations that are perceptibly capable of remunerating the crowd involved in translation. While there are several benefits attached to crowdsourcing translations such as enhancing the visibility of translation, helping smaller languages to gain more visibility online, and showcasing the value of crowdsourcing to the society, there is an underlying concern of devaluing the role of professional translation as well as lowering professional translator's occupational status.

Some translation content also requires austere protection, and therefore cannot be translated by the public. The translators employed in the content of this nature are bound by Non-Disclosure

Agreements (NDAs) against sharing the content outside the project network. The translators are also highly skilled and experienced in these specific domains, where quality cannot be compromised. Examples of such content is patents, contracts, personal data such as medical records, immigration details, marriage information etc. Besides the sensitivity and privacy of this content, its complexity and difficulty level is too risky to entrust to the crowd for translation.

So, not everyone is an admirer of crowdsourcing, progressive as it may be, and advantageous to users. The main antagonists are professional translators and scholars, especially due to the ethical questions it raises and quality concerns associated with its end product.

## **2.5 Pro-crowdsourcing Domains**

Due to the quality issues associated with crowdsourced translations, some sensitive domains may not prefer this method. Instead, they go traditional by outsourcing for vetted professional translators and employ stringent Linguistic Quality Assurance (LQA) processes to ensure premium quality. Some of these domains are legal, medical and engineering, where the margin for errors is very slim. Errors in this domain may lead to serious issues such as death of a patient in case a dosage is wrongly translated, misapplication of a legal concept or malfunctioning of a machine whose manual was mistranslated in one area of the text or another.

Domains whose text has less serious implications if mistranslated have more confidence employing crowdsourcing and enjoy the advantages cited earlier. Digital companies and social media platforms particularly prefer this approach, as we have seen with Facebook, Twitter and Instagram, among others. This is because their content may not have serious legal or even business implications in case of gross translations errors. Facebook, being a social network

where users interact with their friends through posts, pictures, videos, comments and reactions (represented by emojis and acronyms signifying laughter, smiles, sadness, anger or awe) uses easy and non-technical language in its user interface. There are also neologisms and semantic shifts that are particularly common on social media language and online culture. For example, outside social media, the words 'like' is seldom used as a noun as used in Facebook. This complicates translation because while English has made the move of widely adopting neologisms and semantic shifts, smaller languages like Kiswahili are caught unawares. This means that the translator must invent an equivalent term and impose it upon the user, and by extension, to the general Kiswahili audience. This becomes even more complicated when crowdsourcing is used, because the translators, majority of whom are non-professionals, have to assume the role of terminologists at the same time they are translators, something that is quite demanding, even for a professional translator. This might be one of the reasons for the errors witnessed during this research

## **2.6 Crowdsourcing for Kiswahili Translations**

Swahili translations may not have attracted a lot of attention as far as crowdsourcing is concerned. The main outsourcer to call for Kiswahili translations through crowdsourcing is Ushahidi, which is an international tech organization developed to map reports of violence in Kenya after the post-election violence in Kenya in 2008. Crowdsourced translators volunteered to translate Swahili, Sheng and vernacular messages on violence situations into English. Ushahidi then worked with the Kenya Police to handle potential security situations based on the triangulated areas from which the messages originated from.

Ushahidi says that since then, thousands of people have used Ushahidi's crowdsourcing tools to raise their voice. They are a technology leader in Africa, headquartered in Nairobi, with a global team. Ushahidi is a social enterprise that provides software and services to numerous sectors and civil society to help improve the flow of information from bottom-up.

Another organization that has utilized crowdsourced translations is Translators Without Borders (TWB), which is non-profit organization set up to provide translation services for humanitarian non-profits. As of September 22<sup>nd</sup> 2019, had translated over 82 million words in several world languages, including Swahili. In April 2012, TWB International set up a training facility in Nairobi, Kenya, which began its operations by training 20 translators of medical texts, drawn from various professional backgrounds. Since opening, the Kenya Training Center has trained almost 250 people, many of whom continue to work on TWB translations in Kiswahili and a dozen other east African languages. TWB's incentive to crowdsourced translators that volunteer to work with them has ranged from branded T-shirts, \$50 dollar appreciation fees and recognition in the media that cover their work, among others.

As discussed in this study, Facebook is also another major organization that has taken advantage of crowdsourcing to have their content translated into Swahili, and other world languages.

## **2.6 Conclusion**

In conclusion, crowdsourcing does not employ a stringent LQA process like standard outsourcing does, which is a high risk associated with crowdsourced translations. The use of a mix of professional and non-professional translators further compounds this situation. However, Facebook and other social media companies prefer this method due to its advantages in terms of the availability of cheap labor, quick turn-arounds for huge word counts as well as turning users into producers. Scholars and professional translators are the most vocal critics of crowdsourcing. There are some domains that do not favor crowdsourced translations due to the sensitivity and complexity of their content.

## **CHAPTER THREE**

### **3.0 Data Presentation**

This chapter is on presentation of the data collected from the Kiswahili translations on the Facebook User Interface. The collected data has been classified into error categories namely Accuracy, Grammar and Stylistic Errors. Each category is represented by a 3-column table, each column showing the source text, translated texts with the errors, and the suggested translation, with the corrected version of each translation.

We will first discuss the concept of “errors” in brief to enable us to move on to classifying them. Hansen (2010) poses the question, “What is an error and what is a translation ‘error’?” and goes further to explain that an “error” usually means that something is wrong. He says that in written texts, errors can be classified as, for example, pragmatic, semantic, idiomatic, orthographic, linguistic or stylistic errors. A translation error, with the assumption that a translation is the production of a Target Text (TT) which is based on a Source Text (ST), arises from the existence of a relationship between two texts. So if there is an incongruity between the ST and the TT whether in typology, semantics, grammar etc., then there is a translation error.

This raises the question of formal and dynamic equivalence as fronted by Nida (1995). Formal equivalence is concerned with structural equivalence of the ST and TT, in a lexical and syntactic sense, in that sentences and words in the ST are accounted for in the TT, leading to a literal translation. There domains where this approach is critically necessary, such as in legal text, where the spirit and the letter, as it were, have to be preserved. On the other hand, dynamic equivalence describes semantic equivalence of the ST and TT, such that the focus is retaining the

original meaning as much as possible, without endeavoring to maintain the form and structure of the ST at the expense of the intended meaning and message. Indeed, as observed in this study, most of the errors have occurred because the translator has over-applied formal equivalence.

The data in this study is categorized according to Gyde Hansen's classification as discussed in his 2010 work entitled 'Translation "Errors"'. We will also borrow from the Content analysis method, which involves 3 steps namely preparing, organizing and reporting data (Satu Elo and Helvi Kynga's (2007). In this chapter we will work on the first 2 steps, namely preparing and organizing data, while the third step will be handled in the next chapter. The preparation phase starts with selecting the unit of analysis, in this case the Facebook UI content that has been translated from English to Swahili. The sampled errors are a general representation of the translated text from which it is drawn, since it may not be possible to analyze the entire Facebook UI content which has been translated into Swahili.

The data we collected was prepared by taking screenshots of translation errors of Kiswahili Facebook UI through a mobile phone. All errors encountered during the research period were captured indiscriminately. The researcher then compared the ST and TT to determine the validity of the errors, then copied both the ST and TT words in one main table of 2 columns in random order of the errors. In this regard, we will discuss Accuracy, Grammar and Stylistic Errors. Accuracy Errors result from addition of content in the TT that is absent in the ST, ambiguous translation of the ST, incorrect translation of the ST, misunderstanding of the source concept, omission of essential elements, translations that are out of context as well as untranslated texts. Grammatical errors involved spelling and punctuation errors, but more importantly for this study,

syntactical errors, i.e. wrong sentence structure. The last category is on Style, and especially on literal translation and wrong use of tone in the translated text.

The table below is what we arrived at:

<b>SOURCE TEXT</b>	<b>TARGET TEXT</b>
Your post	Chapisho yako
His comment	Maoni zake
[Group Name] has a new post	[Group Name] ana chapisho jipya
[Group Name] has a new photo	[Group Name] ana picha mpya
[Number] people using this	Watu [Number] wanatumia hii
Who can see this	Je, ni nani anaweza kuona hii
Alfred's post	Chapisho ya Alfred
...in Swahili	...kwenye Kiswahili
[App Name] will receive your personal information	[App Name] atapokea maelezo yako ya kibinafsi
[Username] joined April 2007	[Username] amejiunga Aprili 2007

Message (verb)	Ujumbe
SOURCE TEXT	TARGET TEXT
Search nearby	Utafutaji wa karibu
Active	Amilifu
[Number] Views	Hakiki [Number]
[Username] is in a relationship	[Username] katika uhusiano
Back home	Rudi hadi nyumbani
Interact with content...	Kuingiliana na maudhui...
Go to free	Nenda kwa bila malipo
You are on data mode	Uko kwenye modi ya data
Share	Kushiriki
Shared (number of times)	Kushiriki
You became friends	Mulikuwa marafiki
Mutual friends	Marafiki wa karibu
Mutual friends	Marafiki mnaojuana

Uncover photo	Picha ya ufichuzi
Language and region	Lugha na dini
Posts you were tagged in	Chapisho ulizohusika nazo
[Username] likes your comment	[Username] amependezwa na toa maoni yako
[Username] is feeling sick	[Username] anahisi najisikia vibaya
You have one new like	Una imependwa upya mara moja
[Group Name] has 4 new likes and 1 post like	[Group Name] una Imependwa upya mara 4 na POa mpya 1 ya chapisho
[Username] and 313 others have reacted to your post	[Username] na Watu wengine 313 ameonyesha hisia kwenye chapishoendelea yako
Write something to [Username]	Andika jambo katika [Username]
Done	Imekamilika
[Group Name] has a new picture from [Username]	[Group Name] ana picha mpya kutoka kwa [Username]
Like page	Pendezwa na ukurasa
Browse for help	Vinjari Msaada

Closed group	Kundi kilichofungwa
Wave (verb)	Mawimbi
Use your \$5.00 credit	Tumia mkopo wako wa US\$5.00

*Table 3.1. Data Preparation (Researcher's Creation)*

After data was prepared as shown above, it was then organized into error categories according to Gyde Hansen's classification of translation errors. In each category, we have a three-column table showing the ST, TT and Suggested translation. The ST column is a written record of the English Facebook UI whose TT equivalent, in the TT column contained translation errors. The last column is the researcher's suggested correction which offers a solution to the errors shown in the TT column. After this organization, the following categories were arrived at.

### **3.1 Stylistic Errors**

According to Hansen, these type of errors occur when a translation employs a wrong form, art or method in translation due to literalness and wrong tone. In this study, the main stylistic error is literal translation, where the translation gives a word-for-word rendition of the source text resulting to an awkward, unidiomatic and incorrect target text. Here, the translator focuses on each word independently, instead of translating the sentence as a whole, so that instead of conveying the intended message, they render a direct translation.

Let is consider the following errors extracted from Kiswahili Facebook UI.

*Table 3.2 Stylistic Errors*

	<b>SOURCE TEXT</b>	<b>TARGET TEXT</b>	<b>SUGGESTED TRANSLATION</b>
<b>1.</b>	Active	Amilifu	Yupo Mtandaoni
<b>2.</b>	Search nearby	Utafutaji wa karibu	Tafuta maeneo yaliyo karibu nawe
<b>3.</b>	[Number] Views	Hakiki [Number]	Imetazamwa Mara [Number]
<b>4.</b>	[Username] is in a relationship	[Username] katika uhusiano	[Username] ameanzisha uhusiano
<b>5.</b>	Back home	Rudi hadi nyumbani	Rudi mwanzo
<b>6.</b>	Interact with content...	Kuingiliana na maudhui...	Kuchangia katika maudhui...
<b>7.</b>	Go to free	Nenda kwa bila malipo	Tumia data bila malipo
<b>8.</b>	You are on data mode	Uko kwenye modi ya data	Unatumia data ya simu
<b>9.</b>	Write something to [Username]	Andika jambo katika [Username]	Mwandikie [Username] jambo

<b>10.</b>	Done	Imekamilika	Nimekamilisha
<b>11.</b>	[Group Name] has a new picture from [Username]	[Group Name] ana picha mpya kutoka kwa [Username]	[Name of User] amechapisha picha mpya kwenye [Username]
<b>12.</b>	Like page	Pendezwa na ukurasa	Penda ukurasa
<b>13.</b>	Browse for help	Vinjari Msaada	Tafuta Usaidizi
<b>14.</b>	Closed group	Kundi kilichofungwa	Kundi la siri/kundi maalum
<b>15.</b>	Wave (verb)	Mawimbi	Mpungie mkono

### 3.2 Accuracy Errors

Accuracy errors occur when the target text does not respect and convey the source text meaning as much as possible, resulting to an altered meaning. The attempt to have a creative rendition of the target text may result to errors by deviating from the original meaning. Syntactical changes resulting to an altered order of the translation's major elements destroys the text's flow, changes its emphasis or obscures the author's intent. Wrong word choice, semantic discord and wrong lexicon can also lead to accuracy errors.

The following texts are sampled from Facebook to illustrate the Faithfulness errors present on the Kiswahili user interface.

*Table 3.4 Accuracy Errors*

	<b>SOURCE TEXT</b>	<b>TARGET TEXT</b>	<b>SUGGESTED TRANSLATION</b>
1.	Share	Kushiriki	Shiriki
2.	Shared (number of times)	Kushiriki	Imeshirikiwa mara
3.	You became friends	Mulikuwa marafiki	Mlianza urafiki
4.	Mutual friends	Marafiki wa karibu	Marafiki wa wenza
5.	Mutual friends	Marafiki mnaojuana	Marafiki wa wenza
6.	Uncover photo	Picha ya ufichuzi	Fichua picha
7.	Language and region	Lugha na dini	Lugha na eneo
8.	Posts you were tagged in	Chapisho ulizohusika nazo	Chapisho ulizotambuliwa
9.	[Username] likes your comment	[Username] amependezwa na toa maoni yako	[Username] amependezwa na jibu lako
10.	[Username] is feeling sick	[Username] anahisi <b>najisikia</b> vibaya	[Username] anaugua
11.	You have one new like	Una <b>imependwa upya mara moja</b>	Mtu mmoja ameipenda

12.	[Group Name] has 4 new likes and 1 post like	[Group Name] una Imependwa upya mara 4 na POa mpya 1 ya chapisho	Watu 4 wapya wamependa [Group Name] na 1 mgeni kupenda chapisho lako
13.	[Username] and 313 others have reacted to your post	[Username] na Watu wengine 313 ameonyesha hisia kwenye chapishoendelea yako	[Username] na watu wengine 313 ameonyesha hisia kwenye chapisho lako
14.	Use your \$5.00 credit	Tumia mkopo wako wa US\$5.00	Tumia ofa yako ya US\$5.00

### 3.3 Grammatical Errors

This error type occurs when a translation violates the grammatical rules of the target language. Grammatical errors include lack of agreement between a subject and a verb, incorrect verb inflections, and improper declension of nouns, pronouns and adjectives, among other word classes.

The following are some of the errors that were extracted on Kiswahili Facebook UI under this category.

Table 3.5 Grammatical errors

	SOURCE TEXT	TARGET TEXT	SUGGESTED TRANSLATION
1.	Your post	Chapisho <b>yako</b>	Chapisho lako
2.	His comment	Maoni <b>zake</b>	Maoni yake
3.	[Group Name] has a new post	[Group Name] <b>ana</b> chapisho jipya	Chapisho jipya limeandikwa kwenye [Group Name]
4.	[Group Name] has a new photo	[Group Name] <b>ana</b> picha mpya	Picha mpya imechapishwa kwenye [Group Name]
5.	[Number] people using this	Watu [Number] wanatumia hii	Watu [Number] wanatumia programu hii
6.	Who can see this	Je, ni nani anaweza kuona <b>hii</b>	Mtu anayeweza kuona maelezo haya
7.	Alfred's post	Chapisho <b>ya</b> Alfred	Chapisho la Alfred
8.	...in Swahili	... <b>kwenye</b> Kiswahili	...kwa Kiswahili
9.	[App Name] will receive your personal information	[App Name] <b>atapokea</b> maelezo yako ya kibinafsi	[App Name] itapokea maelezo yako ya kibinafsi

<b>10.</b>	[Username] joined April 2007	[Username] <b>amejiunga</b> Aprili 2007	[Username] alijiunga Aprili 2007
<b>11.</b>	Message (verb)	Ujumbe	Mtumie ujumbe

### 3.4 Conclusion

The errors captured during this study were classified into 3 major categories. It is obviously possible to create numerous categories, but in this study we filtered the data into these 3 categories namely Accuracy, Grammar and Stylistic Errors. The errors in each category will be described and analysed in general in the next chapter.

## **CHAPTER FOUR**

### **4.0 Data Analysis, Discussion and Findings**

This study has applied the content analysis method to describe and quantify the collected data. Content analysis is a method that can be used to analyze either qualitative or quantitative data inductively or deductively way. Satu Elo and Helvi Kynga's (2007) say that "qualitative content analysis is commonly used in nursing studies but little has been published on the analysis process and many research books generally only provide a short description of this method." The same authors posit that content analysis is a method of analysing written, verbal or visual communication messages. The latter is the type of data relevant to this study, derived from screenshots extracted from Kiswahili Facebook UI, and therefore is the subject of our analysis.

We have used Content analysis to distil text into fewer content related categories. It is assumed that when classified into the same categories, words, phrases and the like share the same meaning (Cavanagh 1997). We have used content analysis to make replicable and valid inferences from data to their context, with the purpose of providing knowledge, new insights, a representation of facts and a practical guide to action (Krippendorff 1980).

Our aim is to attain a condensed and broad description of the data, and the outcome of the analysis is concepts or categories describing the data. Our purpose is build a group of error categories that help describe the translation errors observed from our data. We have used a deductive approach based on the Gyde Hansen's Classification of Translation Errors, therefore moving from the general to the specific.

Content analysis method involves 3 steps namely preparing, organizing and reporting data (Satu Elo and Helvi Kynga's (2007). The first 2 steps, namely data preparation and organization were handled in the previous chapter on Data presentation. In this chapter, we will go to the final step, which is data reporting, commonly known as data analysis. In order to save ourselves from redundancy, we will analyze a number of errors in each category and use them to draw a general conclusion that applies to the rest of the errors in that specific category.

#### **4.1 Stylistic Errors**

As mentioned in the previous chapter, these type of errors occur when a translation employs a wrong form, art or method in translation by employing a literal approach to the understanding of the ST and composition of the TT. Literal translation is apparently the main stylistic error evidenced by the collected data. The translation gives a word-for-word rendition of the source text resulting to an awkward, unidiomatic and incorrect TT. Here, the translator focuses on each word independently, instead of translating the sentence as a whole, so that instead of conveying the intended message, they render a direct translation, word for word translation. Let's consider the following errors in this category.

*Table 4.1, Stylistic Errors*

	<b>SOURCE TEXT</b>	<b>TARGET TEXT</b>	<b>SUGGESTED TRANSLATION</b>
<b>1.</b>	Active	Amilifu	Yupo Mtandaoni
<b>2.</b>	Search nearby	Utafutaji wa karibu	Tafuta maeneo yaliyo karibu nawe
<b>3.</b>	[Number] Views	Hakiki [Number]	Imetazamwa Mara [Number]
<b>4.</b>	[Username] is in a relationship	[Username] katika uhusiano	[Username] ameanzisha uhusiano
<b>5.</b>	Back home	Rudi hadi nyumbani	Rudi mwanzo
<b>6.</b>	Interact with content...	Kuingiliana na maudhui...	Kuchangia katika maudhui...
<b>7.</b>	Go to free	Nenda kwa bila malipo	Tumia data bila malipo
<b>8.</b>	You are on data mode	Uko kwenye modi ya data	Unatumia data ya simu
<b>9.</b>	Write something to [Username]	Andika jambo katika [Username]	Mwandikie [Username] jambo

10.	Done	Imekamilika	Nimekamilisha
11.	[Group Name] has a new picture from [Username]	[Group Name] ana picha mpya kutoka kwa [Username]	[Name of User] amechapisha picha mpya kwenye [Username]
12.	Like page	Pendezwa na ukurasa	Penda ukurasa
13.	Browse for help	Vinjari Msaada	Tafuta Usaidizi
14.	Closed group	Kundi kilichofungwa	Kundi la siri/kundi maalum
15.	Wave (verb)	Mawimbi	Mpungie mkono

Let's consider a few examples of the errors captured. The standalone word "Active" on the English Facebook UI means that some is currently online and available for a live message chat. Facebook uses "*Amilifu*" as its translation. On the linguistic level, this is correct, but as far as translation is concerned, it is not communicative. First, "*Amilifu*" is an uncommon term, which means that users will need to look it up. Secondly, even if the users readily understood it, it does not quite capture the intended meaning, which is simply "*Yupo mtandaoni*", which roughly translates to "Online now". "[Username] is in a relationship" is used when someone begins a romantic relationship and makes it public. If we replace the Username tag with a name of a person we would arrive at something like "John is in a relationship". The Kiswahili translation on Facebook UI would give us "*John katika uhusiano*". While any Kiswahili reader would

understand what the translation attempts to communicate, this simple message is not delivered a natural and clear style. A good alternative translation would be “*John ameanzisha uhusiano*”.

“Closed group” means a private Facebook group, whose content is not visible to anyone outside the group, until they are enlisted as members. Translating it as “*Kundi kilichofungwa*” is inaccurate on a grammatical level, but more importantly, on a semantic level. The translation implies that the group has been closed, which then misleads the user. A correct translation would be “*Kundi la siri*” which is simply “private group”. The word “Wave” appears on the Facebook Messenger UI to prompt the user to send a hand wave emoji to an active friend. This is therefore a verb. Translating it as “*Mawimbi*”, which is the plural for the noun “wave” totally loses the user, as the translation does not even have a remote relation to the ST. A good alternative would be “*Mpungie mkono*” meaning “Wave to them”. Finally, “Write something to [Username]” appears on a user’s profile on their birthday to encourage a friend to send them birthday wishes. It’s translation; “*Andika jambo katika [Username]*” is awkward and uncommunicative. A simple and straightforward translation would be “*Mwandikie [Username] jambo*”.

If a text is inaccurate, it hinders communication, and in the case of Facebook, affects user experience. This is the case with the above errors as extracted from Facebook.

## **4.2 Accuracy Errors**

Accuracy errors occur when the target text does not respect and convey the source text meaning as much as possible, resulting to an altered meaning. The attempt to have a creative rendition of the target text may result to errors by deviating from the original meaning. Syntactical changes resulting to an altered order of the translation’s major elements destroys the text’s flow, changes

its emphasis or obscures the author’s intent. Wrong word choice, semantic discord and wrong lexicon can also lead to accuracy errors.

*Table 4.2 Accuracy Errors*

	<b>SOURCE TEXT</b>	<b>TARGET TEXT</b>	<b>SUGGESTED TRANSLATION</b>
1.	Share	Kushiriki	Shiriki
2.	Shared (number of times)	Kushiriki	Imeshirikiwa mara
3.	You became friends	Mulikuwa marafiki	Mlianza urafiki
4.	Mutual friends	Marafiki wa karibu	Marafiki wa wenza
5.	Mutual friends	Marafiki mnaojuana	Marafiki wa wenza
6.	Uncover photo	Picha ya ufichuzi	Fichua picha
7.	Language and region	Lugha na dini	Lugha na eneo
8.	Posts you were tagged in	Chapisho ulizohusika nazo	Chapisho ulizotambuliwa
9.	[Username] likes your comment	[Username] amependezwa na toa maoni yako	[Username] amependezwa na jibu lako
10.	[Username] is feeling sick	[Username] anahisi <b>najisikia</b> vibaya	[Username] anaugua

11.	You have one new like	Una <b>imependwa upya mara moja</b>	Mtu mmoja ameipenda
12.	[Group Name] has 4 new likes and 1 post like	[Group Name] una Imependwa upya mara 4 na POa mpya 1 ya chapisho	Watu 4 wapya wamependa [Group Name] na 1 mgeni kupenda chapisho lako
13.	[Username] and 313 others have reacted to your post	[Username] na Watu wengine 313 ameonyesha hisia kwenye chapishoendelea yako	[Username] na watu wengine 313 ameonyesha hisia kwenye chapisho lako
14.	Use your \$5.00 credit	Tumia mkopo wako wa US\$5.00	Tumia ofa yako ya US\$5.00

“Language and region” was translated as “*Lugha na dini*”, the error here being that the translator read and translated “region” as “religion”, and there was no editorial work to correct such a simple but significant error. “Uncover photo” is used when a graphic image has been covered so that the viewer has the discretion to uncover and see it. The Kiswahili equivalent translates it as “*Picha ya ufichuzi*”, which is a noun phrase, instead of a verb phrase, that does not make much pragmatic sense. The closest back-translation we can get is “An uncovering photo”. The translator intended to say “Fichua picha” (uncover/unhide photo) but apparently didn’t give much consideration to the message. In this case, the translation is supposed to be a command that

prompts an action from the user, but one wonders what action the user takes presented with such a translation that inaccurate.

“*Chapisho ulizohusika nazo*” is the translation for “Posts you were tagged in”. The error here is that the meaning has been changed from “you were tagged in” to “you contributed in”. A straightforward equivalent would be “*Chapisho ulizotambuliwa*”, whose back-translation is exactly equivalent to the ST. “[Group Name] una Imependwa upya mara 4 na POa mpya 1 ya chapisho” is a very poor attempt of communicating that “[Group Name] has 4 new likes and 1 post like”. Besides the wrong capitalizations within the sentence and in the word “poa”, the translator used a rather literal approach to translate the ST. The translator forces the ST subject to remain the TT subject, which produces an awkward and unnatural translation. A more fluid and natural rendition would be “*Watu 4 wapya wamependa [Group Name] na 1 mgeni kupenda chapisho lako*”. In the suggested translation, [Group Name] changes to the object without affecting the meaning at all.

### **4.3 Grammatical Errors**

This error type occurs when a translation violates the grammatical rules of the target language. Grammatical errors include lack of agreement between a subject and a verb, incorrect verb inflections, and improper declension of nouns, pronouns and adjectives, among other word classes. Now, this category has errors that may seem very petty and insignificant at face value. The errors involve wrong number inflections here and there, wrong tense and prepositions. Sheng’, (an urban language spoken in Nairobi and other Kenyan urban areas) which heavily borrows from both English and Kiswahili, does not have strict grammatical rules (Mukwana, 2015). What matters in its usage is easy communication. Grammatical rules are broken

chaotically as it is more of a code of communication than standard language like English or Kiswahili. Its influence to Kiswahili, especially in Kenya is undeniable, to the extent that the same noncompliance to grammatical rules may be transferred to Kiswahili (Mukwana, 2015). However, in the corridors of education and scholarship, it should never be permitted that Kiswahili grammatical rules are broken and dismissed as inconsequential.

The following are some of the grammatical errors observed in this study. The errors have been sub-categorized into the following groups, based on their subtypes.

*Table 4.3, Accuracy Errors*

	<b>SOURCE TEXT</b>	<b>TARGET TEXT</b>	<b>SUGGESTED TRANSLATION</b>
<b>1</b>	Your post	Chapisho <b>yako</b>	Chapisho lako
<b>2</b>	His comment	Maoni <b>zake</b>	Maoni yake
<b>3</b>	[Group Name] has a new post	[Group Name] <b>ana</b> chapisho jipya	Chapisho jipya limeandikwa kwenye [Group Name]
<b>4</b>	[Group Name] has a new photo	[Group Name] <b>ana</b> picha mpya	Picha mpya imechapishwa kwenye [Group Name]
<b>5</b>	[Number] people using this	Watu [Number] wanatumia hii	Watu [Number] wanatumia programu hii
<b>6</b>	Alfred's post	Chapisho <b>ya</b> Alfred	Chapisho la Alfred

7	...in Swahili	... <b>kwenye</b> Kiswahili	...kwa Kiswahili
8	[App Name] will receive your personal information	[App Name] <b>atapokea</b> maelezo yako ya kibinafsi	[App Name] itapokea maelezo yako ya kibinafsi
9	[Username] joined April 2007	[Username] <b>amejiunga</b> Aprili 2007	[Username] alijiunga Aprili 2007
10	Message (verb)	Ujumbe	Mtumie ujumbe

### 4.3.1 Wrong Pronouns

In errors 1 to 4, wrong pronouns are used, leading to grammatical errors. In the 1<sup>st</sup> and 2<sup>nd</sup> error, the pronouns “*yako*” and “*zake*” are not in agreement with the noun classes of the preceding words. In errors 3 and 4, the pronoun “*ana*” is normally used for animate things, so using it for the name of a Facebook group is incorrect. The underlying issues seems to be the lack of translator knowledge on how to handle placeholders, such as [App Name]. This error is also reflected in error number 9.

For errors 5 and 6, the pronoun “*hii*” is used without in any referent. While this is permissive in the English text, it is not so in Kiswahili, because Kiswahili pronouns inflect according to the word class of the noun in question. For instance, “*hii*” is correct when used with “*picha*”, but incorrect when used with “*maoni*”. It is therefore important for the translator to mention noun being referred to for clarity.

### **4.3.2 Wrong Prepositions**

Errors 7 and 8 are caused by the use wrong prepositions that are not in agreement with the referents. Instead of “*chapisho la*”, the translator uses “*chapisho ya*” causing a sloppy grammar error. While these could be categorized as minor grammar errors, they are errors, which may not occur if a proper LQA process is performed like in the standard outsourcing process.

### **4.3.3 Wrong Tense Marker**

In error 10, the source text “[Username] joined April 2007” clearly indicates that the time being referred to is way in the past, but using “*amejiunga*” in the translation “[Username] *amejiunga Aprili 2007*” incorrectly implies that the action was performed moments ago rather than several years ago.

### **4.3.4 Wrong Noun Class**

For errors 11 and 12, the source texts are verbs based on their context, which the translator(s) misinterpreted for nouns. This leads to a total breakdown in communication. For example, instead of translating “Wave” as “*Mpungie mkono*” to indicate that the user should wave to another person, the translation stands for a wave, such as a sea or sound wave.

## **Conclusion**

A total of 40 translation errors were sampled from Facebook as shown in Table 1. Majority of the translation errors (37.5%) were stylistic, accuracy errors were at 35% while grammatical errors were 28%. The total number of translated words sampled as 500, where 100 errors were

found. Of these 40 errors were sampled to represent the 100 errors, as indeed there were multiple repeated errors. Figure 4.1 is a visual presentation of the distribution of errors as collected from the Kiswahili Facebook User Interface.

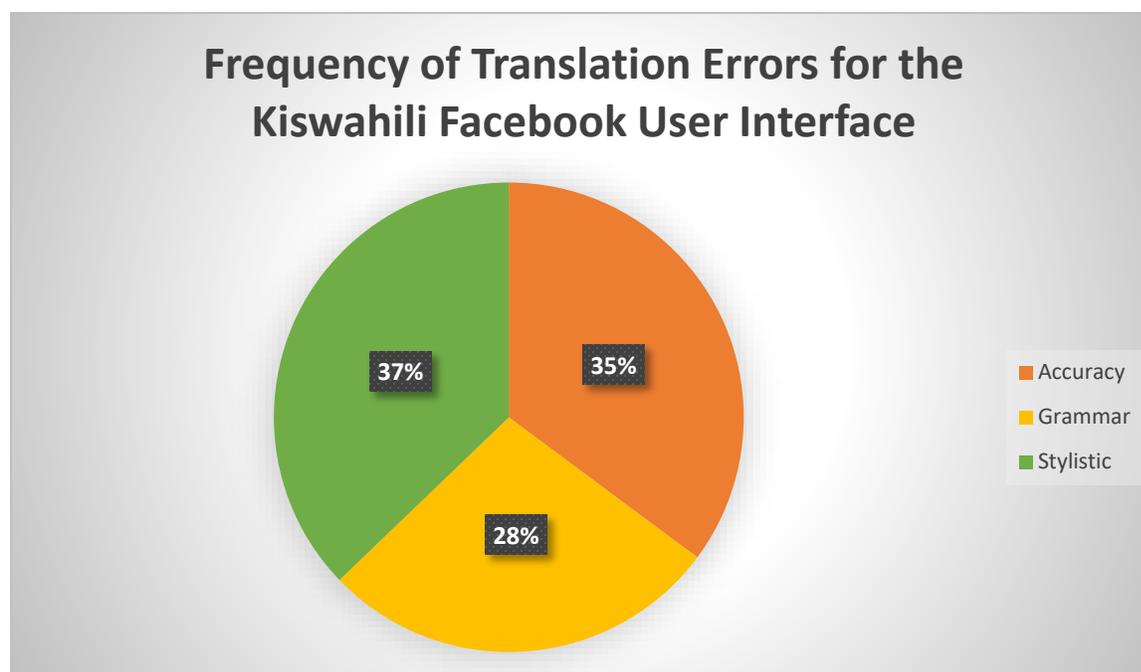


Figure 4.1

*Translation Errors for the Kiswahili Facebook User Interface*

#### **4.4 Apparent Translation Strategies Used by Facebook Translators**

From the data collected and the errors analysed, the researcher can deduce the apparent translation strategies used by Facebook translators for Kiswahili UI. These are literal translation, word-for-word translation and oblique translation. As Nida (1995) argues, in formal equivalence, which yields a literal translation the TT closely resembles the ST in both form and content. When this approach is used, the TT is often compromised as it conforms to the ST even when the

two languages are not in related family groups, just like English and Kiswahili. This leads to several anomalies, including awkwardness, unnaturalness or loss of meaning. Awkwardness and unnaturalness has to do with an expression in the target language whose form and structure reads more like the source language. A good example is “[Group Name] *ana picha mpya*” where, instead of moving the subject to a different location within the TT sentence, the translator forces it to remain at the same location as the ST sentence.

#### **4.4.1 Literal or Word-for-word Translation**

This is a translation strategy where each word in one language is translated exactly into its counterpart in another language (Lorraine Caplan, 2018). It is also called direct translation, where the translator interprets and translates the translation units superficially. This happens, for example, when translating standalone terms such as “wave” and “message”, which are meant to be translated as verbs, but the translator gives them a superficial translation, without considering the context, and therefore translating them as nouns. As expected, this leads to a breakdown in communication since the resulting product is a mistranslation. The remedy for this would have been using dynamic equivalence where an effort is made to convey the ST message in the TT as naturally as possible, without rigidly sticking to the ST structure. It is important to note that word-for-word translation does not always yield to a faulty translation, as indeed at some point every translator has to apply this approach.

Facebook does not provide context description to aid the translator correctly interpret a text in its immediate environment within the UI. This account for some of the literal translations where a text is translated in its value. A description appended to the source text would have easily solved

a major issue such as whether a text is supposed to be translated as a verb or noun, such as in the case of “wave” above.

#### **4.4.2 Oblique Translation**

The second approach is oblique translation, which is used when literal or word-for-word translation is not possible. This is an idiomatic translation where the focus is on transmitting the ST meaning into the TT as naturally and easily as possible, without sticking to the source language structure. Molina, L. & Hurtado Albir, A. (2002) observe that there are several processes involved in this approach, such as modulation - which is a shift in point of view, transposition – which is a shift in word class, and adaptation – a shift in cultural environment, such as replacing a food type in the ST with a local equivalent in the TT for better understanding by the target audience. Oblique translation has been used in the translation of Facebook UI content from English to Kiswahili as observed in the collected data. For example, “Go to free” has been rendered as “*Nenda kwa bila malipo*”, where the word “free” has been replaced by the phrase “*bila malipo*” (unpaid/without payment) because the direct equivalent “bure” would cause a mistranslation that implies the meaning of “useless”. Another example is “Mutual friends” which has been (wrongly) translated as “*Marafiki mnaojuana*” (which should have been “*marafiki wenza*”) because Kiswahili does not have the concept and direct equivalent for “mutual”.

These translation approaches are not inherently erroneous, rather, it is their wrong application that leads to the errors encountered in this study. Therefore, the study does not lay a generalized blame on these approaches, but only finds fault on their wrong use.

#### **4.5 Communicative Implications of the Observed Translation Errors**

As mentioned in the Scope and Limitations of this study, we did not set out to measure the severity of the errors. So our analysis includes both minor and major errors, since they have a bearing on the passage of the intended message in the Facebook UI, and affects the usage of Facebook by Swahili users in general. Admittedly, there were errors that did not hinder communication or Facebook usage, especially grammatical errors. This is because the intended gist was captured, such as in mistranslating a preposition or pronoun. While Kiswahili grammatical rules were flouted, communication or Facebook usage was not hindered by these errors. However, there were several translation errors that left the user confused or stuck as to what action to take, like in the case of translating the verb “wave” as a noun to mean “sea wave” or the verb “message” as a noun. Therefore, the answer to the question as to whether these errors hindered communication and the usage of Facebook through its UI cannot be a resounding yes or no from this study, rather both yes and no, based on each error and its immediate context.

#### **4.5 Conclusion**

This study finds that the errors resulting from crowdsourced Kiswahili translations for Facebook UI fall under 3 main categories namely accuracy, stylistic and grammar errors, with the majority of the errors being stylistic. It is safe to conclude that these errors are avoidable, if only the translators were keener or better prepared for the translation exercise, even within the crowdsourcing process. There wasn't a single case of a concept too difficult to understand or translate, since Facebook UI content is quite easy and straightforward in nature. While these errors do not necessary occur due to the use of crowdsourcing, it was observed that crowdsourcing process did not provide the prerequisite environment for high quality translation.

This environment, found in the standard outsourcing approach includes translator selection, resource preparation and post-translation quality control such as editing and proofreading. This post-translation quality control process would arrest sloppy errors as observed in this study, so the fact that they are glaringly present could indicate that little or no LQA was performed after the crowdsourced translation exercise.

## CHAPTER FIVE

### 5.0 Summary, Conclusion and Recommendations

In this study, we set out to assess the translation errors in crowdsourced translations in Kiswahili Facebook UI content. The study identified crowdsourcing as a translation model that involves two main players—the content owner and the crowd. Depending on the inherent interests, the content owner or the crowd can initiate crowdsourcing. Where a crowd aims to avail content to others in a specific language, then that crowd some interest in the subject domain. Where the crowd bears some interest in the translation of content, the participants are usually intrinsically motivated to translate.

Mostly, this form of crowdsourced content is used for not-for-profit purposes. On the other hand, the content owner like Facebook Inc. can initiate the call for participation and provide a platform like Facebook Translate through which the crowd translates the content. In this form of crowdsourced translations, the crowd does not necessarily bear interest in the subject domain—the crowd is simply called to participation based on a for-profit organization’s interest. In the translation of the English Facebook User Interface into Kiswahili, it is appropriate to argue that Facebook Inc. was more interested in having a Kiswahili User Interface than the crowd was in having a Facebook interface in a language that they can identify with.

These deductions highlight some arguments that were reviewed earlier herein. Dolmaya (2011) points out that the line between ethical and unethical crowdsourced translation is drawn by the motivation behind such translation. Dolmaya’s claims initiate interplay among product-driven, cause-driven, and outsourcing-driven crowdsourced translations. For the first two (product-driven and cause-driven) types of crowdsourced translations, it is arguable that crowdsourcing is

crucial and should, in fact, be encouraged. Product-driven translation aims to localize certain freely available documents or open-source software and cause-driven translation is employed in translation of content for humanitarian - missions mainly by not-for-profit organizations.

Conversely, outsourcing-driven translation is used by for-profit organizations to engage the crowd in translating content without any tangible remuneration. This tendency of by for-profit organizations to rely on free crowdsourced content is the basis for the much publicized clashes between these organizations and professional translators. Ultimately, in what professional translators refer to as unethical practice by for-profit organizations lies a pool of translation errors that can be traced back to the sheer wanton utilization of the crowd without any inclusion criteria.

Digital companies prefer crowdsourcing due to its cheap availability and its ability to have more words translated within a shorter time, albeit at the expense of quality. Among the critics of crowdsourcing are linguistic scholars and professional translators, due to ethical concerns around the practice, as well as its tendency to result to low quality translations. There are domains that favor crowdsourcing, especially for non-technical content and that which would not have serious legal implications in the event of mistranslation. Crowdsourcing for Kiswahili is a new phenomenon both within the Kiswahili speaking world and outside, with the pioneers behind Facebook, TWB and Ushahidi.

### **5.1 Error Types Identified in this Study**

This study extracted and identified errors grouped into 3 major categories in the Kiswahili Facebook UI namely accuracy, grammar, and stylistic errors. Stylistic errors were the majority. These type of errors occur when a translation employs a wrong form, art or method in translation

due to literalness and wrong tone. These error types followed by accuracy errors in frequency. Accuracy errors occur when the target text does not respect and convey the source text meaning as much as possible, resulting to an altered meaning. Finally, grammatical errors were also observed, which occurred due to the violation of the grammatical rules of the target language, which is Kiswahili in this case. All these error types may have occurred due to the lack of a stringent LQA process in the crowdsourcing method.

When crowdsourcing was compared to standard outsourcing in terms of quality control, it was observed that has a more stringent LQA process than crowdsourcing. Translator vetting is one of the steps that crowdsourcing does not apply to a large extend, which is a prerequisite for good quality, besides other pre and post-translation steps. The lack of these steps may account for the avoidable errors observed, which lowered the quality of the translation produced. Had LQA been performed, the frequency and level of sloppiness of these errors would not be in the magnitude observed.

## **5.2 Apparent Translation Strategies Used by Crowdsourced Translators on Kiswahili Facebook UI**

It was observed that the errors identified and discussed in this study were as a result of using two translation strategies namely literal translation/word-for-word translation and oblique translation. Literal translation or word-for-word translation is a translation strategy where each word in one language is translated exactly into its counterpart in another language. This is also commonly known as direct translation. Oblique translation is an idiomatic translation where the focus is on transmitting the ST meaning into the TT as naturally and easily as possible, without sticking to

the source language structure. These strategies are not inherently wrong, but their wrong application in specific translation contexts led to these errors.

### **5.3 The Implications of the Translation Errors on Communication and Usage of Facebook UI**

It was observed that these errors affected communication and Facebook UI usage to some extent. There were some errors that severely obscured meaning, such as when the user is instructed to take an action in the ST, but the same message is not conveyed in the TT. These are both for one-word expressions such as translating the verb “Wave” as “*Mawimbi*”, a noun which literally means (sound/sea) waves; or phrasal expressions such as “uncover image” which was rendered as “*Picha ya ufichuzi*” instead of “*Fichua picha*”.

On the other hand, there were minor errors which did not communicate the gist of the ST, and therefore did not hinder communication and usage of Facebook UI. This was especially so with grammatical rules, since it affected mainly pronouns, prepositions and tense. The user may notice that these were errors, but would still understand the intended message by their own interpretation.

### **5.4 Recommendations**

For scholars who may be interested in expanding this study, it is recommended that they attempt to assess the frequency and severity of the errors that may hinder to a large extent the intended communication or usage of the Kiswahili Facebook UI. This was not determined in this study, as it was outside its scope. An assessment of the frequency and severity of the errors in

crowdsourced translations may offer an objective conclusion on the viability of using the crowdsourcing method of translation in any field.

It would also be interesting to investigate the viability of a crowdsourcing-outsourcing hybrid approach where the beneficial aspects of each approach are merged to form a more efficient solution. As observed in this study, crowdsourcing has the advantage of translating a large word count in a short time by taking leveraging on the power of the crowd. This comes with the disadvantage of low quality translations, as observed. On the other hand, standard outsourcing has better quality translation output due to its stringent LQA process, but at a considerably lower speed of translated words. Therefore, a hybrid on both outsourcing and crowdsourcing that leverages on the strengths from both approaches may be an interesting phenomenon to study.

## Bibliography

- Anastasiou, D., & Gupta, R. (2011). Comparison of crowdsourcing translation with Machine Translation. *Journal of Information Science*, 37(6), 637-659. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.831.3132&rep=rep1&type=pdf>
- Anderson, P. (2012). *Web 2.0 and Beyond: Principles and Technologies*. Chapman and Hall/CRC .
- Brabham, D. C. (2009). Crowdsourcing the Public Participation Process for Planning Projects. *Planning Theory*, 8(3), 242-262.
- Calvo, E. (2011). Translation and/or Translator Skills as Organising Principles for Curriculum Development Practice. *The Journal of Specialised Translation*, 16, 5-25.
- Creswell, J. W. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). California: Sage publications.
- De Wille, M. T. (2017). A Mixed Methods Investigation on Perceived Quality in Crowdsourced Localisation. *Doctoral Dissertation. University of Limerick*.
- Dolmaya, J. M. (2011). The Ethics of Crowdsourcing. *Linguistica Antverpiensia, New Series– Themes in Translation Studies*, (10), 98-110. Retrieved from <https://lans.ua.ac.be/index.php/LANS-TTS/article/viewFile/279/177>.
- Estellés-Arolas, E., & González-Ladrón-De-Guevara, F. (2012). Towards an Integrated Crowdsourcing Definition. *Journal of Information science*, 38(2), 189-200. DOI: 10.1177/0165551500000000.
- European Commission . (2012). Crowdsourcing Translation. *Studies on translation and multilingualism*, Retrieved from <http://www.termcoord.eu/wp-content/uploads/2013/08/Crowdsourcing-translation.pdf>.
- Facebook.Inc. (2019). *How do I get the Translate Facebook app?* Retrieved from Facebook Help Centre: [https://web.facebook.com/help/translations/867502359935199?\\_rdc=1&\\_rdr](https://web.facebook.com/help/translations/867502359935199?_rdc=1&_rdr)

- Flanagan, M. (2016). Cause for Concern? Attitudes Towards Translation Crowdsourcing in Professional Translators' Blogs. *The Journal of Specialised Translation*, 25(1), 149-173. Retrieved from [https://www.jostrans.org/issue25/art\\_flanagan.pdf](https://www.jostrans.org/issue25/art_flanagan.pdf).
- Gass, S. M., & Selinker, L. (2008). *Second Language Acquisition: An Introductory Course 3rd Edition*. Newyork and London: Routledge.
- Guzman, P., Pino, Y. M., Husa, D., & Ayan, N. F. (2018, September 11). *Expanding automatic machine translation to more languages*. Retrieved from Facebook Code Website: <https://code.fb.com/ml-applications/expanding-automatic-machine-translation-to-more-languages/>
- HBS Digital Initiative (2015, October 28). *Leveraging the Collective Intelligence: Facebook Crowdsourced Translations*. Retrieved on September 25<sup>th</sup> 2019 at <https://digital.hbs.edu/platform-digit/submission/facebook-crowd-sourced-translation/>
- Kaminskienė, L., & Kavaliauskienė, G. (2012). Competences in Translation and Interpreting. *Studies about Language*, 20, 138-143.
- Nida, E. A. (1995). *Dynamic equivalence in translating.* " *An Encyclopaedia of Translation: Chinese-English English-Chinese [C/Z]*. Hong Kong: The Chinese University Press.
- O'Hagan, M. (2016). Massively Open Translation: Unpacking the Relationship between Technology and Translation in the 21st Century. *International Journal of Communication (19328036)*, 10.
- Pascoal, S. C., Furtado, M., & Chorão, G. (2017). CROWDSOURCING TRANSLATION AND THE THREATS/CHALLENGES TO THE PROFESSION: A PORTUGUESE TRANSLATOR'S SURVEY." CURRENT TRENDS IN TRANSLATION TEACHING AND LEARNING E 4. *Center for Intercultural Studies, ISCAP, P. Porto*, 229–263.
- Pinto, L. F., & dos Santos, C. D. (2018). Motivations of Crowdsourcing Contributors. *Innovation & Management Review*, 15(1), 58-72.
- Queirós, A., Faria, D., & Almeida, F. (2017). Strengths and Limitations of Qualitative and Quantitative Research Methods. *European Journal of Education Studies*, 3(9), 369-387.
- Ray, R., & Kelly, N. (2011). Crowdsourced Translation. *Common Sense Advisory*.

Ridder, H.-G. (2017). The Theory Contribution of Case Study Research Designs. *Business Research*, 10(2), 281-305. Retrieved from [https://www.econstor.eu/bitstream/10419/177270/1/10.1007\\_s40685-017-0045-z.pdf](https://www.econstor.eu/bitstream/10419/177270/1/10.1007_s40685-017-0045-z.pdf).

Sun, Y. (2015). *Translation and academic journals: the evolving landscape of scholarly publishing*. Springer.

Washbourne, K. (2012). Translation Style Guides in Translator Training: Considerations for Task Design. *The Journal of Specialised Translation*, 17, 2-16.

Wolverton, T. (2018, July 23). *Google CEO Sundar Pichai revealed a jaw-dropping fact about its translation app that shows how much money is still sitting on the table*. Retrieved from A Business Insider Website: <https://www.businessinsider.com/sundar-pichai-google-translate-143-billion-words-daily-2018-7?IR=T>

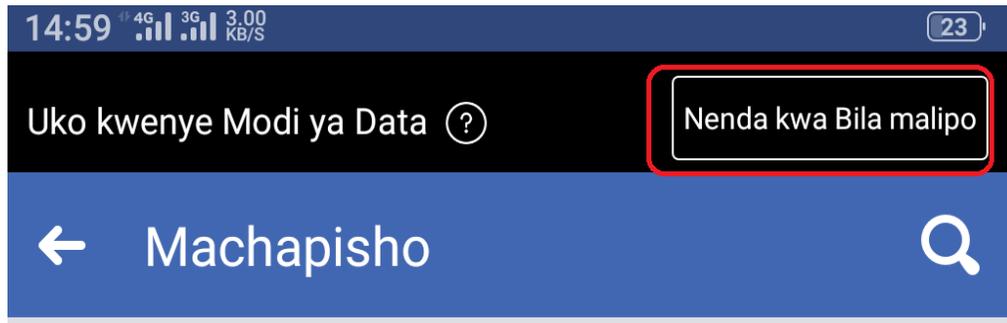
Zaidan, F. O., & Callison-Burch, C. (2011). Crowdsourcing translation: Professional quality from non-professionals. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1* (pp. 1220-1229). Baltimore, USA: Association for Computational Linguistics.

## Appendices

The following are screenshots sampling some of the errors analysed in this study. Only a few screenshots have been included because it would have been too bulky to include all the analyzed errors



Screenshot 1



Screenshot 2



Screenshot 3



Lilly Kavuttih alikualika ujiunge kwenye **kundi kilichofungwa** cha 52-week Savings Challenge Kenya. Dakika 31 zilizopita

Screenshot 4

## Kwa ufupi



Grace Ambundo  
Km 1 · Dakika 11

  
Amepunga  
mkono

## Karibu



Aggrey Samuel  
Km 1 · Saa 1

  
Mawimbi



Dave Koome Mutea  
km 9 · Dakika 57

  
Mawimbi

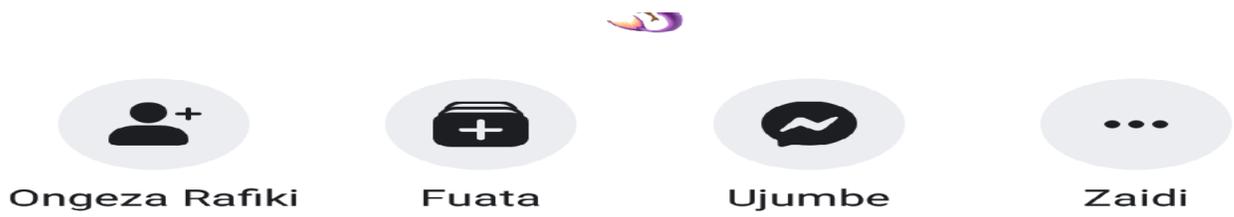


Brielle W. Ashley  
km 8 · Saa 10

  
Mawimbi

Ona Zaidi

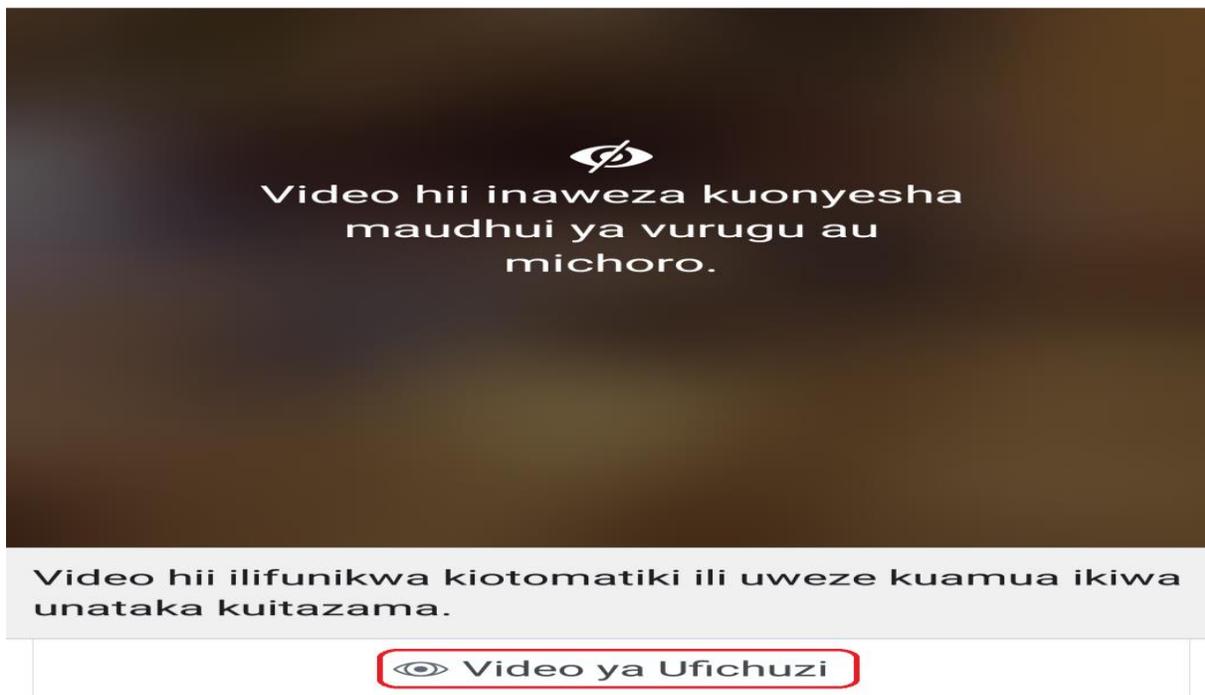
Screenshot 5



 **Amejiunga Aprili 2007**

 **Anafuatwa na watu 234**

Screenshot 6



Screenshot 7



**MemoQ Users** ana chapisho jipya kutoka kwenye **Kevin Lossner**.  
Dakika 24 zilizopita

Screenshot 8



**Futbol 254** ana picha mpya kutoka kwa **Francis Muriithi**.  
Dakika 31 zilizopita

Screenshot 9



**Abraham Okumba** amependezwa na toa maoni yako: "Abraham Okumba That escalated quickly".  
Dakika 13 zilizopita

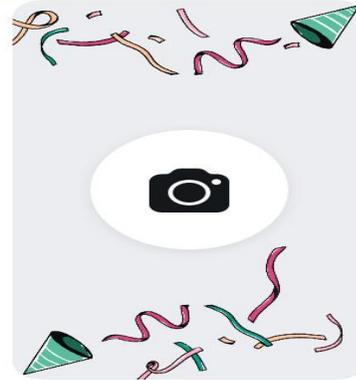
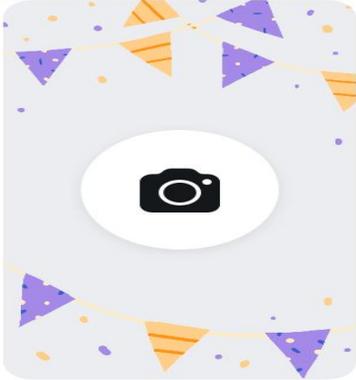
Screenshot 10



**Michubu Jonah Karwash**  
Leo, 11 Mei



**Binafsisha kadi kwa hadithi yake ya siku ya kuzaliwa**



Screenshot 11

← **Maudhui Hayapatikani**

Ukurasa ulioomba hauwezi kuonyeshwa kwa sasa. Inaweza kutopatikana kwa muda, kiungo ulichobofya huenda kimekatika au kimekwisha muda, au huenda usiwe na idhini ya kutazama ukurasa huu.

**Rudi hadi nyumbani**

Screenshot 12



Lilly Kavuttih ●



Charles Kyeti ●

UJUMBE

UJUMBE

Screenshot 13



**Wisdom Stanley** 😊 anahisi najisikia vibaya pamoja na **Stanley Gachago**.

Dakika 17 • 👤

Screenshot 14