# INTEGRATING MACHINE TRANSLATION WITH INSTITUTIONAL REPOSITORIES: A CASE STUDY OF UNIVERSITY OF NAIROBI

**JULIUS NYABUTI MUGOYA**

**A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF MASTER OF ARTS IN TRANSLATION, UNIVERSITY OF NAIROBI**

**NOVEMBER 2018**

## DECLARATION

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

**Signature _____ Date _____**

**JULIUS NYABUTI MUGOYA**

**Y61/8477/2017**

This dissertation has been written and submitted for examination with our approval as the University supervisors.

**Signature_____ Date_____**

**DR. GIDEON MARETE**

**UNIVERSITY OF NAIROBI**

**Signature_____ Date_____**

**MR. JOHN PAUL WARAMBO**

**UNIVERSITY OF NAIROBI**

## ABSTRACT

The general aim of this research is to study the viability and challenges that would face the use of machine translation for digital collections in the University of Nairobi Institutional Repository. The premise of the study is that as universities are churning out their output through IRs, these should be available in the language of the user. Since university funding is getting tighter, open source software should be considered for some tasks such as translation. The specific objectives of the study include: finding out the suitability and convenience of using machine translation in translating IR resources in the UoN repository; evaluating the most suitable open source machine translation system for translating IR resources from English to Kiswahili; identifying the benefits and/or constraints encountered when doing so, and; generally identifying the future role of machine translation in digital information dissemination. The exploratory nature of this study makes it be conducted using a qualitative approach. Purposive sampling is used to choose the texts for study. To examine the perception of quality of machine translation, two open source MT software are used and their quality of translation compared against each other by highlighting their errors. The data emanating from the study is analyzed to identify the translation strategies used by the MT software. It is hoped that the results of the research will galvanize efforts towards offering translation assistance to make IR resources more meaningful to all researchers. Perhaps by having a way through which a user can switch between different language versions one could be able to translate the resources into the language most desirable and therefore eliminate the processes involved in copying and pasting into a translation app or even trying to decipher the meaning.

## DEDICATION

To the Almighty God for His favours that have brought me this far.

To my wife, Teresa, and children, Godwin, Teddy, and Hope, for always being there for me and giving me moral support even when the going got tough.

To my mother for the encouragement to soldier on.

To all my other family members for believing that I am not yet done!

## ACKNOWLEDGEMENTS

I would like to convey my heartfelt indebtedness to my supervisors, Dr. Gideon Marete and Mr. John Paul Warambo, for their generous, humane and patient support, invaluable advice on the successive drafts of the thesis and for always trusting in its finalization. They frequently gave me a lot of expert advice through formal and informal discussions on my thesis writing. Communicating with them always sired forth new ideas and greatly inspired my research.

My special thanks also go to the accommodating and considerate staff and lecturers of the Centre of Translation and Interpretation led by the director, Prof. Jayne Mutiga. Without your constant monitoring, prodding and encouragement, this dissertation would not have been possible. You made the whole course worthwhile. My fellow students encouraged me to continue through their intellectual stimulation and moral support. Every lesson and discussion session was something to look forward to. I can only say "Thank you!"

Finally, infinite gratitude goes to my family without whose support I would not have even attempted to take this course.

## TABLE OF CONTENTS

DECLARATION ..... i
ABSTRACT ..... ii
DEDICATION ..... iii
ACKNOWLEDGEMENTS ..... iv
TABLE OF CONTENTS ..... v
APPENDICES ..... viii
LIST OF FIGURES ..... ix
LIST OF TABLES ..... x
ACRONYMS AND ABBREVIATIONS ..... xi

**CHAPTER ONE**
INTRODUCTION ..... 1
1.0 Introduction ..... 1
1.1 Background to the Study ..... 1
1.1.1 Operational Definitions ..... 7
1.2 Statement of the Problem ..... 9
1.3 Research Objectives ..... 10
1.4 Research Questions ..... 11
1.5 Justification for the Study ..... 11
1.6 Scope and Limitations ..... 12
1.7 Literature Review ..... 12
1.7.1 Institutional Repositories ..... 12
1.7.1.1 University of Nairobi Institutional Repository ..... 13
1.7.2 Machine Translation ..... 14
1.7.3 Types of Machine Translation Systems ..... 18
1.7.4 Machine Translation Systems with Swahili Language Combination ..... 21
1.8 Theoretical Framework ..... 23
1.9 Research Methodology ..... 25

1.9.1 Data Collection ..... 26
1.9.2 Data Analysis ..... 27
1.10 Conclusion ..... 29

**CHAPTER TWO**

OVERVIEW OF GOOGLE TRANSLATE AND BING ..... 30
2.0 Introduction ..... 30
2.1 Google Translate ..... 30
2.2 Microsoft Bing Translator (BT) ..... 31
2.3 Classification of Texts ..... 31
2.4 Classification of Data from the Translation Systems ..... 33
2.5 Assessment Criteria of the Translation Systems ..... 34
2.6 Distribution of Translation Errors ..... 34
2.7 Translation Strategies ..... 35
2.8 Conclusion ..... 35

**CHAPTER THREE**

DATA PRESENTATION ..... 36
3.0 Introduction ..... 36
3.1 Form of Presentation ..... 36
3.2 Results of the Study ..... 38
3.3 Analysis of Errors ..... 40
3.4 Ranking of MTs by Human Translators ..... 43
3.5 Conclusion ..... 44

**CHAPTER FOUR**

DISCUSSION OF FINDINGS ..... 45
4.0 Introduction ..... 45
4.1 Suitable MT for Translating IRs ..... 45
4.2 Errors Encountered ..... 45
4.3 Consequences of Errors ..... 47

4.4 Effects of Translation Strategies ..... 48
4.5 Conclusion ..... 48

**CHAPTER FIVE**
SUMMARY, CONCLUSION AND RECOMMENDATIONS ..... 49
5.0 Introduction ..... 49
5.1 Summary of Findings ..... 49
5.2 Conclusion ..... 51
5.3 Recommendations ..... 54

**REFERENCES** ..... 56

## APPENDICES

Appendix A: Cultural Text 1

Appendix B: Cultural Text 2

Appendix C: Cultural Text 3

Appendix D: Scientific Text 1

Appendix E: Scientific Text 2

Appendix F: Scientific Text 3

Appendix G: Quality Ranking Instrument

Appendix H: MT Ranking Input Sheet

## LIST OF FIGURES

Figure 1. 1: Bernard Vauquois' MT triangle ..... 16
Figure 1. 2: Examples of MT systems ..... 18
Figure 1. 3: Van Doorslaer's translation strategies ..... 24
Figure 5. 1: Knowledge translation scenario using MT ..... 52

## LIST OF TABLES

Table 1.1: Internet content users by language ..... 2
Table 2.1: The randomization of material..... 33
Table 3.1: Applied transtaion strategies in cultural texts..... 38
Table 3.2: Applied transtaion strategies in scientific texts ..... 39
Table 3.3: Distribution of errors that are critical for comprehension of cultural texts ..... 40
Table 3.4: Distribution of errors that are critical for comprehension of scientific texts..... 40
Table 3.5: Overall translation quality rakings of MT by HT ..... 41
Table 3.6: The MT rankings ..... 44

## ACRONYMS AND ABBREVIATIONS

| | |
|---|---|
| CAT | Computer Aid Translator |
| BT | Bing Translator |
| GT | Google Translate |
| HMT | Hybrid Machine Translation |
| HT | Human Translator |
| IR | Institutional Repository |
| LIS | Library and Information Science |
| MT | Machine Translation |
| NER | Named Entity Recognition |
| NLP | Natural Language Processing |
| OA | Open Access |
| RBMT | Rule-Based Machine Translation |
| RTE | Round-Trip Evaluation |
| SL | Source Language |
| SMT | Statistical Machine Translation |
| ST | Source Text |
| TL | Target Language |
| TT | Target Text |
| UoN | University of Nairobi |

# CHAPTER ONE

# INTRODUCTION

## 1.0 Introduction

Translation involves at least two languages (Toury 1995). Therefore this study is examining the results of machine translation of selected abstracts from English to Kiswahili of theses deposited in the University of Nairobi Insitutional Repository. In this chapter, a presentation of information on the background to the study, statement of the problem, the definition of key terms used in this work, the research objectives, questions, the significance of the study, scope and limitations of the study, the theoretical framework, literature review, and the methodology is made.

## 1.1 Background to the Study

The world has witnessed the advancement of online information systems which have also been complemented by the growth of information on the internet regarding size, reach and content. These are also accompanied by diverse user needs and languages in which this information is availed. In turn, these have created many new opportunities and problems in information access and retrieval. Natural Language Processing (NLP) applications and tools such as Machine Translation (MT) and multilingual thesauri are examples of some of the tools that are needed for ensuring access to this information. In the IFLA Trend Report of 2013, a document of the International Federation of Library Associations and Institutions (IFLA), machine translation is mentioned as being among the five critical high-level technologies which are revolutionalising the universal information financial resources. The market for MT is expected to grow and reach an annual turnover of US$983.3 million by 2022 (Hexa Research, 2015). The professional translation industry alone was estimated to offer services worth US$37.19 billion worldwide by

2014 (De Palma, Hegde, Pielmeier, & Stewart, 2014). Therefore the use of MT will push it to astronomical levels.

Globalization and translation complement each other. Translation enables global communication by ensuring that the message from the ST is received by the TT as intended (Cronin, 2013). The translator, whether human or machine, plays the role of a link in the translation process. Although English is heavily used in the internet (Flammia & Saunders, 2007), online documents are now multilingual as seen in table 1 below:

**Table 1.1: Internet content users by language** *(Based on data from Internet World Stats, 2018)*

| Top Ten Languages Used in the Web - December 31, 2017 ( Number of Internet Users by Language ) | | | | | |
|---|---|---|---|---|---|
| **TOP TEN LANGUAGES IN THE INTERNET** | **World Population for this Language (2018 Estimate)** | **Internet Users by Language** | **Internet Penetration (% Population)** | **Internet Users Growth (2000 - 2018)** | **Internet Users % of World (Participation)** |
| **English** | **1,462,008,909** | **1,055,272,930** | **72.2 %** | **649.7 %** | **25.4 %** |
| **Chinese** | **1,452,593,223** | **804,634,814** | **55.4 %** | **2,390.9 %** | **19.3 %** |
| **Spanish** | **515,759,912** | **337,892,295** | **65.5 %** | **1,758.5 %** | **8.1 %** |
| **Arabic** | **435,636,462** | **219,041,264** | **50.3 %** | **8,616.0 %** | **5.3 %** |
| **Portuguese** | **286,455,543** | **169,157,589** | **59.1 %** | **2,132.8 %** | **4.1 %** |
| **Indonesian / Malaysian** | **299,271,514** | **168,755,091** | **56.4 %** | **2,845.1 %** | **4.1 %** |
| **French** | **412,394,497** | **134,088,952** | **32.5 %** | **1,017.6 %** | **3.2 %** |
| **Japanese** | **127,185,332** | **118,626,672** | **93.3 %** | **152.0 %** | **2.9 %** |
| **Russian** | **143,964,709** | **109,552,842** | **76.1 %** | **3,434.0 %** | **2.6 %** |
| **German** | **96,820,909** | **92,099,951** | **95.1 %** | **234.7 %** | **2.2 %** |
| **TOP 10 LANGUAGES** | **5,135,270,101** | **3,209,122,400** | **62.5 %** | **1,091.9 %** | **77.2 %** |
| Rest of the Languages | **2,499,488,327** | **950,318,284** | **38.0 %** | **935.8 %** | **22.8 %** |
| **WORLD TOTAL** | **7,634,758,428** | **4,159,440,684** | **54.5 %** | **1,052.2 %** | **100.0 %** |

Out of the 4.1 million users who accessed the internet in 2017, it is only 1 million who accessed it using English. The other languages are equally important. Due to the multilingualism and multiculturalism in today's international information society, the requirement for translations is ever increasing. This phenomenal growth has necessitated the development of free web-based systems based on automatic MT, such as Google Translate and Microsoft Bing Translator, for those users who do not consider professional services necessary. Free software is sometimes called open-source software.

Universities play a pivotal role in the production and execution of new ideas which have come through study and inquiry. To do research in most fields of knowledge, the previous relevant scholarly writings are generally consulted. These are often made available through journals and in conferences (Harnad, 2003). There is the other scholarly research which is very important but is not always published in journals– theses and dissertations. These have traditionally been kept in libraries. However, with the application of digital technologies, they can now be disseminated universally in digital form without regard to space and time.

To expand the manifestation of a university's research work, universities now are involved in implementing institutional repository services. Lynch (2003) defines an Institutional Repository (IR) as "a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members". An IR is also known as a digital repository or digital commons. IRs generally are maintained by university academic libraries. In an IR system, research output files are uploaded, the necessary bibliographic metadata identified and made available, and then linked to internet

search engines, or to scholarly search engines such as Google Scholar, so that they are made available to internet users. An IR enables users to locate, comprehend and build upon previous work and therefore provide a basis for the following research. Studies have shown that IRs lead to higher citations of articles (Hitchcock, 2013, Piwowar & Vision, 2013). Citations are highly valued by researchers in the scholarly community (Kratz & Strasser, 2015). Furthermore, this improves the webometric ranking of the university. Webometrics is the study of internet-based scholarly output using informetrics or other measurable means (Bjorneborn and lngwersen, 2004). Webometric ranking motivates both the university and scholars to have a web presence that reflects their activities accurately and general performance (www.webometrics.info). As at the end of July 2018, University of Nairobi was ranked number 853 in the world, number 8 in Africa, and in 1st position within Kenya (Webometrics Ranking, 2018).

The University of Nairobi has established a digital repository as the system through which to provide archiving and displaying of its scholarly output. The whole University community is obliged to present their academic output to the repository. The UoN repository includes materials such consists of conference proceedings, students' Ph.D. theses, Masters' dissertations and other multimedia. The University of Nairobi IR is available at this address: http://erepository.uonbi.ac.ke. It is supervised by the University Librarian but managed by the Repository Administrator (www.uonbi.ac.ke).

Going through the contents of the repository, it is noted that most of the research output is in English. In most cases, theses and dissertations are studies carried out about local problems. The irony is that the solutions they offer never get to reach where they are most needed because they

are written in a “foreign” language. In a study conducted by Translators without Borders (TWB) on humanitarian related issues, the majority of respondents (89 percent) indicated that although they preferred receiving information in their mother tongue, this was not available. For instance, during the Ebola outbreak in 2015 in Guinea, Liberia and Sierra Leone, Ebola-related materials were written in English and French, yet the region has over 90 local languages (TWB, 2017). This created barriers to the transfer of relevant knowledge as illiterate people and minority language speakers were left out. Language services are essential for bilateral relationships, especially where the working languages of trade partners are different. Translation opens up further unexploited avenues where the creation of national wealth may be exploited such as when products and services needing technical documentation are being exported from one country to another.

Google Scholar conducted a research in 2014 using 16 languages. The research revealed that of the 75,513 documents on biodiversity conservation published, only 36% were in other languages apart from English (Amano et al., 2016). To increase noticeability of such intellectual works and enhance collaboration with the global research community, it would be ideal to translate it into the desired language of the user. Translation will enable information to be understood and used without minding what the original language was. Machine translation is at the intersection between translation studies and computer science (Giammarresi and Lapalme, 2016). The goal of computer science is to enable computing technology devices to operate and to communicate with humans and with other machines. The discipline of MT is where computer scientists, translators, and linguists come together in a bid to achieve this. As noted by Nzomo (2015) there is a need

for system designers of IR and information scientists to enhance and promote information access and literacy in multilingual environment.

Africa is a continent which has a rich linguistic diversity having an estimated 1500-2000 languages from six language families. The Swahili language (or Kiswahili) is a widely spoken Bantu language. As much as one hundred million people throughout a wide area of East and Central Africa speak it (Ethnologue, 2018). Kiswahili is a lingua franca of long standing which enjoys a reasonably well-accepted status as a medium of communication. It is also aired in international radio broadcasts such as Radio China, Voice of America, Deutsche Welle and the British Broadcasting Corporation among others. Knowledge of Kiswahili enhances the credibility of researchers interested in East Africa. Kiswahili is also taught in universities in Japan, Europe, Mexico, and the USA. Therefore, with content in Kiswahili, a better and deeper understanding of subject matter takes place. Since a lot of academic content already exists in English, using MT is the most practical method of getting this content into Kiswahili quickly.

There are also other initiatives which are promoting the need for having content in Kiswahili. The Government of Kenya passed a regulation in 2015 that requires local content of at least 60% on local media. This means that at some point, some of the content to be aired has to be translated into local languages, Kiswahili being among them (Communications Authority of Kenya, 2015). Also in 2015, the East African Legislative Assembly resolved to make Kiswahili its second official language. The African Union, on the other hand, had earlier on adopted Kiswahili as one of its official languages in 2004. It is also noteworthy that around the same time, Microsoft made software available in Kiswahili (Leclaire, 2004).

### 1.1.1 Operational Definitions

In this research, some terms will be used to refer to specific concepts. Here are the most common terms which will be used and their definitions as used in this study:

*Corpus* – Refers to a collection of texts (Nakayama, Hara & Nishio, 2008; Lenci, Zampolli & Calzolari, 2015).

*Digital library* – This is a gathering of services and information materials that then are made available to users in electronic or digital format.

*Institutional Repositories* – Refers to free, openly available, online preserved researched electronic ouputs of universities.

*Machine Translation* -- This is the technique through which computer software are used to turn a text from one natural language to another using computer devices and without human intervention (Reese, 2015).

*Open Access* -- Refers to the act of making work freely available through the web without any payment, legal or technical barriers.

*Parallel Corpus* – Refers to a collection of texts which are aligned side by side in such a way that one is a translation of the other (Nakayama et al., 2008; Lenci et al., 2015)

*Translation* – The process of rendering the texts or words from one language into another to express the original meaning (Reese, 2015).

*Domain translation* – Translating within a specialized realm or field of knowledge.

*Source language* – Language of the original text.

*Target language* – Language into which the translated text is written.

*Source text* – The original text.

*Target text* – The translated text.

## 1.2 Statement of the Problem

The study posits that provision of a translation facility in the UoN repository will provide an opportunity for greater access and use of the theses and dissertations contained therein. It is therefore imperative to find out which translation tool can be efficient and practical. By so doing, the study will find out if MT can communicate the message effectively given that "computers are becoming very proficient at acquiring skills" (Ford, 2015). In other words, it sets to find out the state of interlingual translation accuracy between English and Kiswahili.

Published literature serves as a foundation for further research in any field of knowledge. It is acknowledged by a significant number of researchers, professors and information professionals that theses and dissertations are of high importance in furthering research and in most cases supervisors urge their research students to refer to these (Vijayakumar, Murthy & Khan, 2007) though they are often underutilized (Horo, S.N. 2006). In instances where they are available through open access, at least they are more used than their paper counterparts hence the reason universities and other bodies are moving to digitize them.

Occasionally the textual content of the research output is multilingual, therefore users who have to access the content in a language they are not proficient in continue to face significant challenges. Many are the times a study or other pertinent information would be available in a researcher's domain but in a language different from that of the researcher. Users want to be able to understand a given text to satisfy their information need. Non-availability of research findings in an understandable language might lead to duplication of effort or their non-use.

Translation would therefore come handy to such users as it would make the context easily discernable (Agosti et al, 2007; Dobreva & Chowdhury, 2010). Throughout human history, human communication has had a translation component, though it was mainly used to transmit religious texts for a long time. According to Roman Jacobson (1959/2004: 139), there are three categories of translation:

1. Intralingual or rewording (translating within the same language, e.g., in subtitling).
2. Interlingual or translation proper (translating from one language into another)
3. Intersemiotic or transmutation (translating verbal signs into non-verbal signs)

This study is concerned with interlingual translation. This study is exploratory to some extent since there are no known previous studies on the integration of machine translation with the institutional repositories of Kenyan universities.

### 1.3 Research Objectives

This study set out with the primary objective of establishing the viability and challenges that would face the use of machine translation for digital collections in the University of Nairobi institutional repository. To meet the primary objective, the study sought to address these objectives:

1) To define the value of machine translation in terms of suitability and convenience in translating IR resources in the UoN repository.
2) To evaluate the most suitable open source MT for translating IR resources in the UoN repository from English to Kiswahili.
3) Identify the benefits and/or constraints encountered in using MT to translate digital resources in the UoN repository.
4) To identify the future role of machine translation in digital information dissemination.

## 1.4 Research Questions

With the above objectives in focus, the research is expected to answer the following questions:

1) Is MT suitable and convenient for translating IR resources in the UoN repository?
2) Which is the most suitable open source MT for translating IR resources from English to Swahili in the UoN repository?
3) What are the benefits and/or constraints encountered when using MT to translate digital resources in the UoN repository?
4) What is the future role of MT in digital information dissemination?

MT evaluation most of the time focuses on functionality and economic utility. MT cannot all translation needs but it is a means to satisfy expressed needs.

## 1.5 Justification for the Study

This research is concerned with the importance of translating the holdings of an IR in order to make it of value to the universe of knowledge. It is hoped that it will be useful to the management of the University of Nairobi and the other universities in Kenya in improving the information services provided through their IRs besides highlighting the changing roles of librarians. This study could also be used as a guideline and reference to scholars and subsequent researchers who have an interest in related undertakings. The study will provide insight that could inform system designers on how to cater to users with diverse linguistic backgrounds and language proficiencies. Web designers could also find the results of the study useful in providing information to support their localization and internationalization efforts.

## 1.6 Scope and Limitations

The University of Nairobi was chosen for study because already it has an established IR. It has also implemented a mandatory deposit of theses in electronic format for all masters and doctoral students. The study intended to exploit only the corpus in the UoN repository. The corpus chosen was for projects undertaken between 2014 and 2017.

The machine translation software to be sampled were limited to Google Translate (GT) and Microsoft Bing Translator (BT). These two are open-source software running on different platforms. Currently, they are the ones offering translations for the English-Swahili language pair. Their choice was informed by the fact that university budgets are getting tighter and should this research be actualized, it would make sense to use software that is freely available and can be improved for the benefit of users.

## 1.7 Literature Review

### 1.7.1 Institutional Repositories

Crow (2002) and Ware (2004) have depicted an institutional repository as a free, multifunctional, cumulative and continuos system that plays a part in scholarly activities by gathering, storing and disseminating the academic content. Generally, institutional repositories perform the following functions:

- They provide a mechanism through which research output is widely disseminated and made available to researchers and the online community in general;
- They create an authentic environment that has credible information;

- They guarantee long-term accessibility of the resources. Therefore repositories also act as libraries.

#### 1.7.1.1 University of Nairobi Institutional Repository

The University of Nairobi (UoN) was set up by an Act of Parliament on 1st July 1970. It was inaugurated on 10 December 1970 and received a charter in 2013. In the charter, the mission of the University is set out as "to provide quality university education and training and to embody the aspirations of the Kenyan people and the global community through creation, preservation, integration, transmission, and utilization of knowledge." (UoN 2018). The UoN has therefore committed to the capturing and preserving of its intellectual output through the signing of the Berlin Declaration on Open Access to Knowledge. A widely accepted definition of Open Access (OA) is work that is made available through the World Wide Web freely without any costs or legal or technical hindrances. However, users are expected to avoid plagiarism by correctly acknowledging and citing those works (Budapest Open Access Initiative, 2002; Drott, 2006). Writers and researchers can make their work OA either by posting to the IR (self-archive) or by publishing in an OA journal (Drott, 2006).

Since its inception up to 2017, UoN has produced over 188,000 graduates. It also boasts of a large number of academic staff. The University offers over 400 academic programmes (UoN information guide, 2017-2018). With such a large number of graduates and academic staff, UoN would be expected to have produced an immense amount of research output. However, UoN had a total of 85,054 items in its IR as of August 2018. Of these, 36,141 are theses and dissertations.

The assumption here is that most of the research output is still in print form. The repository acts as a library since it holds the collection, and as a showcase because the collection is there for the whole world to see the academic exploits of the university. The IR allows organizations, government, and other stakeholders to know the kind of expertise available from the university at a glance.

### 1.7.2 Machine Translation

Machine Translation (MT) is sometimes referred to as natural language processing. It is an automated translation process which makes use of bilingual datasets and other language resources to build language and phrase models that are used to translate text. Simply, it is an automatic translation from one language to another (Abiola O.B & Adetunmbi, O.A., 2015). MT, especially in the area of artificial intelligence, has been of interest to researchers for a long time (Lin, Murakami, Ishida and Tanaka, 2010).

The origin of modern MT can be traced back to 1949. This is when Warren Weaver came up with a model of cryptographic and statistical techniques as used in the field of communication theory to try and apply it in text translation. A demonstration was made in 1954 on the APEXC (All Purpose Electronic X-Ray Computer) machine at Birkbeck College in London of a simple translation of English into French. The need for MT was further fueled by the Cold War as US military sought to use computers to automatically translate Russian documents into English. Similarly, Russia was also researching how to translate from English and French into Russian. Cryptography played a significant role in World War II as it was used to decipher enemy messages. In 1964, the US government formed Automatic Language Processing Advisory Committee (ALPAC) to study the prospects of MT. Its report which came out in 1966 concluded

that “MT was slower, less accurate and twice as expensive as human and that there is no immediate or predictable prospect of useful machine translation” (Hutchins, 1995). This report watered down MT research efforts for some time in the US.

However, MT research continued elsewhere. Machine translation systems were then rule-based (knowledge-based). These systems required a lot of labour. After a domain-specific system, TAUM-METEO, in Canada which translated the weather forecasts from English to French proved successful, research efforts in MT was rekindled from the late 80s. The needs of multilingual communities developed new translation systems in France, Germany, Canada, and Japan. These early systems were based on grammatical rules. However, this changed in the 90s as those based on bodies of texts and examples were developed due to the dynamism of language. Over time, a wide range of MT technologies have evolved, but the main types available include Statistical Machine Translation (SMT), Rule-Based Machine Translation (RBMT), Hybrid Systems which combine RBMT and SMT, and Neural Machine Translation (NMT).

Technology in combination with translation has made it possible for computers to translate automatically and given rise to commercial and free MT software. The online MT systems provide general content translation to data and information which were formerly inaccessible due to language barriers. Babel Fish was the first free translation service to appear on the World Wide Web in 1997 (Yang and Lange 1998, as cited in Aiken, Park and Lindblom, 2010). Others such as Google Translate, Verbalis and Transtext followed.

**Uses and approaches**

Machine translation is mainly used for:

- Assimilation: This is used to enable users understand the general meaning of a text without due regard to grammar.
- Dissemination: This is used when there is a need to publish a text after translation. So the correct grammar is used as much as possible.

Linguistically, MT has three main approaches as described by the Vauquois' triangle in Figure 1 below:

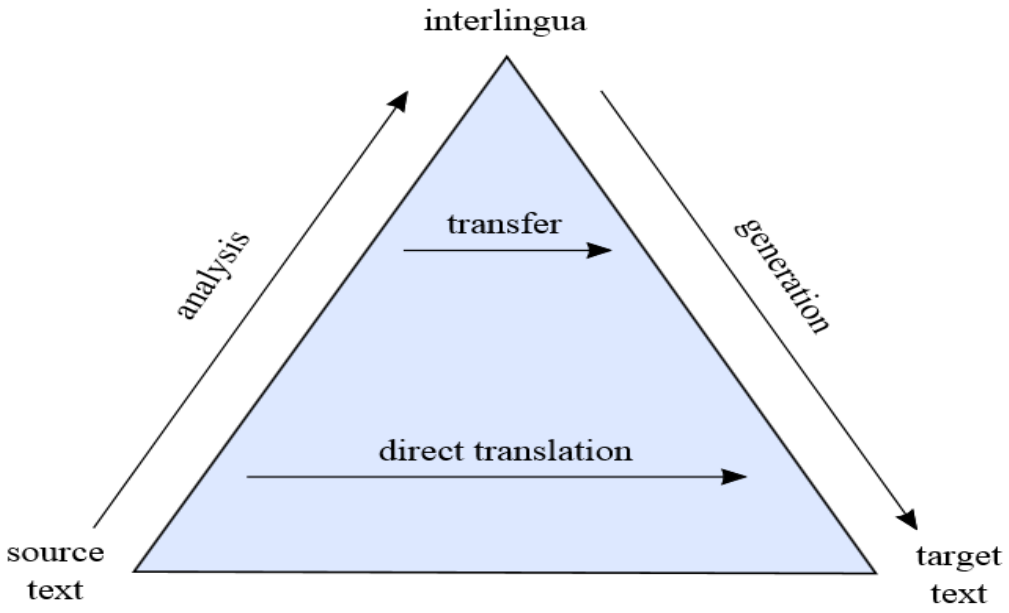


**Figure 1.1: Bernard Vauquois' MT triangle**
**(***Source:http://en.wikipedia.org/wiki/Machine_translation#Interlingual)*

The triangle shows that translation can take place at different linguistic levels using the following approaches:

- Direct Approach: It is the oldest though no longer used in the current MT systems. The system is ideal for a language pair. Specific SL words are connected to particular TL words or expressions.
- Transfer Approach: This approach first converts the SL to any representation. This representation is then remodeled into a similar one for the TL. Finally, the translation is produced from that TL representation. During the analysis stage, some structures could be depending on the dictates of the language involved.
- Interlingua Approach: Like in the transfer approach, the SL is transformed into a transitional stage. The difference here is that the intermediate stage is independent of any language. Then the TL text translation happens directly from the intermediate stage.

The benefits of MT for clients include; reduction of translation costs, reduction of translation time, there is an increase in productivity, and new content can be translated with ease. To translators, MT; provides unique opportunities to be explored, increases productivity, provides higher income due to more work, and there is more consistency in the final product.

### 1.7.3 Types of Machine Translation Systems

In a bid to automate the translation process, various systems have been devised. In general, the process of machine translation involves any of the four types illustrated in Figure 2 below with an additional type, Hybrid MT, which could be a combination of any two.

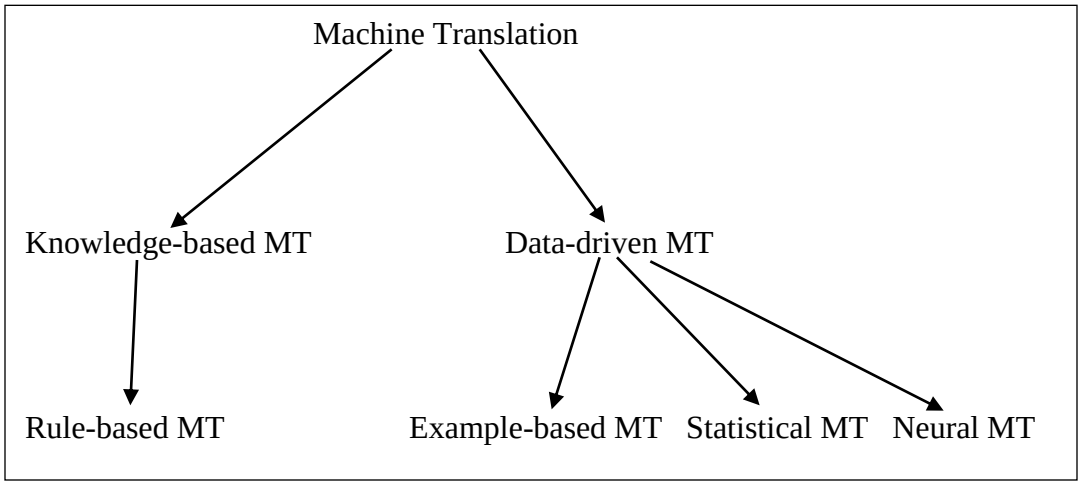


**Figure 1. 2: Examples of MT systems**

***(Source: Researcher)***

**Rule-Based Machine Translation (RBMT)** systems use the transfer approach (interlingua approach) when translating a text from ST to TT hence the reason why sometimes they are called transfer machine translation systems. The rules developed by linguists for this system look at the lexical analysis, morphological information and semantic analysis of both SL and TL among other factors (Wiechetek, 2008). This approach's most significant limitation is many rules and exceptions are required. An example of RBMT is Apertium which is free software. It supports a good number of European languages.

**Statistical Machine Translation (SMT)** use statistical models to train and learn translation of text from a ST to a TT. SMT systems model, train, and decode language. These systems started in the late 1980s due to the increase in computational power and computers were becoming less expensive. The first statistical MT approach was started by IBM in the late 1980s (Brown et al., 1990). LanguageWeaver (now owned by SDL) was the first to be launched in 2002; Moses emerged in 2005; Google and Microsoft changed to SMT in 2007; while Yandex and Baidu started using it in 2011.

SMT technology utilizes statistical models generated by analyzing monolingual and bilingual data. The data is selected using through algorithms by picking the most recurrent words or phrases. To train the MT engine for a specific language pair or domain, a minimum of two million words is required. A notable example of SMT is Google Translate developed by Google Inc. It incorporates a lot of languages spoken in the world. GT is an acceptable way of getting clued in on what the source message is about (Ignacio, 2010; Benjamin, 2014). Each month GT is said to be visited by more than two hundred million users, and their chief scientist asserts that they provide "most of the translation on the planet" (Benjamin, 2014). Google Translate uses the translated documents between multiple languages in the United Nations, European Union and websites from around the world as the foundation for the translation services resulting in a large corpus to rely on (Och, 2012; Koehn, 2007). Translations produced by SBMT are more powerful and accurate than those produced by RBMT systems. Many of the current statistical machine based translation engines require human based parallel text input in order to process translations. This is often the weak point in translation engines since creating parallel texts takes time, and people get quickly bored of the mundane and repetitive task.

**Example-Based Machine Translation (EBMT)** which is also known as memory based translation was first proposed by Makoto Nagao in 1981. The Example-based translation system basically recalls or finds similar examples of language pairs. This system is fed a set of sentences in ST and their corresponding translations of each sentence in the TT with point to point matching. These examples are then used to translate similar types of sentences of ST to the TT later. Basically the idea is that if the previously translated sentence occurs again, the same kind of translation is likely to be correct if used (Nagao, 1984).

**Neural Machine Translation (NMT)** uses neural network models to learn statistical models. NMT is a new breed of corpus-based machine translation. It is trained on vast corpora of pairs of source language segments and their translations. It is similar to the SMT technology but uses an entirely different computational approach: neural networks (Forcada, 2017). This approach's main advantage is that it is trained directly on the ST and TT without having to use the specialized systems typically used in statistical machine learning.

**Hybrid Machine Translation (HMT).** This system's strength is that it takes advantage of the strengths of both statistical and rule-based translation methodologies. This approach is used in a number of ways. In some of the cases, MTs are initially performed by using a rule-based approach then followed up by making adjustments or corrections to the results by using statistical approach. In the other cases, rules are applied to the pre-process state when data is being input and at the post-process production state. This technique has more potential,

versatility, and checks in translation. Systran is an example of a HMT with over 140 language combinations (Karlbom, H. 2016).

### 1.7.4 Machine Translation Systems with Swahili Language Combination

It is estimated that there are approximately 7079 languages that are spoken in the world today (Ethnologue, 2018). Out of these languages, the ones which offer the best results when undergoing machine translation are English, French, Italian, German and Spanish (Benjamin & Radetzky, 2014). The United Nations considers Arabic, Chinese, English, French, Russian and Spanish as its official languages (United Nations, 2016). In the online system, Wikipedia, the English version contained 5.5 million articles as of May 2018 while the Kiswahili version had only 38,000 pages (Wikipedia, 2018). No wonder machine translation between English and French is almost 100% precise. Swahili language lacks NER in the online MT engines hence the engines cannot extract the concepts of people or places in the Swahili sentences. Despite this, several machine translation systems which can translate into Swahili have been developed. They include:

**The Kamusi Project:** Using the same concepts as Wikipedia, The Kamusi Project makes use of online collaborator communities comprising of volunteers, experts and authorities in the Swahili language in order to generate content (The Kamusi Project, 2016). Kamusi crowd-sources translations from the general public and holds them in an on-line database against which searches may be performed.

The Kamusi Project describes itself as being more of a dictionary rather than a machine translation engine (Benjamin, 2016). Indeed the word kamusi is Kiswahili for a dictionary. The Kamusi Project offers excellent resources for one-word translations. However, it does not give the contextual meaning of words and phrases.

**Microsoft (Bing) Translator (BT) :** By offering only two African languages out of the over 60 languages on its menu, Microsoft Translator vividly outlines the underrepresentation which African languages face in machine translation software systems. Since 2015, Microsoft has collaborated with the *Translators Without Borders* team based in Nairobi to provide the Kiswahili MT engine. *Translators Without Borders* is a non-profit organization composed of professional translators that collaboratively work on translation related to humanitarian issues (Translators Without Borders, 2018).

**Google Translate (GT):** This engine supports over 100 languages including Kiswahili. It is the most popular machine translation engine by virtue of usage. Developed by Google Inc., Google Translate has more than 200 million active users each month (Benjamin, 2014). Google Translate works by analyzing statistical text corpora.

## 1.8 Theoretical Framework

In order to observe and define strategies adopted by MTs for translating the texts, the theoretical framework of this study is based on the translation strategies which were developed by Luc van Doorslaer. The aim is to map out the strategies activated by MTs and to detect the norms that are prevalent when translating into Kiswahili. Van Doorslaer (2007:223) has subdivided the strategies into four groups; those that are of lingual mode (interlingual, intralingual), media printed, audiovisual, electronic), mode, and field. These subdivisions are mapped into the following approaches: functional, foreignizing, exoticizing, localization, and domestication among others. These are represented in Figure 3 as follows:

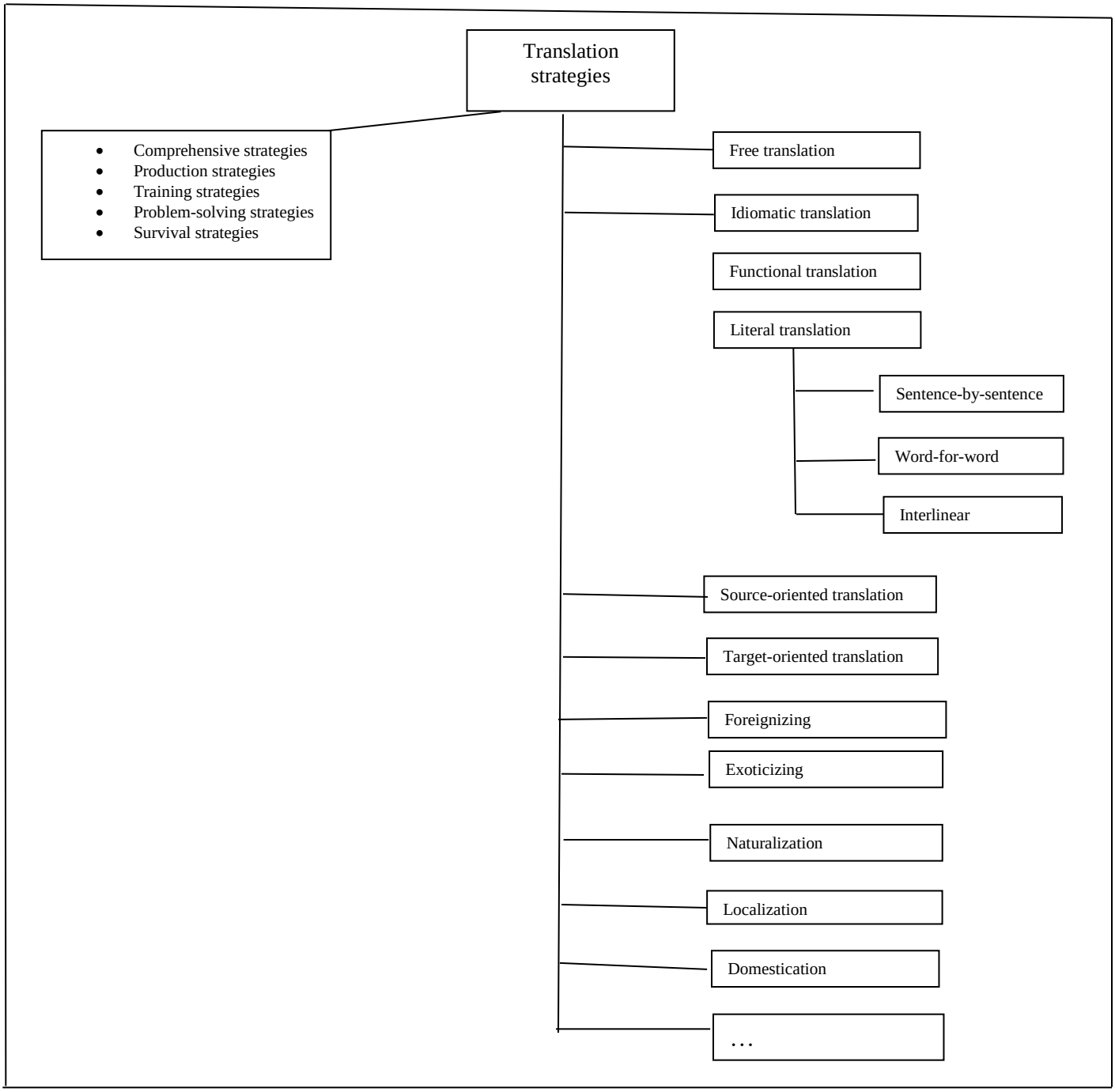


**Figure 1.3: Van Doorslaer's translation strategies (**Van Doorslaer, 2007**)**

To this end, the study will highlight the strategies that are used in MT systems to translate since most of the time, the ST and the TT do not necessarily originate from the same culture, which

will cause some cultural constraints in translation. In addition, some phrases cannot be translated directly into the TT language because they do not exist in that culture. Therefore, to facilitate the transfer of meanings and concepts, MTs utilize various strategies, some of which are more functional and make more sense than others. This study hypothesizes that MTs employ most of Van Doorslaer's translation strategies to convey the cultural otherness of the original text.

### 1.9 Research Methodology

Research methodology refers to the design to be used to get the results of a given research problem. Therefore the research design arrived at should provide credence to the findings derived from the study. It gives assurance that the study is reliable and valid. The methodologies used in translation studies are often based on the principles related to either linguistics or allied disciplines (Saldanha & O'Brien, 2014). The process also included identification and classification of errors, and their quantification.

This study investigates translation strategies, procedures and the role of the MTs in terms of the strategies of Van Doorslaer. Both the original and translated texts are read to discover differences between Swahili and English culture. This process involved several steps. First, culture-specific items (CSIs) in the ST are identified and then compared to their equivalents in the TT to discover how they CSIs are rendered in the TT by the MTs. In the next step, the research identifies the strategies used to convey the items in the TT according to Van Doorslaer.

The primary criterion which served as a guide in the selection of the material for this study was the need to get a sufficiently large corpus that should also cover various genres so that the findings can be considered representative.

### 1.9.1 Data Collection

In order to demonstrate the difficulties facing the use of MT systems for digital resource translation, a descriptive research design was used. This design, according to Kothari (2003), describes the state of affairs as it is. This design involved conducting a study on the UoN repository texts. Details of the texts used and their translations are provided in Appendices A-F.

The study also made use of current academic literature which included books, theses, dissertations, journal articles, and e-databases such as Agora, Emerald, Hinari, Sage, Jstor, Wiley, Springer link and Ebsco.

In order to get reliable conclusions, the MT analysis was administered on a corpus of texts from various genres, of varied content, different language uses and target audiences. The study was performed using well known and easily accessible MT systems. The two MT systems used are available online for free, which was an important consideration for this research. The testing involved using six samples of abstracts of theses already in the IR. These are drawn from the theses posted on the IR for studies done between 2014 and 2017.

### 1.9.2 Data Analysis

Analysis of the study data was conducted using simple descriptive statistics. This aided in providing realization of the study objectives. Secondary data was gathered from textbooks, journal articles, and electronic resources. Restricting the attention of the research to MT, it is pointed out that there two tasks which required an evaluation of the quality:

1. Determining whether the product of an MT system can be of use to a specific discipline;
2. Doing comparative analysis of the systems with one another.

Automatic translation mechanisms work well when it comes to languages that have a similar structure of words. Because this is not the case when it comes to English and Kiswahili, this study did not use an automated method to test the results. Therefore, the analysis mechanism is human-based evaluation in the form of a quality ranking instrument.

## 1.10 Conclusion

University libraries are now involved in managing scholarly resources. They are at the centre of the evolution in the scholarly communication. The evolution of IT, internet and other new information agencies have pushed the library into new realms. Since the academic libraries have embraced IRs which have taken a lot of effort and resources into their creation, it would be of importance to have the resources utilized and have their contents be meaningful to all researchers. Perhaps by having a way through which a user can shift between the various language versions being offered, these invaluable resources would be utilized. Such a system should be able to translate the resource into the language set on a user's device and therefore eliminate the hassle of copying and pasting into a translation app or trying to decipher the meaning by using language dictionaries.

However, it should be noted that African languages are not well catered for in the MT translation front (Benjamin & Radetzky, 2014; Benjamin, 2014). Most of the MTs are developed with European languages in mind; therefore the methods and rules of machine translation are not working well since there are structural differences between European and African languages. African languages use verbs with a complex conjugation rule set that cannot be deployed within the rule sets of European languages. This calls for the development of a translation engine which is relevant to the Swahili language.

This chapter has shown the need for content translation to African languages now that interest in MT is just beginning to take root in Kiswahili as an example of an African language. This chapter has also highlighted the importance of translations in academic usage. The next chapter

discusses in depth the two free MT systems used in this research and outlines the various ways in which the data from the study is organized. It will also delve into the different methods through which MT engines are evaluated.

# CHAPTER TWO

# OVERVIEW OF GOOGLE TRANSLATE AND BING

## 2.0 Introduction

This chapter starts with a presentation in detail of the two MT systems under study which offers translation for the English-Swahili language pair: Google Translate and Microsoft Bing Translator. Thereafter the discussion delves into the various structures adopted in classifying and evaluating the data collected from using the two translation systems.

## 2.1 Google Translate

Basically, Google Translate (GT) was build and made freely available by Google in 2006. This MT system traces its roots from IBM's Candide system. GT was based on Systran software from its inception up to October 2007. As from October 2007, Google made its own technology which it has been using (Sfetcu, N. 2014). GT started out by using statistical models of existing translation data mainly from the United Nations and European Parliament (Smith, D.A. 2006). Now Google does data mining from various web pages on the internet to feed the GT system. Unlike rule-based systems, it does not apply grammatical rules when translating but looks for a matching word from its database to enable it to come up with the best or closest translation. The original creator of the system, Franz Josef Och, now favours empirical approaches instead of the rule-based algorithms (Och, 2009).

To increase translation accuracy, Google has had to overhaul the GT by switching to using Neural Machine Translation (NMT) in 2016. By using this system, now GT translates whole sentences instead of word for word. GT engine incorporates a lot of languages spoken in the world. It supports over 100 languages (Ignacio, 2010; Benjamin, 2014). When compared to

Microsoft Bing Translator (BT), GT has a lot more African language coverage. However, it still lacks enough African language data to enable it to give more accurate results when it comes to various aspects of African languages (Benjamin & Radetzky, 2014).

## 2.2 Microsoft Bing Translator (BT)

Microsoft provides it as part of its Bing services. It started as Live Search Translator and then Windows Live Translator around the year 2000. The Windows Live Translator later in the year 2007 changed to Bing Translator. It also offers a statistical approach to translation. By 2014, speech translation had been added. BT can handle about 70 languages as of June 2018. BT can be freely used to translate up to 2 million characters per month.

In May 2018, Microsoft updated its Application Programming Interface (API) which makes Neural Machine Translation as its default method. BT offers to developers of web pages a free widget which they can add to their sites. It also has a system known as Collaborative Translations Framework (CTF) and the Microsoft Translator Hub through which users can suggest alternative or better translations.

## 2.3 Classification of Texts

De Beaugrande and Dressler (1981), regard text as a communicative medium which obeys the seven standards of textuality. These are acceptability, cohesion, intertextuality, coherence, informativity, situationality, and intentionality. Texts enable readers to get knowledge on a specific issue. Texts are the carriers of messages. For purposes of this study, the texts on focus were divided into two broad areas – scientific and cultural.

Scientific texts have a lot of numbers, formulas, diagrams and symbols, which must not be ignored in the process of translation. The purpose of scientific texts is to discuss a scientific problem which forms the subject of their study. The language of scientific texts requires its readers to have background knowledge in order to understand them. There could also be some technical terms which can only be understood by target readers. Scientific text translation is mostly concerned with the conveyance of facts and ideas. Cultural context of the text is of little concern here. Clarity and conciseness are the main stylistic goals of scientific translation. The scientific texts used in this study are on Appendices D to F.

Cultural texts emanate from humanities and social sciences. The study avoided poems and literary texts. Culture is defined as a community's whole way of life (Geertz, 1975, Lado, 1968). Halliday (1989) sees it "a set of semiotic systems, a set of systems of meaning, all of which interrelate" (p. 4). De Mooji (2004) observed that some texts are deeply imbedded in their culture that they fail when they are transplanted into another culture. The cultural texts chosen for the study are in Appendices A to C. In the library world, the fundamental technique adopted for the organization of knowledge is called classification (Mitchell, J.S. 1996). However for the purposes of this study, the texts chosen were aligned to the colleges of the University of Nairobi. Three abstracts were picked from colleges which offer science-based disciplines, hence they are called scientific texts, while the other three were from the College of Humanities and Social Sciences.

## 2.4 Classification of Data from the Translation Systems

The study was divided into two phases. The first phase had the goal of establishing the output of each MT system. The second phase sought to evaluate the results of each MT system so as to come up with one which could best meet the needs of users. Two different machine translation systems were used in the study. An important consideration was that the systems be freely available online and support the English–Swahili language pair. Then three human translators were used to evaluate the six different translated raw versions of each text whereby each translator got two texts – one from each category. To safeguard against bias by the HT reviewers, the systems were given anonymous identities and the material randomized as tabulated in Table 2.1 below:

**Table 2.1: The randomization of material** (Source: Researcher, 2018)

| **Cultural Text 1** | **Cultural Text 2** | **Cultural Text 3** |
|---|---|---|
| System A = Google Translate | System A = Microsoft Bing | System A = Microsoft Bing |
| System B = Microsoft Bing | System B = Google Translate | System B = Google Translate |
| **Scientific Text 1** | **Scientific Text 2** | **Scientific Text 3** |
| System A = Microsoft Bing | System A = Microsoft Bing | System A = Google Translate |
| System B = Google Translate | System B = Google Translate | System B = Microsoft Bing |

## 2.5 Assessment Criteria of the Translation Systems

According to White (2003), evaluation does not have a strict, rigid form. It can range from using complicated automatic metrics to human evaluation. In this study, the resulting translations were given to humans for review. As Forcada (2010) says, traditional manual evaluation measures use human judges to score the translation. Many judges should be used to review, and their scores averaged out to get a more stable indicator.

The general assessment criteria used was that of fidelity and intelligibility. Fidelity is the determination of accuracy to the meaning of the original while intelligibility looks at the style which enables the text to be easily understood and comfortable to read. The reviewers were asked to rate the translations against the source text using a scale of 1 to 3, where 1 was the best score and 3 the poorest. The criteria used by Fiederer & O'Brien (2009) was adopted with a slight change as shown in Appendix G.

## 2.6 Distribution of Translation Errors

Translation errors are classified variously by translation scholars. They can be based on conventions, format, style, semantics, register, origin, grammar and syntax, and construct among other aspects. Analyzing these errors can either be measurable or immeasurable (Solano-Flores, G. et al, 2009). This study modified Temnikova's (2010) classification of error types and used them to measure these errors by noting the number of instances particular error types occurred in each MT. The error types were grouped into incorrect word, extra word, missing word, not translated, wrong word order, and miscellaneous errors. It is important to list the errors as they are a pointer to the quality of a system. It should also be noted that the translation errors are listed depending on text types.

## 2.7 Translation Strategies

In the process of translation, a set of actions or strategies are applied in order to achieve a successful translation. Strategies are generally procedures or techniques which are used to chieve a goal. Translation strategies are usually applied when translators are transferring the meaning of the ST to the TT. The list of translation strategies adapted from Van Doorslaer (2007) have been applied to examine the translation of various IR resources into the Swahili language. Prior to the in-depth structural textual analysis of translation strategies, preliminary analysis was carried out.

## 2.8 Conclusion

This chapter has given an overview of the way data from the study is classified for analysis. In addition, the two systems under study were discussed in detail. The following chapter is going to present the actual results of the study.

# CHAPTER THREE

# DATA PRESENTATION

## 3.0 Introduction

This chapter presents the results of machine translations of various texts. It also gives analysis of those findings. Data analysis is the process of reducing large amounts of collected data to make sense out of them. Qualitative analysis is the data analysis technique used. Qualitative data analysis is suitable for non-numeric information such as interview transcripts, notes, video, and audio recordings, images and text documents. As mentioned earlier, the analysis is mainly based on the translation strategies described by Luc Van Doorslaer (2007).

## 3.1 Form of Presentation

The findings are presented in form of narrations, tables, and charts where applicable. Some of the methods of qualitative data interpretation (Creswell J.W., 1998; Miles, M.B., 1994) included:

- *Evaluating words and phrases* – scanning the primary data for words and phrases used and how they are translated;
- *Primary and secondary data* – comparing findings of any other translations with the findings of this study;
- *Noting missing information* – discussions about the aspects of the translations which did not happen as expected.

To achieve its objectives, the study used Van Doorslaer's (2007) strategies. The following six strategies were selected to find out the approach most frequently utilized by the MT systems in Kiswahili translation of the various abstracts:

1. Free Translation: Use of this strategy produces a TT which is independent of the style, form, or content of the ST.
2. Foreignizing: This strategy retains information which was in the ST thereby deliberately ignoring the conventions and norms of the TT (Gile, 2009).
3. Literal Translation: This strategy retains the forms of the ST as much as possible. Since it translates lexical words singly, they are normally out of context (Newmark, 1988b, p.81).
4. Source-oriented strategy: the focus of translation is on original idea.
5. Target-oriented strategy: the texts created are literary, but are not the exact translations of the originals (Toury 1993).
6. Functional translation: This strategy opines that the quality and quantity of a translation depends on its skopos (Nord, 1991, p.28).

In instances where the MTs were unable to translate because of lexical and syntactical differences between the two languages, it was noted that the Human Translators (HT) also came up with their own strategies and procedures to render the translation complete. They used the following oblique translation methods when giving Kiswahili translations of the abstracts:

1. Borrowing: Refers to reproduction or transliteration of the ST (Harding & Riley (1986).
2. Adaption: This is a free form of translation though used mainly for plays (comedies) and poetry (Newmark, 1988b, p.81). Used in situations when cultural differences occur between the ST and the TT.
3. Calque: This is a special type of borrowing in which the borrowed expression is literally translated into the target language.
4. Literal translation: It is a strategy which renders ST into the appropriate idiomatic or grammatical equivalent in the TT.
5. Equivalence: this refers to rendering two situations by different stylistic and structural methods; these two texts include the ST and its equivalent TT.

### 3.2 Results of the Study

According to Van Doorslaer's (2007) translation strategies, data collection and analysis should be carried out in a comparative descriptive framework. In other words, each term needs to be interpreted in the original language (English) and then its Kiswahili translation. These two interpretations were compared by using Van Doorslaer's (2007) translation strategies and procedures. Finally, the study categorized these strategies and procedures and calculated their frequencies.

According to Table 3.1 below there were 55 instances of Literal Translation, 103 instances of Source-oriented translation, 98 of Target-oriented translation, 11 instance for Foreignizing, 8 of Functional translation, and 21 instance for Free Translation.

**Table 3.1. Applied translation strategies in cultural texts**

| Strategies | Frequency | Percentage |
|---|---|---|
| Free translation | 21 | 7 |
| Foreignizing | 11 | 4 |
| Literal translation | 55 | 19 |
| Functional translation | 8 | 3 |
| Source-oriented translation | 103 | 35 |
| Target-oriented translation | 98 | 33 |
| Total number | 296 | 100 % |

When the cultural texts (Table 3.1) are compared with the scientific texts (Table 3.2), the observation is that source-oriented translation has been the dominant strategy in both texts. Scientific texts also tend to use the literal translation strategy as opposed to target-oriented strategy in cultural texts. Source-oriented and literal translation strategies support each other in scientific texts. Infact, some of the terms are carried over from the ST into the TT. Therefore scientific translation using MT is prone to more errors.

**Table 3.2. Applied translation strategies in scientific texts**

| Strategies | Frequency | Percentage |
|---|---|---|
| Free translation | 54 | 12 |
| Foreignizing | 68 | 15 |
| Literal translation | 97 | 21 |
| Functional translation | 31 | 7 |
| Source-oriented translation | 128 | 27 |
| Target-oriented translation | 89 | 19 |
| Total number | 467 | 100 % |

Overall, the MTs rated fairly for correctness, as only a few sentences were judged incorrect either with regard to meaning or with regard to language. As Temnikova (2010) observed, word form errors are minor while errors emanating from missing words, wrong idiomatic expressions and wrong word order affect meaning and are therefore major.

### 3.3 Analysis of errors

To explore the differences in each MT, a more detailed error analysis was performed on the sentences identified as particularly difficult. Tables 3.3 and 3.4 below give the results as organized according to the MTs.

**Table 3.3: Distribution of errors that are critical for comprehension of cultural texts**

| Error type | No. of instances in GT | No. of instances in Bing |
|---|---|---|
| Incorrect word | 89 | 108 |
| Extra word | 17 | 19 |
| Missing word | 19 | 29 |
| Not translated | 78 | 95 |
| Wrong word order | 123 | 149 |
| Miscellaneous errors | 77 | 83 |
| Total | 403 | 483 |

**Table 3.4: Distribution of errors that are critical for comprehension of scientific texts**

| Error type | No. of instances in GT | No. of instances in Bing |
|---|---|---|
| Incorrect word | 105 | 127 |
| Extra word | 13 | 18 |
| Missing word | 26 | 31 |
| Not translated | 125 | 142 |
| Wrong word order | 87 | 99 |
| Miscellaneous errors | 93 | 101 |
| Total | 449 | 518 |

After having analyzed the errors in translation of both scientific and cultural texts, it can be concluded that GT produces a lower number of errors and therefore it is more qualitative.

This conclusion is also supported by the overall ranking given by the human translators as Table 3.5 below shows.

**Table 3.5: Overall translation quality rankings of MT by HT**

| | Google Translate | | | Bing Translator | | |
|---|---|---|---|---|---|---|
| | Grammar | Fidelity | Style | Grammar | Fidelity | Style |
| **Scientific Text** | 1.3 | 1.3 | 2 | 2.3 | 2 | 2.3 |
| **Cultural Text** | 1.3 | 1 | 1.6 | 2.3 | 1.6 | 2 |

An overall quantitative analysis of the texts show that the errors are mainly on the morphological level (incorrect word forms and typographical errors) and on the lexical level (incorrect style and missing words). Some of the errors also contain errors related to word order and idioms.

An example of a successfully machine translated sentence is the following:

Example 1

ST: *Work-family conflict has adverse effects on behaviour, emotions, and health.*

MT: *Migogoro ya kazi-familia ina athari hasi kwa tabia, hisia na afya.*

On the other hand, some sentences have multiple errors as in the case of Example 2:

Example 2

ST: *The research employed descriptive research design to examine the work-family conflict among the employees of the State Department of Coordination, Ministry of Interior and Coordination of National Government.*

MT: *Utafiti uliofanywa utafiti unaoelezea utafiti kuchunguza migogoro ya kazi na familia kati ya wafanyakazi wa Idara ya Udhibiti wa Nchi, Wizara ya Mambo ya Ndani na Uratibu wa Serikali ya Taifa.*

The above example shows that the MT system used a word-for-word translation strategy which made the whole sentence meaningless. This was more noticeable in scientific texts than in the cultural texts.

There are other errors which are categorised as untranslated word and incorrect punctuation. For instance, the Bing Translator has not been able to translate ‘satisfaction’ and ‘data’ while Google Translate left ‘package’, ‘frequency’ and ‘masculinity’ among other words in English. However in some instances, it was noted that if the ST word was wrongly transcribed, then it does not get translated at all. An example is the word ‘incooporates’ in Appendix C. Among the texts with the most number of untranslated words or terms were the scientific texts. Perhaps this is owing to the fact the terms are yet to be localized. The errors from scientific texts could be considered cognitively more difficult for MTs as clearly none of the MTs had been able to figure out the correct meaning.

Example 3

ST: *radiographs*

MT: *radiographs*

In some instance, words are missing from the translation thereby changing the meaning significantly. In some translation cases, even one error in an expression may severely affect the sentence. An example of where an error in one word form significantly change the entire sentence, is given below:

Example 4

ST: *A cross sectional study was done at…*

MT: *Utafiti wa sehemu ya msalaba ulifanyika ...*

In the study, since not all errors were listed, others were lumped under miscellaneous errors. There were errors in punctuation which do not seem to affect the meaning of the MT output. However, those where incorrect words or incorrect word order forms occurred are crucial since they could change the meaning of the translation. The results of the study indicate that MT texts of different genres exhibit different distributions of error categories.

### 3.4 Ranking of MTs by Human Translators

Using the assessment criteria in Appendix G, the human translators were asked to provide a comparative ranking for each text. These rankings are presented in Table 3.6 below, and an average score was calculated for each MT system.

**Table 3.6: The MT rankings**

| | Google Translate | | | Microsoft Bing Translator | | |
|---|---|---|---|---|---|---|
| Scientific Text | Grammar | Fidelity | Style | Grammar | Fidelity | Style |
| Translator 1 | 1 | 2 | 2 | 2 | 2 | 3 |
| Average | 1.6 | | | 2.3 | | |
| Translator 2 | 2 | 1 | 2 | 3 | 2 | 2 |
| Average | 1.6 | | | 2.3 | | |
| Translator 3 | 1 | 1 | 2 | 2 | 2 | 2 |
| Average | 1.3 | | | 2 | | |
| Cultural Text | | | | | | |
| Translator 1 | 2 | 1 | 2 | 3 | 2 | 2 |
| Average | 1.6 | | | 2.3 | | |
| Translator 2 | 1 | 1 | 2 | 2 | 1 | 2 |
| Average | 1.3 | | | 1.6 | | |
| Translator 3 | 1 | 1 | 1 | 2 | 2 | 2 |
| Average | 1 | | | 2 | | |

## 3.5 Conclusion

Based on the findings of this study, the dominant strategy used in using MTs to translate was Source-oriented Translation. This is closely followed by Literal and Target-oriented translation. The other strategies are also used but to a minor extent. This means that MTs transfer terms to the TT with almost exact content of the ST.

# CHAPTER FOUR

# DISCUSSION OF FINDINGS

## 4.0 Introduction

This chapter presents indepth discussion of the data presented in Chapter three. It was observed that the nature of the information being translated dictates the kind of MT engine to be used since not all content, file formats and language pairs can use the same engine.

## 4.1 Suitable MT for translating IRs

To answer the research question on the most suitable and convenient MT which could be used to translate IRs in the UoN repository, the frequencies of the errors occurring in Kiswahili translation of various English abstracts in the IR were determined separately.

Based on the findings, the texts translated by GT were more accurate than those produced by BT. Other studies have suggested that GT can even be more reliable with the involvement of a human translator (Aiken and Balan, 2011; Khosravizadeh and Pashmforoosh, 2011). GT selects proper lexis for cultural texts when compared to BT. Perhaps GT has a richer corpus of literary terms and so is able to identify the proper meanings of the terms. In evaluating scientific texts (such as technology, medical), BT has better performance. This is also confirmed by a study done by Dhakar, Sinha, and Pandey (2013) in the English-Hindi language pair.

## 4.2 Errors Encountered

Both MT systems had some errors in spellings. As shown in tables 3.3 and 3.4, the most frequently occurring errors were from category of “incorrect word”, “not translated word”, and

"wrong word order". For example, part of the translation done by Microsoft Bing Translator goes like this:

> "*Utafiti huu inachunguza jukumu la intertextuality na utendaji katika uumbaji wa maana na aesthetic kukata rufaa katika maneno ya mashairi ya washairi kumi na nne wa Kenya* …" (Appendix C).

The above does not make sense at all. Full details of error analysis revealed that amongst the errors, there were also those of rendition, subtle differences, and misinterpretation of the source text. The MTs committed insufficient rendering and misinterpretation perhaps due to lack of the correct terms.

Most of the errors were noted to emanate from words that were not translated and left intact in the TT. In the same Appendix C, words like "intertextuality, aesthetic, jazz, hip-hop, heterogeneity" were left untranslated by Microsoft Bing Translator. Google Translate translated all except "jazz, hip-hop and dub". This would be as a result of word limitations in the MTs' dictionaries. Some would be as a result of: native language transfer interference; intralingual errors due to differences between the two languages; sociolinguistic situation; the way in which the MT is exposed to the TL and the manner of producing a text; the complexity of the particular texts to be translated. (Chomsky, 1969, p. 49).

Going by the list of taxonomies of errors proposed by the American Translation Association (ATA, 2009, p. 17), the following are the ones commonly noted in the study: mistranslation into target language, addition or omission, terminology, word choice, register, too freely translated, too literal, word-for-word translation, ambiguity, grammar, syntax, cCase (upper case/lower

case), word form, usage, and style and form. However, in this study all the above taxonomies were compressed into four units; incorrect word, extra word, missing word, not translated, wrong word order.

When the human translators were asked to rank the MTs, they noted inappropriate register in some cases. MTs are insensitive to cultural and social norms, therefore the terms might come out differently from the expected (Solano-Flores, G. et al, 2009). Grammatical mistakes which also included awkward expressions were noted. In some instances there was excessive free translation as well as cases of omission. In total, miscellaneous errors were the most frequently errors followed by language errors.

Of the two categories of texts, the scientific register category had more errors than the cultural category. According to Webster's ninth New Collegiate Dictionary, register is a variety of a language that is appropriate to a particular subject or occasion (Webster, 2009). Perhaps this implies that the MT engines have had very little to learn from scientific texts in Kiswahili.

### 4.3 Consequences of Errors

Possible reasons for the occurrence of errors would be that the MTs misunderstood the source texts or were unable to produce the target text in some instances.

As noted by Albir (1995, as cited in Waddington, 2001), errors would lead to:

- Inappropriate translations which affect how ST is understood,
- Improper translations which would change the expression of the TT

- Inadequate translations which would affect the transfer of the primary or secondary function of the ST.

The above assertion is also affirmed by Solano-Flores, G., et al (2009) who opine that the inaccurate use and/or omission of technical terms may lead to wrong interpretation of a technical text.

### 4.4 Effects of Translation Strategies

The second objective of this study was to evaluate the most suitable open source MT system which could be used to translate IR resources. To determine this, the study sought to identify the way in which culture-specific items were translated and the effects of the strategies used. This also partly answers the first objective in terms of suitability. It is common knowledge that translation involving human translators is influenced by background knowledge of a language and culture. The use of different translation strategies will have certain effects on a translation.Therefore this study sought to find out if the same applied to MTs through highlighting the strategies used.

### 4.5 Conclusion

The findings of the study affirm that MT has not yet gained as much accuracy as human translation when translating texts from English to Kiswahili. Some segments of the translated text depart from the accepted norms of the Swahili language. However, the translations rendered are still suitable and give an indication of what the content is all about.

# CHAPTER FIVE

# SUMMARY, CONCLUSION AND RECOMMENDATIONS

## 5.0 Introduction

This chapter provides the summary of findings and conclusions emanating from the study. Lastly, recommendations derived from the study are also made.

## 5.1 Summary of Findings

This study's broad objective was to establish the viability and challenges that would face the use of machine translation for digital collections in the University of Nairobi institutional repository.

The challenges facing the use of MT were identified to be caused by broad factors, chiefly among them being: translation inadequacies and omissions. Translation inadequacies can lead to meaning loss and distortion of the message. The analysis also shows that MT is also facing the challenges of linguistic inconsistency, and in some cases, vagueness among others. Most MT errors are due to their incompetence of lexical or structural complexities between English and Kiswahili. Misunderstanding of several words and choosing an inadequate solution can lead or produce meaningless structures. These inadequacies prune the aesthetic value of the expression or playing words. These challenges are presented by lack of equivalence and non recognition of the cultural concepts in the TT.

The two MT systems in the study show that although they can perform a literal translation of ST, they don’t create a dynamic equivalence between the TT and ST perfectly. Most of the problems

identified are due to ambiguities as well as the varying styles and usage of language in the different subject areas. Therefore interlingual automatic translation requires a bit of pre-editing and post editing at the moment. Overall, MT can benefit human translators who can now devote more time to refine their translations and deepen their craft, rather than translating regular or repetitive texts. Therefore to a certain extent, information professionals may depend upon these MT tools.

While MT has performed remarkably well in controlled language domains (like weather forecasts, computer manuals, and stock exchange reports), it has not been able to reach the richness, depth, and accuracy of human translation, mainly when dealing with general language (Giammarresi and Guy Lapalme, 2016). In this case, the community of practice experts in those specific domains have shaped their language over time, and all the terms are codified. As Melby (1995: 52) states, in general language, a word is like clay while it is a hard stone in a domain-specific language. Therefore general language use has very loose (clay) words that keep changing depending on usage, context, speaker and audience, among other factors while domain-specific language is based on a hardened (hard stone) network of terms and meaning, where there is no ambiguity. MT is perfect for some languages; however, it leaves a lot of room for improvement for most language pairs, including Kiswahili. The primary source of this problem is noted to be lack of high-quality data. MT systems require sufficient high quality data to train them. It is therefore imperative upon the speakers and experts in those languages which are underrepresented to supply as much data as possible to enable MTs in those languages to improve. Generally, language is dynamic, ambiguous and rooted in ethical, economical,

pragmatic relationships. Language does not have a single approach to translation as it is evolutionary in nature.

## 5.2 Conclusion

The results of this study imply that MTs need to be trained to translate meaning and not just texts. Misinterpretation occurs when they do not consider culture. When they translate the cultural words literally, the meaning is distorted and the target language readers or audience will be culturally shocked. Besides, each language has unique characteristics and one of the main translation problems is to find possible strategies/procedures and analyze specific translations. Generally, it is observed that translation over time has been skewed in favour of religious and literary works (Smith, D.A. 2006). These works have been translated over and over and in several languages while the rest of the universe of knowledge never gets to be translated. Perhaps this kind of scenario can be seized upon so that those works which have been massively translated can be used to build and train MT systems in the languages which are underrepresented. Other works which are moderately translated can be used to align existing texts and provide training to upcoming translations, and finally, new MT systems could provide multilingual finding aids to aid in translating new texts (Figure 5.1). This is where this study comes in.

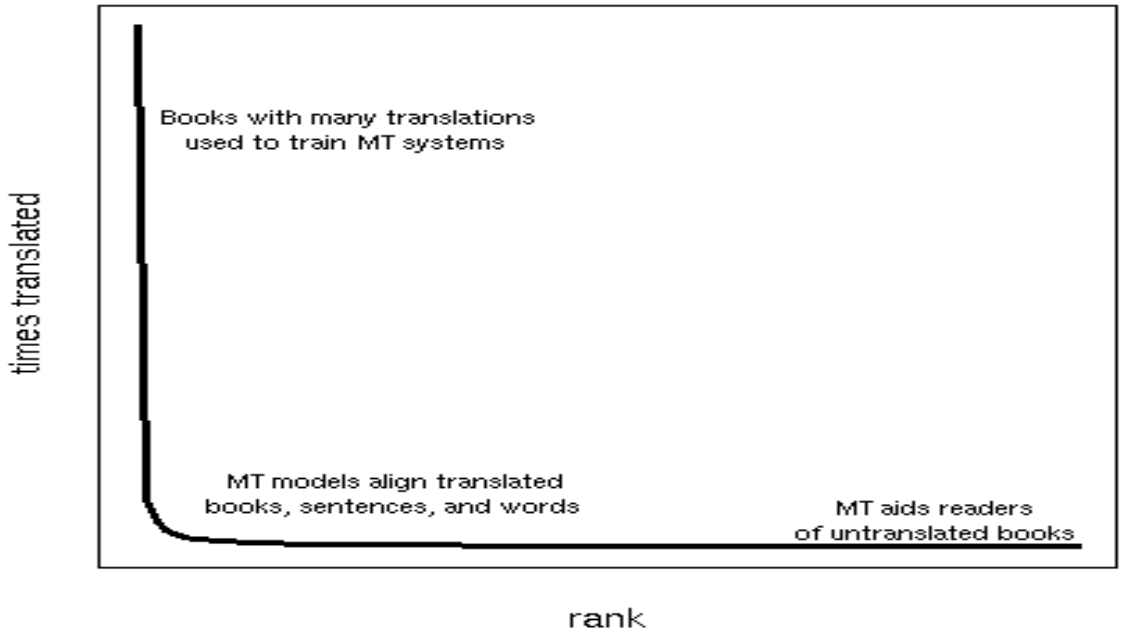


**Figure 5.1: Knowledge translation scenario using MT** *(Adapted from Smith, D.A. (2006))*

In the face of globalization, the assertion of Lori Thicke, co-founder of Lexcelera and Translators Without Borders, that lack of knowledge rather than diseases has led to the death of many people in developing countries since they are unable to access global knowledge in a language they understand is true.

Though the quality of human translation is high, the time taken to produce a translation is long, and it is a costly affair compared to MT. Also, human translators who are subject experts are few. It is not feasible to produce consistent translations when working with large quantities of texts manually.

The use of technology in translation is necessary in the profession for translators to be able to achieve high quality and increase productivity. Soon this technology will be available in mobile

devices as it is evolving. This does not mean that technology is taking over the job of human translators. Instead, it is complementing their work, but it calls for translators to acquire new skills to understand and join the digital revolution. At its International Annual Meeting on Language Arrangements, Documentation and Publications, held in Paris in June 2010, the United Nations (2010) declared: “Without a new generation of trained linguists and professionals with language skills, international organizations will be unable to perform their vital tasks.” Though in the process of doing so, some knowledge may be lost at the expense of learning a new one (Mutte, 2010), seasoned translators would still be valuable because of their vast amounts of tacit knowledge (Nonaka, 1994) which could be handy in situations where technology to translate.

The mission of the university and the objectives of the library can be achieved if the university administrators and library management provide opportunities to address the challenges identified by this study. According to Makori & Mauti (2016), information management is now changing and the use of innovative technologies and new multimedia technology is the way to go. It is high time libraries adopted technological solutions, emerging technologies and mobile devices in their service provision. Several studies show that the information infrastructure in the African education sector is laden with human resources with a dearth of adequate knowledge, skills and competencies to face the digital revolution (Kandiri, 2012, Zaman et al., 2011). Libraries and information systems cannot perform well if their intellectual and capital assets are below par (Canadian Association of Research Libraries 2010). Digital technology has permeated all sectors of life. Libraries have to adapt to social and technological changes. Librarians should take up the role of guiding scholars and other users towards getting valuable information instead of leaving them to grope blindly through the digital world.

### 5.3 Recommendations

Information technology has proved that MT processes can be automated despite challenges confronting its advancement. This technology can also be used to cross-link IR resources. Information technology processes need to evolve, process and disseminate information into an intuitive interface which consistently gives the desired results in the language of the user. The study recommends that a widget to aid users of the IR to be included on the university website. This would detect and align the required translations to the source text. An IR provides a good ground for translation modelling. Apparently, the resources are in digital format and from a broad range of knowledge. The resources will be of value when they can be read by the prospective users in a language they are comfortable with. Provision of MT services that can help IR users make some sense of the content are significant, however flawed that service may be.

Libraries are an essential component in educational institutions. Since there is an increasing number of virtual users, the time has come whereby information technology has to be embraced in full. Information technology linkage with library collections will allow simultaneous and multiple access to content regardless of geographical location, time or language. Overall, the university's profile will be raised. A similar study is recommended to be done with a broader scope involving more universities which will incorporate research into the end users' acceptance of MT output and its implications.

This research investigated the translation errors in the context of a limited sample of cultural and scientific texts, further research should be conducted in literary texts, religious texts, and other genres. Furthermore, the study would not determine the effect of each error category or all the errors on the whole. Therefore, the frequency of errors and their impact on and correlations to users might also be a subject for further research.

## REFERENCES


Abiola O., Adetunmbi, O. and Oguntimilehin, A. (2015). A Review of the various approaches for text to text machine translations. *International Journal of Computer Applications 120* (18) pp. 7-12.

Agosti, M., Angelaki, G., Coppotelli, T., and Di Nunzio, G. (2007). "Analysing HTTP Logs of a European DL Initiative to Maximize Usage and Usability." In *Asian Digital Libraries. Looking Back 10 Years and Forging New Frontiers*: 10th International Conference on Asian Digital Libraries, ICADL 2007, Hanoi, Vietnam, December 10-13, 2007. Proceedings, edited by Dion Hoe-Lian Goh, Tru Hoang Cao, Ingeborg Torvik Sølvberg, and Edie Rasmussen, 35–44. Berlin, Heidelberg: Springer Berlin Heidelberg. Viewed August 8, 2018, from http://dx.doi.org/10.1007/978-3-540-77094-7_9.

Aiken, M., & Balan, S. (2011). An analysis of Google Translate accuracy. *Translation Journal and the Author, 16* (2). Viewed September 11, 2018, from: http://translationjournal.net/journal/56google.htm.

Amano, T., González-Varo, J.P., Sutherland, W.J. (2016). Languages are still a major barrier to global science. *PLoS Biol 14*(12): e2000933. Viewed May 8, 2018, from: https://doi.org/10.1371/journal.pbio.2000933

American Translators Associations (2009). *ATA Certification Program Certification Exam: An Informational Presentation*. Retrieved on September 1, 2018 from http://www.atanet.org/certification/aboutexams_presentation.php

Amollo, B. A. (2011). *Digitization for libraries in Kenya. 2nd International Conference on African Digital Libraries and Archives*.

Andy, W. and Gough, N. (2005). Comparing example-based and statistical machine translation. *Natural Language Engineering*, pp. 295-309.

Association of College and Research Libraries (ACRL). 2000. *Guidelines for distance learning library services*. Viewed on April 11, 2018, from http://www.ala.org/Content/NavigationMenu/ACRL/Standards_and_Guidelines/Guidelines_for_Distance_Learning_Library_Services.htm

Baikady, M.R., Jessy, A., Bhat, K.S. (2014). Off Campus Access to Licensed E-resources of Library: A Case Study. *DESIDOC Journal of Library & Information Technology, 34*(6) 2014. Viewed May 8, 2018, from http://publications.drdo.gov.in/ojs/index.php/djlit/article/view/7509.

Benjamin, M. (2016). *Comparing Kamusi and Google Translate*. The Kamusi Project. Viewed July 21, 2018, from: http://kamusi.org/google-translate.

Benjamin, M. (2014). *How does Google translate work? Do they have database for all words of a particular language?* Quora. Retrieved July 4, 2018 from: https://www.quora.com/How-does-Google-translate-work-Do-they-have-database-for-all-words-of-a-particular-language.

Benjamin, M. & Radetzky, P. (2014). Multilingual Lexicography with a Focus on Less-Resourced Languages: Data Mining, Expert Input, Crowdsourcing and Gamification. In *Workshop on Collaboration and Computing for Under-Resourced Languages in the Linked Open Data Era Language Resources and Evaluation Conference*, Reykjavik, Iceland. Universit di Pisa, Dipartimento di Linguistica.

Bjorneborn, L. & Ingwersen, P. (2004). Towards a basic framework for Webometrics. *Journal of the American Society for Information and Technology. 55*(14), 1216-27.

Bowker, Lynne (2014). "*Can machine translation facilitate outreach to newcomers? : a pilot study investigating the needs of Spanish-speaking users of the Ottawa Public Library*." 2013 OCLC/ALISE research grant report published electronically by OCLC Research.

Bryman, A & Emma, B. (2004). *Business research methods*. Oxford: Oxford University Press.

Budapest Open Access Initiative. (2002). *Budapest Open Access Initiative*. Retrieved on 7 April, 2018, from: http://www.soros.org/openaccess/

Burkhardt, A. (2010). Social media: a guide for college and university libraries. *College & Research Libraries News, 71*(1): 10-24.

Canadian Association of Research Libraries (2010). *Core competencies for 21st century CARL librarians.*

Carbonell, J. G. and Brown, R. D. (2004). *Example based machine translation.* Available at: http://www.cs.cmu.edu/~ralf/ebmt/ebmt.html (Accessed on 25th April 2018)

Chomsky, N. (1969). Quine's Empirical Assumptions. In D. Davidson, J. Huntikka (eds.), *Words and Objections; Essays on the work of W. V. Quine.* Reidel: Dordrecht, Netherland.

Chu, S.K.W. & Du, H.S. Social networking tools for academic libraries. *Journal of Librarianship and Information Science, 45*(1)

Colins, B. & Moonen, J. (2002). *Flexible learning in a digital world: experiences and expectations.* London: Routledge.

Collins, G. & Quan-Haase, A. (2012). *Social media and academic libraries: current trends and future challenges.* 49(1): 64-75.

Collins, G. & Quan-Haase, A. (2014). Are social media ubiquitous in academic libraries? A longitudinal study of adoption and usage patterns. *Journal of Web Librarianship, 8*(1): 48-68.

Communications Authority of Kenya (2015). *The Programming Code for Free-to-Air Radio and Television Services in Kenya.* Communications Authority of Kenya.

Cronin, M. (2013). Translation and globalization. In C. Millan & F. Bartrina (eds.). *The Routledge handbook of translation of translation studies* (pp. 491-502). London: Routledge.

Crow, R. (2002). *SPARC institutional repository checklist & resource guide*. Washington, SPARC. Viewed May 8, 2018, from: http://www. arl.org/sparc/IR/IR_Guide.html.

De Beaugrande, R. A. & Dressler, W. (1981). *Textlinguistics*. New York: Longman.

Debowski, S. (2003). *Information seeking in tertiary settings: the challenge for educators*. In Partners in Learning: Proceedings of the 12th Annual Teaching Learning Forum, 11-12 February 2003. http://lsn.curtin.edu.au/tlf/tlf2003/abstracts/debowski-abs.htm. (Accessed 24 April 2018).

De Mooij, Marieke. (2000). Translating advertising. *The Translator10* (2), 179-198.

Dhakar, B. S., Sinha, S. K., & Pandey, K. K. (2013). A survey of translation quality of English to Hindi online translation systems (Google and Bing). Retrieved from: *International Journal of Scientific and Research Publications, 3*(1), p. 1-4. Retrieved July 4, 2018 from: http://www.ijsrp.org/research-paper-1301/ijsrp-p1355.pdf.

Dobreva, Milena, and Sudatta Chowdhury. 2010. "A User-Centric Evaluation of the Europeana Digital Library." *In The Role of Digital Libraries in a Time of Global Change*. ICADL 2010, edited by Gobinda Chowdhury, Chris Koo, and Jane Hunter, 6102:148–57. Lecture Notes in Computer Science. Springer Berlin / Heidelberg.

Drott, M.C. (2006). Open Access. *Annual Review of Information Science and Technology (ARIST) (40),*79-109.

Epstein, B.J. (2012). *Translating Expressive Language in Children's Literature*. By Any Other Name? Translating Names. [Ebook Central]. Peter Lang AG, Internationaler Verlag der Wissenschaften. [Cited 27.3.2017]. 67–98. Limited access: https://ebookcentral-proquest-com.proxy.uwasa.fi/lib/tritonia- ebooks/reader.action?docID=1056074&ppg=81

Ethnologue (2018). *Languages of the World. SIL International Publications*. Retrieved on August 3, 2018 from https://www.ethnologue.com/

Feeney, M. (2004). Centralizing information about library services and resources: delivering the library to users at any distance. *Internet Reference Services Quarterly, 9*(1-2): 129-146.

Flammia, M. & Saunders, C. (2007). Language as power on the Internet. *Journal of the American Society for Information Science and Technology, 58*(12): 1899-1903.

Forcada, M. L. (2017). Making sense of neural machine translation. *Translation Spaces 6:2* (2017) 291–309

Fulkerson, D. M. (2012). *Remote access technologies for library collections: tools for library users and managers*. Hershey, PA: IGI Global.

Gall, M. D., Borg, W. R. & Gall, J. P. (1996). *Educational research: an introduction*. 8th ed. New York: Longman.

Gay, L. R. (1992). *Educational research: competence for analysis and applications*. 4th ed. New York: Macmillan Publishers.

Geertz, C. (1979). *The interpretation of culture: Selected essays*. London: Fontana Press.

Giammarresi, S. and Lapalme, G. (2016). Computer science and translation: natural languages and machine translation. In *Border Crossings : Translation Studies and other disciplines*. Edited by Yves Gambier and Luc van Doorslaer. John Benjamins Publishing Company.

Google Translate. (2016). *Google Translate*. Alphabet, Inc. Retrieved from http:// www.translate.google.com

Gough, N. and Way, A. (2004). Example-Based Controlled Translation. In: *Proceedings of the Ninth EAMT Workshop,* pp. 73-81.

Gudo, C. O., Olel, M. A. & Oanda, I. O. (2011). University expansion in Kenya and issues of quality education: challenges and opportunities. *International Journal of Business and Social Science, 2* (20), 203-214.

Gust, K. J. & Haka, C. H. (2006). Bringing users back to the library: a case history. *New Library World, 107*, 141-148.

Gutt, Ernst-August. (2000). *Translation and Relevance: Cognition and Context*. Manchester & Boston: St. Jerome Publishing.

Halliday, MAK. & Hassan, R. (1989). *Language, context, and text: Aspects of language in a social semiotic perspectives.* London: Edward Arnold.

Haliso, Y. (2011). Factors affecting information and communication technologies (ICTs) use by academic librarians in Southwestern Nigeria. *Library Philosophy and Practice (e-journal), Paper 571*. Viewed 16 February 2018, from http://digitalcommons.unl.edu/libphilprac/571.

Harnard, S. (2003). Eprints: Electronic preprints and postprints. In: *Encyclopedia of Library and Information Science*, Marcel Dekke.

Hitchcock, S. (2013). *The effect of open access and downloads ('hits') on citation impact: a bibliography of studies.* University of Southampton.

Horo, S. N. (2006). *Utilization of electronic information resources in academic libraries in Kenya: a comparative study of Jomo Kenyatta University of Agriculture and Technology and United States International University*. Unpublished Masters Dissertation. Nairobi: Kenyatta University.

Hutchins J, Example based machine translation – a review and commentary. In *Recent advances in example-based machine translation.* (Ed. Michael Carl and Andy Way) (2003). Available at: http://www.hutchinsweb.me.uk/MTJ- 2005.pdf (Accessed on 25th April 2018)

International Federation of Library Associations and Institutions (IFLA). (2013). Riding the Waves or Caught in the Tide? Navigating the Evolving Information Environment. *Insights from the IFLA Trend Report*. The Hague: IFLA. http://trends.ifla.org/files/trends/assets/insights-from-the-ifla-trend-report_v3.pdf (Retrieved March 2, 2018).

Jiangping Chen, Olajumoke Agozu, Wenqian Zhao, Cheng Chieh Lien, Ryan Knudson, & Ying Zhang: *Integrating machine translation with digital collections for multilingual information access.* AMTA-2012: the Tenth Biennial Conference of the Association for Machine Translation in the Americas. Proceedings, San Diego, CA, October 28 – November 1, 2012;

Kandiri, M. (2012). *A survey on ICT access and use in Kenya secondary schools*. Summit Strategies Ltd, Nairobi: Kenya.

Khosravizadeh, P. & Pashmforoosh, R. (2011). *Google translation: a semantic structure analysis*. Retrieved August 11, 2018 from: https://researchgate.net/publication/216597782-_Google_Translation_A_Semantic_Structure_Analysis

Koehn, Philipp (2010). "Enabling monolingual translators: Post-editing vs. options." *NAACL HLT 2010: Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*. Proceedings (Los Angeles, California, June 2010), 537-545. http://aclweb.org/anthology-new/N/N10/N10-1078.pdf (Retrieved on September 12, 2018)

Koponen, Maarit and Leena Salmi (2012). "*Does prior use of machine translation systems help in post-editing?"* Paper presented at International Workshop on Expertise in Translation and Post-editing: Research and Application (Copenhagen Business School, 17-18 August 2012).

Kothari C R, (2003). *Research methodology : method and techniques*. 2nd ed. New Delhi : Wishwa Prakashan.

Kratz, J.E. & Strasser, C. (2015). Researcher perspectives on publication and peer review of data. *PLOS ONE, 10*(2): e0117619. Viewed May 8, 2018, from http://doi.org/10.1371/journal.pone.0117619

Lado, R. (1968). *Linguistics across cultures: Applied linguistics for Language teachers.* Ann Arbor: The University of Michigan Press.

Leclaire, A. (2004). *African Union recognizes Kiswahili*. http://www.powerofculture.nl/en/taxonomy/term/128.html (Accessed on Sept. 13, 2018)

Lenci, A., Zampolli, A. & Calzolari, N. (2015). *International standards for multilingual resource sharing*. The ISLE Computational Lexicon Working Group, Universit di Pisa, Dipartimento di Linguistica.

*Lost for Words* (2017). Translators Without Borders.

Lynch, C. A. (2003). *Institutional repository: essential infrastructure for scholarship in the digital age*. Retrieved April 25, 2018, from http://www.arl.org/resources/pubs/br/br226/br226ir.shtml

Makori, E.O. and Mauti, N.O. (2016). "Digital technology acceptance in transformation of university libraries and higher education institutions in kenya". *Library Philosophy and Practice (ejournal). 1379*. http://digitalcommons.unl.edu/libphilprac/1379 (Retrieved Sept. 5, 2018)

Mitchell, J. S. et al (ed), *Dewey Decimal Classification.* 21st ed. New York, Forest Press, 1996 pp lxix.

Munday, Jeremy (2012). *Introducing translation studies: theories and applications*, 3rd ed. London: Routledge.

Mutte, Jean-Louis (2010). "*Connaissances ou compétences: que faut-il privilégier et gérer?*" Communication presented at the 13th CRHA Annual Conference (Montréal, 27 Sept. 2010).

Nakayama, K., Hara, T. & Nishio, S. (2008). *Wikipedia Mining*. Department of Multimedia Engineering, Graduate School of Information Science and Technology.

Nonaka, Ikujiro (1994). "A dynamic theory of organizational knowledge creation." *Organization Science 5*(1), 14-37.

Nord, Christiane. (2007). *Translating as a Purposeful Activity: Functionalist Approaches Explained*. Manchester: St. Jerome Publishing.

Nzomo, P. I., "*Multilingual Information Access: Practices and Perceptions of Bi/multilingual Academic Users*" (2015). Electronic Thesis and Dissertation Repository. 2812. Viewed May 8, 2018, from https://ir.lib.uwo.ca/etd/2812.

Palumbo, Giuseppe. (2009). *Key Terms in Translation Studies*. London: Continuum.

Piwowar, H.A. & Vision, T.J. (2013). *Data reuse and the open data citation advantage*. PeerJ, 1: e175. Viewed June 8, 2018, from https://doi.org/10.7717/peerj.175

Reese, R. (2015). *Natural Language Processing with Java*. Packt Publishing Ltd.

Saldanha, G. and O'Brien, S. (2014). *Research Methodologies in Translation Studies*. New York: Routledge.

Saraiva, R., & Rodrigues, E. (2010). Open access in Portugal. In L. Anglada & E. Abadal (eds.), *Open access in southern European countries* (pp.83–99). Madrid: FECYT. Retrieved May 6, 2018 from http://www.accesoabierto.net/sites/accesoabierto.net/files/OASouthEurope.pdf.

Sfetcu, N. (2014). *Google products, services and tools*.

Shah, R., Lin, B., Gershman, A. & Frederking, R. (2010). Synergy: A Named-Entity-Recognition system for resource-scarce languages such as Swahili using online machine translation. In *Proceedings of the Second Workshop on African Language Technology (AfLaT)* (pp. 21–26).

Smith, D. A. (2006). Debabelizing Libraries: Machine Translation by and for Digital Collections. *D-Lib Magazine 12*(3).

Solano-Flores, G., Backhoff, E., & Contreras-Niño, L. Á. (2009). Theory of test translation error. *International Journal of Testing*, *9*(2), 78-91. Viewed August 14, 2018, from http://dx.doi.org/10.1080/15305050902880835

Tatsumi, Midori (2009). "Correlation between Automatic Evaluation Metric Scores, Post-Editing Speed, and Some Other Factors." *MT Summit XII: Proceedings of the twelfth Machine Translation Summit*, (Ottawa, Canada, August 2009), 332-339. Viewed August 27, 2018, from http://www.mt-archive.info/MTS-2009-Tatsumi.pdf.

TAUS (2010). *Machine Translation Post-editing Guidelines*. Viewed September 9, 2018, from https://evaluation.taus.net/resources/guidelines/post-editing/machine-translation-post-editing-guidelines.

Temnikova, Irina (2010). "Cognitive Evaluation Approach for a Controlled Language Post-Editing Experiment." *LREC 2010: Proceedings of the seventh international conference on Language Resources and Evaluation*, (Valletta, Malta, May 2010), 3485-3490. Viewed September 12, 2018, from http://www.lrec-conf.org/proceedings/lrec2010/pdf/437_Paper.pdf.

Thicke, Lori (2013). "The industrial process for quality machine translation." *The Journal of Specialised Translation 19*, p. 8-18. Viewed August 27, 2018, from http://www.jostrans.org/issue19/art_thicke.pdf.

Toury, G. (1995). *Descriptive translation studies and beyond*. Amsterdam-Philadelphia: John Benjamins.

United Nations (2010). *The Paris Declaration*. Viewed July 15, 2018, from http://ec.europa.eu/dgs/scic/docs/paris_decl_en.pdf .

*University of Nairobi Library Services strategic plan, 2005-2010*. Nairobi: University of Nairobi.

*University of Nairobi Library System brochure* (2012). Nairobi: University of Nairobi.

UoN (2008) *University of Nairobi Strategic Plan 2008-2013*. Nairobi: University of Nairobi.

Van Doorslaer, L. (2007), 'Risking Conceptual Maps: Mapping as a keywords-related tool underlying the online Translation Studies Bibliography', *Target, 19*(2), p. 217-233.

Vauquois, Bernard. 1968. A survey of formal grammars and algorithms for recognition and transformation in mechanical translation. In *Proceedings of IFIP Congress*, 1114-1122. Edinburgh.

Waddington, C. (2001). Different methods of evaluating student translation: the question of validity. *Meta, XLVI*, p. 2.

Ware, M. (2004). *Pathfinder research on web-based repositories*. Viewed April 14, 2018, from Available at: http://mrkwr.files.wordpress.com/2006/11/pals-report-on-institutional-repositories.pdf

www.webometrics.info (Accessed September 1, 2018)

Zaman, M., Shamim, R. & Clement K. (2011). Trends and issues to integrate ICT in teaching and learning for the future world of education. *International Journal of Engineering & Technology, 11*(3): pp. 114-119.

# APPENDICES

## Appendix A:
## Cultural Text 1

**(Thesis: Effects of work-family conflict on job and life satisfaction among staff of The State Department Of Coordination, Ministry Of Interior And Coordination Of National Government by Kiunga, Martha R. 2017)**

Work-family conflict has adverse effects on behaviour, emotions, and health. Underestimating the importance of helping workers to reduce the impacts, to date, few authors have examined work-family conflict among employees in public sector. This research, therefore, aimed to examine the association between work-family conflict and job and life satisfactions, as well as how gender influenced the associations. The research employed descriptive research design to examine the work-family conflict among the employees of the State Department of Coordination, Ministry of Interior and Coordination of National Government. Data was collected using questionnaires targeting different employee levels in the ministry. The researcher used chi-square tests and analysis of variance (ANOVA) to obtain the relationship between work-family conflict, and job and life satisfaction. The researcher also compared men to women individual experiences at different levels of work relating to the work-family conflict, and job and life satisfactions. Current research helped to understand work and family roles including working women with children in job and life satisfaction, giving a consolidated overview of the current work and family theories within a conceptual and unifying manner.

**Google Translate**

Migogoro ya kazi ya familia ina athari mbaya juu ya tabia, hisia, na afya. Ukizingatia umuhimu wa kuwasaidia wafanyakazi kupunguza madhara, hadi leo, waandishi wachache wamechunguza migogoro ya kazi na familia kati ya wafanyakazi katika sekta ya umma. Utafiti huu, kwa hiyo, ulikuwa na lengo la kuchunguza ushirikiano kati ya migogoro ya kazi-familia na ufanisi wa kazi na maisha, pamoja na jinsi jinsia ilivyoathiri vyama. Utafiti uliofanywa utafiti unaoelezea utafiti kuchunguza migogoro ya kazi na familia kati ya wafanyakazi wa Idara ya Udhibiti wa Nchi, Wizara ya Mambo ya Ndani na Uratibu wa Serikali ya Taifa. Takwimu zilikusanywa kwa kutumia maswali yaliyoelezea viwango vya mfanyakazi tofauti katika huduma. Mtafiti alitumia vipimo vya ki-mraba na uchambuzi wa tofauti (ANOVA) ili kupata uhusiano kati ya migogoro ya kazi na familia, na kuridhika kazi na maisha. Mtafiti pia alilinganisha wanaume na uzoefu wa wanawake binafsi katika viwango tofauti vya kazi zinazohusiana na migogoro ya kazi-familia, na kazi na fikira za maisha. Utafiti wa sasa ulisaidia kuelewa majukumu ya kazi na familia ikiwa ni pamoja na kazi za wanawake na watoto katika kuridhika kazi na maisha, kutoa maelezo ya jumla ya kazi na sasa ya familia katika nadharia na umoja**.**

**Microsoft Bing Translator**

Migogoro ya kazi-familia ina athari hasi kwa tabia, hisia na afya. Harakati ya umuhimu wa kusaidia wafanyakazi ili kupunguza athari, hadi sasa, waandishi wachache na kuchunguza kazi-familia mgogoro miongoni mwa wafanyakazi katika sekta ya umma. Utafiti huu, kwa hiyo, lengo la kuchunguza chama kati ya migogoro ya kazi-familia na satisfactions wa kazi na maisha, kama vile jinsi jinsia kusukumwa vyama. Utafiti walioajiriwa utafiti fafanuzi Sanifu kuchunguza migogoro ya kazi-familia miongoni mwa wafanyakazi wa Idara ya hali ya uratibu, Wizara ya mambo ya ndani na uratibu wa serikali. Data zimekusanywa kutumia hojaji kuwalenga ngazi tofauti ya wafanyakazi katika Wizara. Mtafiti alitumia majaribio ya chi-square na uchambuzi wa tofauti (ANOVA) kupata uhusiano kati ya migogoro ya kazi na familia, na kuridhika ya kazi na maisha. Mtafiti pia ikilinganishwa wanaume kwa wanawake binafsi uzoefu katika ngazi tofauti za kazi zinazohusiana na kazi-familia mgogoro, na kazi na maisha satisfactions. Utafiti wa sasa alimsaidia kuelewa kazi na majukumu ya familia ikiwemo wanawake kufanya kazi na watoto katika kazi na maisha ya kuridhika, kutoa maelezo ya jumla ya masaa ya kazi sasa na nadharia familia ndani ya njia ya dhana na unganishi.

**Human Translator**

Migogoro ya kazi-familia ina athari hasi kwa tabia, hisia na afya. Kwa kutozingatia umuhimu wa kuwasaidia wafanyakazi kupunguza madhara, hadi leo, ni waandishi wachache wamechunguza migogoro ya kazi-familia miongoni mwa wafanyakazi katika sekta ya umma. Utafiti huu, kwa hivyo, ulikuwa na lengo la kuchunguza ushirikiano kati ya migogoro ya kazi-familia na ufanisi wa kazi na maisha, pamoja na jinsi jinsia ilivyoathiri vyama. Utafiti ulitumia mbinu ya utafiti fafanuzi kuchunguza migogoro ya kazi-familia miongoni mwa wafanyakazi wa Idara ya Udhibiti wa Nchi, Wizara ya Mambo ya Ndani na Uratibu wa Serikali ya Taifa. Takwimu zilikusanywa kwa kutumia hojaji zilizowalenga wafanyikazi wa ngazi tofauti tofauti katika wizara. Mtafiti alitumia vipimo vya ki-mraba na uchambuzi wa tofauti (ANOVA) ili kupata uhusiano kati ya migogoro ya kazi-familia, na kuridhika kikazi na kimaisha. Mtafiti pia alilinganisha uzoefu wa binafsi kati ya wanaume na wanawake katika viwango tofauti tofauti vya kazi zinazohusiana na migogoro ya kazi-familia, na ridhaa ya kazi na maisha. Utafiti wa sasa ulisaidia kuelewa majukumu ya kazi na familia ikiwa ni pamoja na wanawake wafanya kazi na watoto katika kazi na maisha ya kuridhisha, kutoa maelezo ya jumla ya kazi ya sasa na nadharia ya familia ndani ya njia ya dhana na unganishi.

# Appendix B: Cultural Text 2

**(Thesis: Causes and consequences of domestic violence against men in Mukurwe-Ini constituency, Kenya by Njuguna, Joseph K. 2014)**

This research project describes the work undertaken in Mukurwe-ini concerning Domestic Violence Against Men (DVM). The research was carried out in October 2014. The research question explored the causes of physical violence against men in Mukurwe-ini and investigated the consequences of violence against men in the study area from the victims of female perpetrated domestic violence with the focus on understanding the surrounding issues of society, local authority and masculinity. From the in depth qualitative interviews conducted, the findings conclude that the help and support available for male victims is virtually nonexistent and that which exist is of poor quality. The study interviewed Key informants who were the Deputy County commissioner and the OCPD in Mukurwe-ini. Quantitative data collected from closed ended questions was analyzed by the use Statistical Package for Social Sciences (SPSS Version 21) and the data presented by use of descriptive statistics such as percentages, means, standard deviations and frequencies. The findings and literature in this area suggest the underlying reason for why men are battered, due to patriarchal society and the relevant authorities' not actively recognizing domestic violence to be a male issue as well as a female issue, their only focus is towards female victims. Therefore male victims remain invisible. This societal repression has created stereotypes that impact upon the way authorities deal with domestic violence cases, in that they tend to discriminate against male victims. These stereotypes and discrimination cause male victims to feel reluctant to seek help. So, in turn, it seems domestic violence towards men is not a widespread problem hence there is no pressure for change. However this is not the case. This research piece develops these arguments fully to provide insight into the taboo subject of domestic violence against males.

**Google Translate**

Mradi huu wa utafiti unaelezea kazi iliyofanyika katika Mukurwe-ini kuhusu Ukatili wa Dhidi ya Wanaume (DVM). Utafiti ulifanyika mnamo Oktoba 2014. Swali la utafiti lilibainisha sababu za unyanyasaji wa kimwili dhidi ya wanaume huko Mukurwe-ini na kuchunguza matokeo ya unyanyasaji dhidi ya wanaume katika eneo la utafiti kutoka kwa waathirika wa unyanyasaji wa kike wa ndani wa kike na kuzingatia kuelewa masuala yanayozunguka jamii, mamlaka za mitaa na masculinity. Kutoka kwa kina mahojiano ya ubora yaliyofanywa, matokeo hayo yanahitimisha kwamba usaidizi na msaada unaopatikana kwa waathirika wa kiume ni karibu haipo na kile kilichopo ni cha ubora duni. Utafiti huo uliohojiwa na waandishi wa habari muhimu ambao walikuwa Naibu Kamishna wa Wilaya na OCPD katika Mukurwe-ini. Dalili za takwimu zilizokusanywa kutoka kwa maswali yaliyofungwa imefungwa na kutumia Package ya Takwimu za Sayansi za Jamii (SPSS Version 21) na data iliyowasilishwa kwa kutumia takwimu zinazoelezea kama vile asilimia, inamaanisha, kupoteza kiwango na frequency. Matokeo na vichapo katika eneo hili zinaonyesha sababu ya msingi ya kwa nini wanaume wanakabiliwa, kutokana na jamii ya patriar na mamlaka husika si kutambua kikamilifu unyanyasaji wa nyumbani kuwa suala la kiume na suala la kike, lengo lao pekee ni kwa waathirika wa kike. Kwa hiyo waathirika wa kiume hubakia asiyeonekana. Ukandamizaji huu wa kijamii umesababisha maoni ambayo yanayoathiri mamlaka ya kukabiliana na kesi za unyanyasaji wa nyumbani, kwa kuwa huwa na ubaguzi dhidi ya waathirika wa kiume. Ubaguzi huu na ubaguzi husababisha waathirika wa kiume kujisikia kusita kutafuta msaada. Kwa hiyo, inaonekana kuwa unyanyasaji wa ndani kwa wanaume sio tatizo la kuenea kwa hiyo hakuna shinikizo la mabadiliko. Hata hivyo hii sio kesi. Kipande hiki cha utafiti kinaendelea hoja hizi kwa kikamilifu kutoa ufahamu katika suala la uhalifu wa ndani dhidi ya wanaume.

**Microsoft Bing Translator**

Mradi huu utafiti inaeleza kazi inayoendeshwa Mukurwe-ini kuhusu ukatili wa majumbani dhidi ya watu (DVM). Utafiti ulifanyika katika Oktoba 2014. Swali la utafiti kutalii visababishi vya unyanyasaji wa kimwili watu katika Mukurwe-ini na kuchunguzwa na matokeo ya vurugu dhidi ya watu katika eneo la kujifunza kutoka kwa waathirika wa kike unaofanywa ndani vurugu na lengo juu ya ufahamu wa jirani masuala ya jamii, mamlaka za mitaa na kiume. Kutoka kina katika mahojiano ubora unafanywa, matokeo ya kuhitimisha kwamba msaada na Tegemeza iliyopo kwa ajili ya waathirika wa kiume ni watatoweka karibu na kwamba ambayo kuwepo ni duni. Utafiti ilimhoji Watakaotoa taarifa muhimu waliokuwa Kamishna wa Kaunti ya Naibu na OCPD KUWEZA katika Mukurwe-ini. Kiasi takwimu kutoka maswali funge ukaisha alikuwa kuchambuliwa kwa kutumia Kifurushi cha takwimu kwa ajili ya sayansi ya jamii (SPSS toleo la 21) na data kilichowasilishwa na matumizi ya takwimu fafanuzi kama vile asilimia, njia, kiwango kupotoka na masafa. Matokeo na fasihi katika eneo hili kupendekeza sababu ya msingi kwa nini watu ni iliyoonekana, kutokana na jamii za baba mkuu na mamlaka husika si kikamilifu kutambua ukatili wa majumbani kuwa suala la kiume kama vile suala la kike, dhamira yao pekee ni kuelekea waathirika wa kike. Kwa hivyo waathirika wa kiume kubaki asiyeonekana. Ukandamizaji huu jamii imeunda ubaguzi kwamba athari juu ya njia mamlaka kushughulikia kesi za ukatili wa majumbani, kwa kuwa wao huwa zina ubaguzi dhidi ya waathirika wa kiume. Haya mambo ya kuiga na ubaguzi kusababisha waathirika kiume kuhisi wanasita kutafuta msaada. Hivyo, kwa upande, inaonekana ndani vurugu kuelekea watu sio tatizo sana hivyo kuna hakuna shinikizo kwa ajili ya mabadiliko. Hata hivyo hii si kesi. Kipande hii utafiti yanaendelea hoja hizi kikamilifu kutoa ufahamu katika mada ya mwiko wa ukatili wa majumbani dhidi ya wanaume.

**Human Translator**

Mradi huu wa utafiti unaelezea kazi iliyofanyika katika Mukurwe-ini kuhusu Ukatili Dhidi ya Wanaume (DVM). Utafiti ulifanyika mnamo Oktoba 2014. Swali la utafiti lilibainisha sababu za unyanyasaji wa kimwili dhidi ya wanaume huko Mukurwe-ini na kuchunguza matokeo ya unyanyasaji dhidi ya wanaume katika eneo hilo kutoka kwa waathirika wa unyanyasaji wa kike na kuzingatia kuelewa masuala yanayozunguka jamii, mamlaka za mitaa na uume. Kutoka kwa mahojiano ya kina ya ubora yaliyofanywa, matokeo hayo yanahitimisha kwamba usaidizi na msaada unaopatikana kwa waathirika wa kiume ni kama kwamba haupo na ule ambao upo ni duni. Utafiti huo ulihoji watoa habari muhimu ambao walikuwa naibu kamishna wa wilaya na mkuu wa polisi (OCPD) katika Mukurwe-ini. Takwimu zilizokusanywa kutoka kwa maswali zilichambuliwa kwa kutumia kifurushi cha takwimu kwa ajili ya sayansi ya jamii (SPSS toleo la 21)) na habari iliyowasilishwa kwa kutumia takwimu fafanuzi kama vile asilimia, kiwango kupotoka na masafa. Matokeo na maandishi katika eneo hili zinaonyesha sababu ya msingi ya wanaume kukabiliwa, kutokana na jamii ya ukoo na mamlaka husika hawatambui kikamilifu unyanyasaji wa nyumbani kuwa suala la kiume na suala la kike, lengo lao pekee ni kwa waathirika wa kike. Kwa hiyo waathirika wa kiume hubakia

kutoonekana. Ukandamizaji huu wa kijamii umesababisha maoni ambayo yanaathiri mamlaka ya kukabiliana na kesi za unyanyasaji wa nyumbani, kwa kuwa huwa na ubaguzi dhidi ya waathirika wa kiume. Ubaguzi huu husababisha waathirika wa kiume kujisikia kusita kutafuta msaada. Kwa hivyo, inaonekana kuwa unyanyasaji wa nyumbani kwa wanaume sio tatizo la kuenea na hivyo hakuna shinikizo la mabadiliko. Hata hivyo hii sio kweli. Kipande hiki cha utafiti kinaendelea kutoa hoja hizi kwa kikamilifu na ufahamu katika mada ya mwiko wa ukatili majumbani dhidi ya wanaume.

# Appendix C: Cultural Text 3

**(Thesis: Heterogeneity and performance of spoken word in Kenya by Ekesa, Beatrice J. 2016)**

This study investigates the role of intertextuality and performance in the creation of meaning and aesthetic appeal in the spoken word poetry of fourteen Kenyan spoken word poets namely Brigeddia Poet, Dan Oballa, Dorphan, El-Poet, Elsaphan Njora, Imani woomera, Kennet B., Mufasa, Namatsi Lukoye, Ngwatilo Mawiyoo, Raya Wambui, Teardrops, Tess Aura, and Wanjiku Mwaura. The interrogation of the numerous ways in which spoken word texts create meaning through intertextuality focuses on the oral tradition, the novel and other forms of oral poetry like jazz, hip-hop and dub, while creation of meaning through performance takes into account the techniques used by the poets to enhance the interpretation of their poems. Thus, the research draws upon the theory of intertextuality to examine how the interpretation of spoken word relies on other texts, and how this contributes to the heterogeneity of this genre. The study also adopts the performance theory to establish how the poets manipulate the performance techniques and other theatrical elements to enhance the interpretation of their poems as well as add to the general aesthetic appeal. A close textual analysis of spoken word poetry reveals that the genre is diverse in nature since it incooporates other art forms and cannot be studied in isolation without paying attention to other texts embedded within it which play a major role in the interpretation of the text. However, in as much as spoken word constitutes other genres within it, it still retains its own unique features that seperate it from the rest. These unique features are mainly realised through performance, because it is the manner in which spoken word is articulated that distinguishes it from other forms of oral poetry.

**Google Translate**

Utafiti huu unachunguza jukumu la kuingiliana na utendaji katika kuundwa kwa maana na ushauri wa kupendeza katika mashairi ya maneno yaliyozungumzwa ya washairi wa maneno kumi na wanne wa Kenya, yaani Mchungaji wa Brigeddia, Dan Oballa, Dorphan, El-Poet, Elsaphan Njora, Imani woomera, Kennet B. , Mufasa, Namatsi Lukoye, Ngwatilo Mawiyoo, Raya Wambui, Teardrops, Tess Aura, na Wanjiku Mwaura. Kuhojiwa kwa njia mbalimbali ambazo maneno ya maneno yaliyozungumza husababisha maana kwa njia ya uingiliano unazingatia mila ya mdomo, riwaya na aina nyingine za mashairi ya mdomo kama jazz, hip-hop na dub, wakati uumbaji wa maana kupitia utendaji unazingatia mbinu zilizotumiwa na washairi kuimarisha tafsiri ya mashairi yao. Kwa hiyo, utafiti unakaribia nadharia ya kuingilia kati ili kuchunguza jinsi tafsiri ya neno lililosema inategemea kwenye maandiko mengine, na jinsi hii inachangia urithi wa aina hii. Utafiti pia unachukua nadharia ya utendaji ili kuanzisha jinsi washairi wanavyofanya mbinu za utendaji na vipengele vingine vya maonyesho ili kuongeza ufafanuzi wa mashairi yao pamoja na kuongeza kiti cha ujumla cha upasuaji. Uchunguzi wa karibu wa maneno ya mashairi yaliyozungumzwa unaonyesha kwamba aina hiyo ni tofauti kwa asili kwa sababu haiingii fomu za sanaa nyingine na haiwezi kujifunza kwa kutengwa bila kulipa kipaumbele kwa maandiko mengine yaliyomo ndani yake ambayo yana jukumu kubwa katika tafsiri ya maandiko. Hata hivyo, kwa vile vile neno lililozungumzwa linajenga aina nyingine ndani yake, bado ina sifa zake za kipekee ambazo zinazitoka kutoka kwa wengine. Makala haya ya pekee yanatambuliwa kwa njia ya utendaji, kwa sababu ni namna ambayo neno lililozungumzwa linaelezewa kuwa linafautisha kutoka kwa aina nyingine za mashairi ya mdomo.

**Microsoft Bing Translator**

Utafiti huu inachunguza jukumu la intertextuality na utendaji katika uumbaji wa maana na aesthetic kukata rufaa katika maneno ya mashairi ya washairi kumi na nne wa Kenya amesema neno yaani mshairi Brigeddia, Dan Oballa, Dorphan, El-mshairi, Elsaphan Njora, Imani woomera, B. Kennet, Mufasa, Namatsi Lukoye, Ngwatilo Mawiyoo, Raya Wambui, Teardrops, Tess Aura na Wanjiku Mwaura. Mahojiano ya njia mbalimbali katika neno ambalo amesema mafungu kuunda maana kupitia intertextuality hulenga mapokeo simulizi, riwaya na aina nyingine za mashairi simulizi kama jazz, hip-hop na Black, huku uumbaji wa maana kupitia utendaji katika akaunti mbinu zinazotumiwa na washairi wa kuongeza Tafsiri ya mashairi yao. Hivyo, utafiti huchota juu nadharia ya intertextuality kuchunguza jinsi Tafsiri ya maneno hutegemea mafungu mengine, na jinsi hii inachangia kwa heterogeneity ya Ghana hili. Utafiti pia antar nadharia ya utendaji kuanzisha jinsi washairi wa kuendesha mbinu za utendaji na elementi nyingine ushiriki wa jamii msingi kuongeza Tafsiri ya mashairi yao pamoja na kuongeza rufaa aesthetic mkuu. Karibu Ilidhihirika uchambuzi wa mashairi ya maneno hufunua kwamba tanzu ni tofauti katika asili tangu incooporates sanaa nyingine aina na haiwezi kusomwa katika kutengwa bila kuzingatia mafungu mengine iliyoingia ndani ambayo jukumu kubwa katika tafsiri ya matini. Hata hivyo, katika kama vile maneno kinachojumuisha tanzu nyingine ndani yake, ni bado hubakia makala yake mwenyewe kipekee kuwa tofauti na wengine. Vipengele hivi kipekee hasa walitambua kupitia utendaji, kwa sababu ni namna ambamo maneno ni imewekwa ambayo haitofautishi kutoka aina nyingine za ushairi simulizi.

**Human Translator**

Utafiti huu unachunguza jukumu la kuingiliana na utendaji katika uumbaji wa maana na ushauri wa kupendeza katika mashairi ya maneno yaliyozungumzwa ya washairi kumi na wanne wa Kenya ambao ni Brigeddia Poet, Dan Oballa, Dorphan, El-Poet, Elsaphan Njora, Imani woomera, Kennet B., Mufasa, Namatsi Lukoye, Ngwatilo Mawiyoo, Raya Wambui, Teardrops, Tess Aura na Wanjiku Mwaura. Mahojiano ya njia mbalimbali ambazo maneno yaliyozungumzwa husababisha maana kwa njia ya uingiliano unazingatia mila simulizi, riwaya na aina nyingine za mashairi simulizi kama jazi, huku uumbaji wa maana kupitia utendaji ukizingatia mbinu zinazotumiwa na washairi ili kuimarisha tafsiri ya mashairi yao. Hivyo, utafiti unatumia nadharia ya kuingilia kati ili kuchunguza jinsi tafsiri ya neno lililosemwa unategemea maandiko mengine, na jinsi huo inachangia urithi wa aina hii. Utafiti pia unachukua nadharia ya utendaji ili kuanzisha jinsi washairi wanavyofanya mbinu za utendaji na vipengele vingine vya maonyesho ili kuongeza ufafanuzi wa mashairi yao pamoja na kuongeza ujumi kwa jumla. Uchambuzi wa karibu wa maneno ya mashairi yaliyozungumzwa unaonyesha kwamba tanzu hiyo ni tofauti kwa asili kwa sababu haiingii fomu za sanaa nyingine na haiwezi kufunzwa kwa kutengwa bila kuzingatia mafungu mengine yaliyomo ndani yake ambayo yana jukumu kubwa katika tafsiri ya matini. Hata hivyo, kama vile neno lililozungumzwa linajenga aina nyingine ndani yake, bado lina sifa zake za kipekee ambazo zinatoka kutoka kwa wengine. Vipengele hivi vya kipekee hasa utambuliwa kupitia utendaji, kwa sababu ni namna ambavyo neno linavyoelezewa ambayo inalitofautisha kutoka kwa aina nyingine za ushairi simulizi.

# Appendix D: Scientific Text 1

**(Thesis: Genetic diversity of indigenous chicken (*gallus domesticus*) population in Kenya by Sinoya, Kevin S. 2017)**

This study was conducted with the aim of analyzing the diversity, relationship and population structure of the local indigenous chicken ecotypes in Kenya at the genetic level. Ecotypes from different regions were compared to identify regions which have been subjected to selection (selection signatures). A total of 384 free-ranging chicken sampled from eight counties in four geograpical regions (Western, North Rift, South Rift and Coast) in Kenya were genotyped using 12 microsatellite markers. Identification of signatures of selection was done using whole genome-resequencing data. The total number of alleles for all codominant data was 140, while the mean number of different alleles (NA) was 8.094 ± 0.516 and the effective number of alleles (NE) 4.452 ± 0.297. Observed (HO) and expected (HE) heterozygosities were 0.714 ± 0.011 and 0.726 ± 0.009 respectively for the whole population. The fixation indices – effects of sub-populations to the total population (FST), variance among subpopulations within groups (FIS), and variance among groups relative to total variance (FIT) for all populations were 0.029, 0.066 and 0.093 respectively. A total of 21 private alleles were observed in all populations. Bomet and West-Pokot ecotypes were closely related (0.997) than the rest of the population while the most distantly related were Lamu and Narok (0.779). The studied chicken population showed two clusters at K=8 when analyzed using STRUCTURE software. On identifying selection signatures, a total of 36,026 SNP variants were identified in these chicken populations with 30 Z-transformed outlier values ranging between 5 and 8 all in the 8th chromosome. Ninety nine percent (99%) of the variants were modifiers. Of the total biotype, 50.96 were in the non-coding region, 26.22 in the protein coding region. The consequences were expressed in upstream and downstream regions of the DNA, intron, intron non-coding transcript region and intergenic regions. The study concluded that with this rich genetic diversity, management and conservation measures should be undertaken for sustainable utilization of the indigenous chicken and reduce genetic dilution.

**Google Translate**

Utafiti huu ulifanyika kwa lengo la kuchambua tofauti, uhusiano na idadi ya watu wa aina za asili za kuku za asili nchini Kenya katika kiwango cha maumbile. Ecotypes kutoka mikoa tofauti ikilinganishwa na kutambua mikoa ambayo yamewekwa chini ya uteuzi (ishara ya uteuzi). Jumla ya kuku 384 ya bure iliyopangwa kutoka kwa wilaya nane katika mikoa minne ya geograpical (Magharibi, Kaskazini Rift, Kusini Rift na Pwani) nchini Kenya zilikuwa zinajitokeza kwa kutumia alama 12 za microsat-ellite. Utambuzi wa saini za uteuzi ulifanyika kwa kutumia data kamili ya urekebishaji wa jenome. Idadi ya jumla ya takwimu zote zilizotawala ilikuwa 140, wakati namba ya maana ya alleles tofauti (NA) ilikuwa 8.094 ± 0.516 na nambari ya jumla ya alleles (NE) 4.452 ± 0.297. Kuzingatia (HO) na kutarajia (HE) heterozygosi-ties walikuwa 0.714 ± 0.011 na 0.726 ± 0.009 kwa mtiririko kwa idadi ya watu wote. Matukio ya kurekebisha - madhara ya idadi ndogo ya watu kwa jumla ya idadi ya watu (FST), tofauti kati ya idadi ndogo ya vikundi (FIS), na tofauti kati ya vikundi kuhusiana na tofauti ya jumla (FIT) ya watu wote walikuwa 0.029, 0.066 na 0.093 kwa mtiririko huo. Mikokoteni ya Bomet na Magharibi-Pokot yalihusiana sana (0.997) kuliko watu wengine wote wakati wa kizazi cha mbali zaidi walikuwa Lamu na Narok (0.779). Katika kutambua saini za kuchagua, jumla ya vipimo 36,026 vya SNP zilibainishwa katika watu hawa wa kuku na maadili ya nje ya Z zenye kati ya 5 na 8 yote katika chromosomu ya 8. Asilimia tisini tisa (99%) ya vipengee walikuwa marekebisho. Kati ya jumla ya biotype, 50.96 walikuwa katika kanda isiyokosa coding, 26.22 katika kanda ya coding kanda. Matokeo yalifanyika katika mikoa ya chini na ya chini ya DNA. Utafiti huo ulihitimisha kwamba kwa udhibiti huu wa maumbile, usimamizi na hatua za uhifadhi zinapaswa kufanyika kwa ajili ya matumizi endelevu ya kuku wa asili na kupunguza dilution ya maumbile.

**Microsoft Bing Translator**

Utafiti huu ulifanyika kwa lengo la kuchambua muundo utofauti, uhusiano na idadi ya kuku ndani ya asili eco-aina nchini Kenya katika ngazi ya kijenetiki. Ecotypes kutoka mikoa mbalimbali walikuwa ikilinganishwa na kutambua maeneo ambayo wanakabiliwa na uteuzi (uteuzi saini). Jumla ya 384 kuku bure-kuanzia sampled kutoka nane kaunti katika geograpical mikoa minne (Magharibi, North Rift, ufa wa Kusini na pwani) nchini Kenya walikuwa genotyped kutumia alama 12 wa microsat-ellite. Utambulisho wa saini ya uteuzi ilifanyika kutumia data nzima genome-resequencing. Jumla ya idadi ya alleles kwa data zote wa kubwa mwenza alikuwa 140, wakati idadi ya maana ya alleles tofauti (rangi) 8.094 ± 0.516 na ufanisi namba-kujaza ya alleles (NE) 4.452 ± 0.297. Aliona (HO) na inatarajiwa (yeye) heterozygosi-tai walikuwa ± ± 0.714 0.011 na 0.726 0.009 mtiririko huo kwa idadi ya watu wote. Fahirisi kuwabainishia – madhara ya idadi ndogo kwa jumla watu (FST), tofauti kati ya subpopulations ndani ya vikundi (FIS), na tofauti miongoni mwa makundi jamaa jumla tofauti (FIT) kwa ajili ya wote wa popula-tions walikuwa 0.029, 0.066 na 0.093 mtiririko huo. Ecotypes ya Bomet na Pokot Magharibi walikuwa uhusiano wa karibu (0.997) kuliko wananchi wengine wakati zaidi wa mbali relat-ed Lamu na Narok (0.779). Katika kutambua ziainishwa saini, jumla ya 36,026 SNP variants walikuwa kutambuliwa katika watu hawa kuku na thamani 30 wa kubadilishwa Z outlier kuanzia kati ya 5 na 8 yote katika shurutisho 8. Asilimia tisini na tisa (99%) ya variants na walikuwa modifiers. Biotype jumla, 50.96 walikuwa mkoani zisizo na mpangilio, 26.22 katika protini mpangilio wa mkoa. Matokeo ya waliotajwa katika mabwawa na kijito mikoa ya DNA. Ya kujifunza conclud-ed kuwa na utofauti huu tajiri ya kijenetiki, hatua za usimamizi na uhifadhi ichukuliwe kwa ajili ya matumizi endelevu ya kuku wa asili na kupunguza dilution kijenetiki.

**Human Translator**

Utafiti huu ulifanyika kwa lengo la kuchambua tofauti, uhusiano na muundo wa idadi ya aina za asili za kuku za kiasili nchini Kenya katika kiwango cha maumbile. Aina mbalimbali kutoka mikoa tofauti ililinganishwa na kutambua mikoa ambayo yamewekwa chini ya uteuzi. Jumla ya kuku huria 384 zilikuwa sampuli kielelezo kutoka kaunti nane katika maeneo manne ya (Magharibi, Bonde la Ufa Kaskazini, Bonde la Ufa Kusini na Pwani) nchini Kenya ziliwekwa alama 12 za kijenetiki. Utambuzi wa saini za uteuzi ulifanyika kwa kutumia habari kamili ya urekebishaji wa jenome. Jumla ya idadi ya takwimu zote zilizotawala ilikuwa 140, wakati namba ya maana ya alleles tofauti (NA) ilikuwa 8.094 ± 0.516 na nambari ya jumla ya alleles (NE) 4.452 ± 0.297. Kuzingatia (HO) na kutarajia (HE) heterozygosi-ties walikuwa 0.714 ± 0.011 na 0.726 ± 0.009 kwa mtiririko kwa idadi ya watu wote. Matukio ya kurekebisha - madhara ya idadi ndogo ya kuku kwa jumla ya idadi ya kuku (FST), tofauti kati ya idadi ndogo ya vikundi (FIS), na tofauti kati ya vikundi

kuhusiana na tofauti ya jumla (FIT) ya kuku zote ilikuwa 0.029, 0.066 na 0.093 kwa mtiririko huo. Idadi ya Bomet na Pokot Magharibi ilihusiana sana (0.997) kuliko kuku zingine zote na wakati huo kuku wa kizazi cha mbali zaidi walikuwa Lamu na Narok (0.779). Katika kutambua saini za kuchagua, jumla ya vipimo 36,026 vya SNP zilibainishwa katika idadi ya kuku na maadili ya nje ya Z zenye kati ya 5 na 8 yote katika chromosomu ya 8. Asilimia tisini tisa (99%) ya vipengee zilikuwa za kurekebisha. Kati ya jumla ya 50.96 zilikuwa katika kanda iliyokosa alama, na 26.22 katika kanda ya alama. Matokeo yalibaini hali ya DNA. Utafiti huo ulihitimisha kwamba ili kudhibiti maumbile, usimamizi na hatua za uhifadhi zinapaswa kufanyika kwa ajili ya matumizi endelevu ya kuku wa asili na kupunguza kufifilisha kuku kijenetiki.

# Appendix E: Scientific Text 2

**(Thesis: The urban block as a tool for urban design by Njagi, Rael W. 2015)**

The urban block is the link between city user's everyday lives and the urban space. This study calls for this urban component to be treated as an essential element of the city when new urban plans and models are developed. The aim of this study was to investigate how evolution of the urban block has influenced design of urban spaces, in order to develop principles that could be adopted for Parklands, Nairobi where the urban block is rapidly changing. Examples of urban projects where the urban block was a key element of design were evaluated, drawn from the traditional city to the post-modernist city. Desk research was used to study urban blocks outside Nairobi and as an introductory strategy to form basis for analysing the Parklands urban blocks in context. Typo-morphological approach was taken to understand the physical and spatial structures in different urban blocks in Parklands guided by a case study protocol. In order to analyse the process of transformation, urban blocks that could be paradigms of subsequent historical periods were selected. The study revealed that rapid urbanization and policy change directed the new urban form in Parklands and therefore urban blocks were constantly changing either to accommodate commercial function or higher residential densities. Buildings and the related open spaces are seen as complimentary units of space by the users and therefore changes in individual blocks translate to changes in the urban fabric whether guided by design or not. This study therefore makes a case for the adoption of the urban block as an intermediary level of planning and design and using it to achieve desired urban form.

**Google Translate**

Vikwazo vya mijini ni kiungo kati ya maisha ya kila siku ya mtumiaji wa jiji na nafasi ya mijini. Utafiti huu unaomba kwamba sehemu hii ya mijini ipatikane kama kipengele muhimu cha jiji wakati mipango mipya na mifano ya mijini inapatikana. Lengo la utafiti huu lilikuwa kuchunguza jinsi mageuzi ya miji ya mijini imesababisha mipangilio ya maeneo ya miji, ili kuendeleza kanuni ambazo zinaweza kupitishwa kwa Parklands, Nairobi ambako block ya mijini inabadilika haraka. Mifano ya miradi ya mijini ambapo block ya miji ilikuwa kipengele muhimu cha kubuni ilipimwa, inayotokana na mji wa jadi hadi mji wa baada ya kisasa. Utafiti wa dawati ulitumika kujifunza vitalu vya mijini nje ya Nairobi na kama mkakati wa utangulizi wa kuunda msingi wa kuchunguza vitalu vya miji ya Parklands katika mazingira. Njia ya maumbile ya kimaumbile ilichukuliwa ili kuelewa miundo ya kimwili na anga katika vitalu tofauti vya mijini huko Parklands inayoongozwa na itifaki ya utafiti wa kesi. Ili kuchambua mchakato wa mabadiliko, vitalu vya mijini ambavyo vinaweza kuwa vielelezo vya vipindi vya kihistoria vilivyochaguliwa. Utafiti huo umebaini kuwa miji ya haraka na mabadiliko ya sera zilielezea fomu mpya ya mijini huko Parklands na hivyo vitalu vya mijini vilibadilika kila mara ama kuwezesha kazi ya kibiashara au dalili za juu za makazi. Majengo na maeneo yanayohusiana yanaonekana kama vitengo vyema vya nafasi na watumiaji na kwa hiyo mabadiliko katika vitalu vya mtu binafsi hutafsiri mabadiliko ya kitambaa cha mijini ikiwa imeongozwa na kubuni au la. Kwa hiyo utafiti huu hufanya kesi kwa kupitishwa kwa kuzuia mijini kama ngazi ya mpatanishi wa kupanga na kubuni na kuitumia ili kufikia fomu ya miji inayofaa.

**Microsoft Bing Translator**

Umbo la mjini ni kiungo baina ya maisha ya kila siku ya mji mtumiaji na nafasi mijini. Utafiti huu wito wa kijenzi mjini kutibiwa kama kipengele muhimu cha mji wakati mipango ya mpya ya miji na mifano ni maendeleo. Lengo la Utafiti huu lilikuwa kuchunguza jinsi mageuzi ya umbo la mjini Imebadilisha muundo wa mazingira mijini, ili kuendeleza kanuni ambazo inaweza kuchukuliwa kwa ajili ya jijini Nairobi ambapo umbo la mjini inabadilika kwa kasi. Mifano ya miradi mijini ambapo umbo la mjini ilikuwa kipengele muhimu ya kubuni walikuwa tathmini, inatoka mji jadi mji modernist baada. Dawati utafiti kutumika kujifunza vitalu mijini nje ya Nairobi na kama mkakati wa utangulizi kwa msingi wa fomu kwa ajili ya kuchambua vitalu ya mijini ya Parklands alipokuwa anaiba katika muktadha. Njia kosa morphological alipelekwa kuelewa kimwili na nafasi miundo katika tofauti mjini huzuia katika Parklands kuongozwa na itifaki Uchunguzi kifani. Ili kuchambua mchakato wa mageuzi, vitalu mijini ambazo zinaweza msimano wa vipindi wa kihistoria wa baadae walichaguliwa. Utafiti ulionesha kwamba mabadiliko haraka ya ukuaji wa miji na sera kwa madhumuni ya fomu mpya ya mijini katika Parklands na kwa hivyo vitalu mijini walikuwa kubadilika ama kwa ajili ya malazi kazi kibiashara au wingi juu ya makazi. Majengo na nafasi wazi husika ni kuonekana kama vitengo bure ya nafasi kwa watumiaji na kwa hivyo mabadiliko katika vitalu binafsi kutafsiri mabadiliko katika kitambaa mijini kama kuongozwa na kubuni au la. Utafiti huu kwa hivyo hufanya kesi kwa ajili ya kupitishwa kwa umbo la mijini kama kiwango cha mpatanishi wa kupanga na kubuni na kutumia kufikia fomu taka mijini.

**Human Translator**

Umbo la mjini ni kiungo kati ya maisha ya kila siku ya mtumiaji wa jiji na nafasi ya mijini. Utafiti huu unatoa wito kwamba sehemu hii ya mijini ipatikane kama kipengele muhimu cha jiji wakati mipango mipya na mifano ya mijini inapatikana. Lengo la utafiti huu lilikuwa kuchunguza jinsi mageuzi ya miji ya mijini imesababisha mipangilio ya maeneo ya miji, ili kuendeleza kanuni ambazo zinaweza kupitishwa kwa Parklands, Nairobi ambako block ya mijini inabadilika haraka. Mifano ya miradi ya mijini ambapo block ya miji ilikuwa kipengele muhimu cha kubuni ilipimwa, inayotokana na mji wa jadi hadi mji wa baada ya kisasa. Utafiti wa dawati ulitumika kujifunza vitalu vya mijini nje ya Nairobi na kama mkakati wa utangulizi wa kuunda msingi wa kuchunguza vitalu vya miji ya Parklands katika mazingira. Njia ya maumbile ya kimaumbile ilichukuliwa ili kuelewa miundo ya kimwili na anga katika vitalu tofauti vya mijini huko Parklands inayoongozwa na itifaki ya utafiti wa kesi. Ili kuchambua mchakato wa mabadiliko, vitalu vya mijini ambavyo vinaweza kuwa vielelezo vya vipindi vya kihistoria vilivyochaguliwa. Utafiti huo umebaini kuwa miji ya haraka na mabadiliko ya sera zilielezea fomu mpya ya mijini huko Parklands na hivyo vitalu vya mijini vilibadilika kila mara ama kuwezesha kazi ya kibiashara au dalili za juu za makazi. Majengo na maeneo yanayohusiana yanaonekana kama vitengo vyema vya nafasi na watumiaji na kwa hiyo mabadiliko katika vitalu vya mtu binafsi hutafsiri mabadiliko ya kitambaa cha mijini ikiwa imeongozwa na kubuni au la. Kwa hiyo utafiti huu hufanya kesi kwa kupitishwa kwa kuzuia mijini kama ngazi ya mpatanishi wa kupanga na kubuni na kuitumia ili kufikia fomu ya miji inayofaa.

# Appendix F: Scientific Text 3

**(Thesis: Dental age estimation in children attending a university dental hospital by Kihara, Eunice 2016)**

There are various circumstances when the chronological age of a child needs to be verified but due to lack of legally approved documentary evidence, age has to be conventionally estimated. Determination of age provides an important biological data which plays a vital role in the identification of victims of mass disasters and crimes. Additionally, it is an important requirement during school admission, marriage, child adoption, in determination of criminal responsibility, imprisonment of minors and in management of various orthodontic as well as pediatric conditions and pathologies. Various methods have been applied to estimate the age of children mainly through the assessment of developmental and morphological changes of bone and teeth. However, in Kenya there is hardly any approved method that can be used to achieve this purpose. Hence, the need to determine the applicability of the available methods in estimating the age of a given Kenyan population. Objective: To determine the performance of Willems' model of age estimation in children visiting the University of Nairobi Dental Hospital. Materials and method: A cross sectional study was done at The University of Nairobi Dental Hospital. It involved examination of panoramic radiographs of 401 children aged 3.0 – 16.99 years old who had previously visited the pediatric and orthodontic clinic. The sample was divided into one year age cohorts from age 3 – 16 years. The radiographs were assessed in order to determine the tooth maturity stages for the first 7 mandibular teeth. In addition, each maturity stage was assigned a corresponding score as described in Willems' tables. The age difference was calculated through subtraction of the dental age from the chronological age. Descriptive and inferential statistics were done using SPSS 20.0 and presented in tables and various types of figures. Results: A sample of 401 radiographs was included in the study, 188 (46.9) belonged to females while 213 (53.1%) were for males. The mean chronological age was 9.73±3.60 years. The mean chronological age for girls and boys was 9.83±3.65 and 9.65±3.56 years respectively. The overall dental age was 10.01±3.60 years. Willems' method overestimated overall age by -0.27±1.30 years. There was statistically significant difference between the overall CA and DA, (t (400) = -4.185, p= 0.000). The 95% confidence interval was -0.40 - -0.14 years. Overall, there was a strong, positive correlation between the chronological and estimated age, r=0.935, n=401 p=0.000. The overall dental age for the girls was 9.93±3.60 years. The study found that the girls had a mean age difference of -0.10±1.37 which was not statistically significant, (t (187) = -1.017, p= 0.311) with 95% confidence interval of -0.30 - 0.10 years. The boys had an overall dental age of 10.07 ± 3.60 years. The mean age difference was -0.42±1.22 years which was statistically significant (t (212) = -5.041, p= 0.000) with a 95% confidence interval of -0.26 - -0.59 years. The most accurately estimated age by Willems' method was for age cohort 9 which had a mean age difference that was less than a month (-0.07 years). About a third (150, 37%) of the children had their age estimated within 6 months of their chronological age while about two thirds (258, 64%) were within one year. There was complete development of about 50% of the teeth which had already achieved the final maturity stage H. The youngest boy and girl with fully matured 7 mandibular teeth were aged 12.91 and 13.02 years respectively. In general, there was no statistical difference between the maturity for girls and boys in most tooth stages. However, girls were significantly ahead in crown development of the lateral incisor and root development of the 1st premolar and canine. The maturity of the children in the same age group revealed variations in tooth stages. Conclusion and recommendation: Use of Willems' method resulted in statistically significant overestimation of the age. The method performed better in estimating the age of the girls as compared to boys who were significantly over aged. Majority of the children had their age estimated within one year of their actual age. Generally, there was no statistical difference between the tooth maturity for girls and boys in most of the maturity stages. However, girls were significantly ahead of the boys in the root development of the canine. There existed different patterns of tooth maturity in children of the same age group. The current findings should be validated with a larger sample size that is representative of the Kenyan population. This will inform whether there is a need to modify Willems' method.

**Google Translate**

Kuna hali tofauti wakati umri wa mtoto unahitajika kuthibitishwa lakini kutokana na ukosefu wa ushahidi wa waraka ulioidhinishwa kisheria, umri unapaswa kuwa wastani wa wastani. Uamuzi wa umri hutoa data muhimu ya kibiolojia ambayo ina jukumu muhimu katika kutambua waathirika wa maafa ya masuala na uhalifu. Zaidi ya hayo, ni sharti muhimu wakati wa uandikishaji wa shule, ndoa, utunzaji wa watoto, uamuzi wa uhalifu wa kifungo, kifungo cha watoto na usimamizi wa hali mbalimbali za watoto na vilevile watoto. Njia mbalimbali zimetumika kwa kukadiria umri wa watoto hasa kupitia tathmini ya mabadiliko ya maendeleo na maumbile ya mfupa na meno. Hata hivyo, nchini Kenya hakuna njia yoyote inayoidhinishwa ambayo inaweza kutumika kufikia lengo hili. Kwa hiyo, haja ya kuamua uwezekano wa mbinu zilizopo katika kupima umri wa idadi ya watu wa Kenya. Lengo: Kuamua utendaji wa kiwango cha Willems 'wa umri wa makadirio katika watoto kutembelea Chuo Kikuu cha Nairobi Hospitali ya meno. Vifaa na mbinu: Utafiti wa sehemu ya msalaba ulifanyika katika Hospitali ya Madaktari ya Chuo Kikuu cha Nairobi. Ilihusisha uchunguzi wa radiographs za panoramic ya watoto 401 wenye umri wa miaka 3.0 - 16.99 ambao walitembelea kliniki ya watoto na orthodontic hapo awali. Sampuli ilikuwa imegawanywa katika makundi ya umri wa miaka moja kutoka umri wa miaka 3 hadi 16. Radiographs zilipimwa ili kuamua hatua za ukomavu wa jino kwa meno 7 ya kwanza ya mandibular. Aidha, kila hatua ya ukomavu ilipewa alama zinazofanana na ilivyoelezwa kwenye meza za Willems. Tofauti ya umri ulihesabiwa kwa njia ya kuondoa kwa umri wa meno kutokana na umri wa wakati. Takwimu zilizoelezea na zisizo za kutosha zilifanyika kwa kutumia SPSS 20.0 na zinawasilishwa katika meza na aina mbalimbali za takwimu. Matokeo: Sampuli ya radiographs 401 ilijumuishwa katika utafiti, 188 (46.9) walikuwa wa wanawake wakati 213 (53.1%) walikuwa kwa wanaume. Kiwango cha wakati wa wakati ulikuwa ni miaka 9.73 ± 3.60. Wakati wa maana wa wasichana na wavulana ulikuwa 9.83 ± 3.65 na 9.65 ± 3.56 miaka kwa mtiririko huo. Umri wa kawaida wa meno ulikuwa miaka 10.01 ± 3.60. Njia ya Willems 'ilipunguza umri wa jumla kwa miaka -0.27 ± 1.30. Kulikuwa na tofauti ya takwimu kati ya CA na DA, (t (400) = -4.185, p = 0.000). Kipindi cha ujasiri cha 95% kilikuwa -0.40 - -0.14 miaka. Kwa ujumla, kulikuwa na uwiano mzuri, unaofaa kati ya umri na kiwango cha wastani, r = 0.935, n = 401 p = 0.000. Urefu wa meno kwa wasichana ulikuwa miaka 993 ± 3.60. Utafiti huo uligundua kuwa wasichana walikuwa na umri wa maana tofauti ya -0.10 ± 1.37 ambayo haikuwa muhimu sana, (t (187) = -

1.017, p = 0.311) na kipindi cha ujasiri wa 95% -0.30 - 0.10. Wavulana walikuwa na umri wa meno kwa jumla ya miaka 10.07 ± 3.60. Tofauti ya umri wa umri ilikuwa -0.42 ± 1.22 miaka ambayo ilikuwa na takwimu muhimu (t (212) = -5.041, p = 0.000) na kipindi cha kujiamini kwa 95% ya miaka -0.26 - -0.59. Njia bora ya Willems 'ilikuwa sahihi kwa umri wa miaka 9 ambayo ilikuwa na tofauti ya umri wa miaka ambayo ilikuwa chini ya mwezi (-0.07 miaka). Kuhusu asilimia tatu (150, 37%) ya watoto walikuwa na wastani wa miaka yao ndani ya miezi 6 ya umri wao wa muda wakati karibu theluthi mbili (258, 64%) walikuwa ndani ya mwaka mmoja. Kulikuwa na maendeleo kamili ya asilimia 50 ya meno ambayo tayari yalifikia hatua ya mwisho ya ukomavu H. Mvulana mdogo zaidi na msichana aliye na meno 7 mandibular yaliyo na matunda yalikuwa na umri wa miaka 12.91 na 13.02 kwa mtiririko huo. Kwa ujumla, kulikuwa hakuna tofauti ya takwimu kati ya ukomavu kwa wasichana na wavulana katika hatua nyingi za jino. Hata hivyo, wasichana walikuwa na nguvu zaidi katika maendeleo ya taji ya incisor ya kuimarisha na maendeleo ya mizizi ya premolar 1 na canine. Ukomavu wa watoto katika kikundi hicho cha umri umeonyesha tofauti katika hatua za jino. Hitimisho na mapendekezo: Matumizi ya njia ya Willems ilisababishwa na upunguzaji wa umri wa takwimu. Njia hiyo ilifanya vizuri zaidi katika kupima umri wa wasichana ikilinganishwa na wavulana ambao walikuwa zaidi ya wazee. Wengi wa watoto walikuwa na wastani wa umri wao ndani ya mwaka mmoja wa umri wao halisi. Kwa ujumla, kulikuwa na tofauti ya takwimu kati ya ukuaji wa jino kwa wasichana na wavulana katika hatua nyingi za ukomavu. Hata hivyo, wasichana walikuwa wengi mbele ya wavulana katika maendeleo ya mizizi ya canine. Kulikuwa na mifumo tofauti ya ukomavu wa jino kwa watoto wa kikundi cha umri sawa. Matokeo ya sasa yanapaswa kuthibitishwa kwa ukubwa wa sampuli kubwa ambayo ni mwakilishi wa idadi ya watu wa Kenya. Hii itajulisha ikiwa kuna haja ya kurekebisha njia ya Willems.

**Microsoft Bing Translator**

Kuna hali mbalimbali wakati umri wa mtoto miaka chronological zinahitajika kuyakinishwa lakini kutokana na ukosefu wa ushahidi wa maandishi wa kupitishwa kisheria, umri ina kukadiriwa conventionally. Uamuzi wa umri hutoa deta ya muhimu ya kibiolojia ambayo ina jukumu muhimu katika utambulisho wa waathirika wa majanga ya halaiki na uhalifu. Zaidi ya hayo, ni sharti muhimu wakati wa uandikishaji shule, ndoa, kuasili mtoto, katika uamuzi wa jukumu ya jinai, kufungwa jela ya watoto na katika usimamizi wa hali ya orthodontic kama vile upasuaji na pathologies mbalimbali. Mbinu mbalimbali kutumika kukadiria umri wa watoto hasa kupitia tathmini ya mabadiliko ya maendeleo na morphological mifupa na meno. Hata hivyo, katika Kenya kuna njia yoyote vigumu kupitishwa ambayo inaweza kutumika kufikia lengo hili. Hivyo, haja ya kuamua utekelezaji wa mbinu zilizopo katika ukadiriaji wa umri wa idadi fulani ya Kenya. Lengo: Kuamua utendaji wa Willems' mfano wa hesabu ya umri watoto kutembelea Chuo Kikuu cha Nairobi meno hospitali. Vifaa na mbinu: A msalaba kujifunza sectional ilifanyika katika hospitali ya meno ya Chuo Kikuu cha Nairobi. Kushiriki uchunguzi wa radiographs panoramic 401 watoto wa kati ya umri wa miaka ya 3.0 – 16.99 ambaye alikuwa awali alitembelea kliniki ya upasuaji na orthodontic. Sampuli iligawanyika katika nyongeza ya umri wa mwaka mmoja kutoka wenye umri wa miaka 3 – 16. Radiographs kwenye overweight walikuwa tathmini ili kuamua hatua ya ukomavu ya jino kwa meno mandibular 7 kwanza. Aidha, kila hatua ya ukomavu ilikuwa imepangiwa alama sambamba kama ilivyoelezwa katika majedwali ya Willems'. Tofauti ya umri alikuwa mahesabu kupitia ya umri meno umri chronological. Takwimu za kimaelezo na inferential kufanyika kutumia SPSS 20.0 na iliyotolewa katika majedwali na aina mbalimbali za takwimu. Matokeo: Sampuli ya 401 radiographs alikuwa pamoja katika utafiti, 188 (46.9) ni wanawake wakati 213 (53.1%) walikuwa kwa ajili ya wanaume. Maana chronological umri miaka 9.73±3.60. Umri maana ya chronological kwa wasichana na wavulana ni 9.83±3.65 na 9.65±3.56 miaka kwa mtiririko huo. Umri meno jumla ilikuwa miaka ya 10.01±3.60. Mbinu ya Willems' overestimated umri wa jumla kwa - miaka 0.27±1.30. Kulikuwa na tofauti kubwa Kihesabu kati CA jumla na DA, (t (400) =-4.185, p = 0.000). Asilimia 95 imani nafasi alikuwa-0.40 --0.14 wa miaka. Kwa ujumla, kulikuwa na uwiano wa nguvu, chanya kati ya wendo na inakadiriwa umri, r = 0.935, n = 401 wa p = 0.000. Umri jumla ya meno kwa wasichana ilikuwa miaka ya 9.93±3.60. Utafiti uligundua kwamba wasichana walikuwa tofauti ya umri maana ya -0.10±1.37 ambayo ilikuwa si muhimu Kihesabu, (t (187) =-1.017, p = 0.311) na nafasi ya imani ya asilimia 95 ya-0.30 - miaka 0.10. Wavulana walikuwa umri meno jumla ya 10.07 ± miaka 3.60. Tofauti ya umri maana alikuwa - miaka ya 0.42±1.22 ambayo ilikuwa muhimu Kihesabu (t (212) =-5.041, p = 0.000) na muda wa imani asilimia 95 ya-0.26--0.59 miaka. Umri zaidi usahihi makadirio ya mbinu ya Willems' alikuwa kwa hutendeka katika umri wa miaka 9 ambayo ilikuwa tofauti ya umri maana kwamba alikuwa chini ya mwezi (-0.07 miaka). Kiasi cha thuluthi (150, 37%) ya watoto walikuwa umri wao inakadiriwa ndani ya miezi 6 ya umri wao chronological wakati karibu theluthi mbili (258, 64%) walikuwa ndani ya mwaka mmoja. Kulikuwa na maendeleo kamili ya asilimia 50 ya meno ambayo alikuwa tayari mafanikio hatua ya ukomavu mwisho H. Mvulana mdogo na msichana na kukomaa kikamilifu 7 mandibular meno walikuwa umri wa miaka 12.91 na 13.02 mtiririko huo. Kwa ujumla, kulikuwa hakuna tofauti ya takwimu kati ya ukomavu kwa ajili ya wasichana na wavulana katika hatua zaidi ya jino. Hata hivyo, wasichana walikuwa kwa kiasi kikubwa mbele katika maendeleo ya taji ya maendeleo pembeni incisor na mizizi ya 1 premolar na canine. Ukomavu wa watoto katika umri huo alifunua tofauti katika hatua ya jino. Hitimisho na mapendekezo: matumizi ya mbinu ya Willems' ulisababisha overestimation muhimu kihesabu ya umri. Mbinu ya kutumbuiza vyema katika ukadiriaji wa umri wa wasichana si kitu ukilinganisha na wavulana ambao walikuwa kwa kiasi kikubwa zaidi umri. Wengi wa watoto alikuwa umri wao inakadiriwa ndani ya mwaka mmoja ya umri wao halisi. Kwa ujumla, kulikuwa hakuna tofauti ya takwimu kati ya jino ukomavu kwa ajili ya wasichana na wavulana katika zaidi ya hatua ya ukomavu. Hata hivyo, wasichana walikuwa kwa kiasi kikubwa kabla ya wavulana katika ukuaji wa mizizi ya canine ya. Kuna kuwepo aina tofauti tofauti ya ukomavu wa meno kwa watoto wa umri huo huo. Matokeo sasa unapaswa kuhakikishwa na sampuli kubwa kwamba ni mwakilishi wa watu wa Kenya. Hii taarifa kama kuna haja ya kurekebisha mbinu ya Willems'.

**Human Translator**

Kuna hali mbalimbali wakati umri wa mtoto unahitajika kuthibitishwa lakini kutokana na ukosefu wa ushahidi wa hati iliyoidhinishwa kisheria, umri unapaswa kuwa wastani wa wastani. Uamuzi wa umri hutoa habari muhimu ya kibiolojia ambayo ina jukumu muhimu katika kutambua waathirika wa maafa ya masuala na uhalifu. Zaidi ya hayo, ni sharti muhimu wakati wa uandikishaji wa shule, ndoa, utunzaji wa watoto, uamuzi wa uhalifu wa

kifungo, kifungo cha watoto na usimamizi wa hali mbalimbali za watoto na vilevile watoto. Njia mbalimbali zimetumika kwa kukadiria umri wa watoto hasa kupitia tathmini ya mabadiliko ya maendeleo na maumbile ya mfupa na meno. Hata hivyo, nchini Kenya hakuna njia yoyote inayoidhinishwa ambayo inaweza kutumika kufikia lengo hili. Kwa hiyo, haja ya kuamua uwezekano wa mbinu zilizopo katika kupima umri wa idadi ya watu wa Kenya. Lengo: Kuamua utendaji wa kiwango cha Willems 'wa umri wa makadirio katika watoto kutembelea Chuo Kikuu cha Nairobi Hospitali ya meno. Vifaa na mbinu: Utafiti wa sehemu ya msalaba ulifanyika katika Hospitali ya Madaktari ya Chuo Kikuu cha Nairobi. Ilihusisha uchunguzi wa radiographs za panoramic ya watoto 401 wenye umri wa miaka 3.0 - 16.99 ambao walitembelea kliniki ya watoto na orthodontic hapo awali. Sampuli ilikuwa imegawanywa katika makundi ya umri wa miaka moja kutoka umri wa miaka 3 hadi 16. Radiographs zilipimwa ili kuamua hatua za ukomavu wa jino kwa meno 7 ya kwanza ya mandibular. Aidha, kila hatua ya ukomavu ilipewa alama zinazofanana na ilivyoelezwa kwenye meza za Willems. Tofauti ya umri ulihesabiwa kwa njia ya kuondoa kwa umri wa meno kutokana na umri wa wakati. Takwimu zilizoelezea na zisizo za kutosha zilifanyika kwa kutumia SPSS 20.0 na zinawasilishwa katika meza na aina mbalimbali za takwimu. Matokeo: Sampuli ya radiographs 401 ilijumuishwa katika utafiti, 188 (46.9) walikuwa wa wanawake wakati 213 (53.1%) walikuwa kwa wanaume. Kiwango cha wakati wa wakati ulikuwa ni miaka 9.73 ± 3.60. Wakati wa maana wa wasichana na wavulana ulikuwa 9.83 ± 3.65 na 9.65 ± 3.56 miaka kwa mtiririko huo. Umri wa kawaida wa meno ulikuwa miaka 10.01 ± 3.60. Njia ya Willems 'ilipunguza umri wa jumla kwa miaka -0.27 ± 1.30. Kulikuwa na tofauti ya takwimu kati ya CA na DA, (t (400) = -4.185, p = 0.000). Kipindi cha ujasiri cha 95% kilikuwa -0.40 - -0.14 miaka. Kwa ujumla, kulikuwa na uwiano mzuri, unaofaa kati ya umri na kiwango cha wastani, r = 0.935, n = 401 p = 0.000. Urefu wa meno kwa wasichana ulikuwa miaka 993 ± 3.60. Utafiti huo uligundua kuwa wasichana walikuwa na umri wa maana tofauti ya -0.10 ± 1.37 ambayo haikuwa muhimu sana, (t (187) = -1.017, p = 0.311) na kipindi cha ujasiri wa 95% -0.30 - 0.10. Wavulana walikuwa na umri wa meno kwa jumla ya miaka 10.07 ± 3.60. Tofauti ya umri wa umri ilikuwa -0.42 ± 1.22 miaka ambayo ilikuwa na takwimu muhimu (t (212) = -5.041, p = 0.000) na kipindi cha kujiamini kwa 95% ya miaka -0.26 - -0.59. Njia bora ya Willems 'ilikuwa sahihi kwa umri wa miaka 9 ambayo ilikuwa na tofauti ya umri wa miaka ambayo ilikuwa chini ya mwezi (-0.07 miaka). Kuhusu asilimia tatu (150, 37%) ya watoto walikuwa na wastani wa miaka yao ndani ya miezi 6 ya umri wao wa muda wakati karibu theluthi mbili (258, 64%) walikuwa ndani ya mwaka mmoja. Kulikuwa na maendeleo kamili ya asilimia 50 ya meno ambayo tayari yalifikia hatua ya mwisho ya ukomavu H. Mvulana mdogo zaidi na msichana aliye na meno 7 mandibular yaliyo na matunda yalikuwa na umri wa miaka 12.91 na 13.02 kwa mtiririko huo. Kwa ujumla, kulikuwa hakuna tofauti ya takwimu kati ya ukomavu kwa wasichana na wavulana katika hatua nyingi za jino. Hata hivyo, wasichana walikuwa na nguvu zaidi katika maendeleo ya taji ya incisor ya kuimarisha na maendeleo ya mizizi ya premolar 1 na canine. Ukomavu wa watoto katika kikundi hicho cha umri umeonyesha tofauti katika hatua za jino. Hitimisho na mapendekezo: Matumizi ya njia ya Willems ilisababishwa na upunguzaji wa umri wa takwimu. Njia hiyo ilifanya vizuri zaidi katika kupima umri wa wasichana ikilinganishwa na wavulana ambao walikuwa zaidi ya wazee. Wengi wa watoto walikuwa na wastani wa umri wao ndani ya mwaka mmoja wa umri wao halisi. Kwa ujumla, kulikuwa na tofauti ya takwimu kati ya ukuaji wa jino kwa wasichana na wavulana katika hatua nyingi za ukomavu. Hata hivyo, wasichana walikuwa wengi mbele ya wavulana katika maendeleo ya mizizi ya canine. Kulikuwa na mifumo tofauti ya ukomavu wa jino kwa watoto wa kikundi cha umri sawa. Matokeo ya sasa yanapaswa kuthibitishwa kwa ukubwa wa sampuli kubwa ambayo ni mwakilishi wa idadi ya watu wa Kenya. Hii itajulisha ikiwa kuna haja ya kurekebisha mbinu ya Willems.

**Appendix G: Quality Ranking Instrument**

| |
|---|
| ***Grammatical accuracy – Is the translation grammatically correct?***<br><br>1. Flawless in terms of grammar<br>2. A few errors, but good<br>3. Impossible to understand. Full of errors. |
| ***Fidelity – The extent to which the TT contains the similar information as that in ST***<br><br>1. The similar information<br>2. Similar information to some extent<br>3. Target doesn't contain any of the information in the source text |
| ***Style – Is the Swahili natural and idiomatic?***<br><br>1. Everything is correct, fluent and natural to read<br>2. Most of it isn't appropriate in terms of register or naturalness<br>**3.** The language is totally inappropriate and the register unsuitable |

## Appendix H: MT Ranking Input Sheet

| | MT System A | | | MT System B | | |
|---|---|---|---|---|---|---|
| | Grammar | Fidelity | Style | Grammar | Fidelity | Style |
| **Scientific Text** | | | | | | |
| **Cultural Text** | | | | | | |