Available online at www.koozakar.com/journal/koozakar-festschrift

Koozakar Festschrift
OPEN-ACCESS | PEER-REVIEWED

Design and Implementation of English to Yorùbá Verb Phrase Machine Translation System

Safiriyu I. Eludiora and Benjamin A. Ayoade*
Department of Computer Science and Engineering, Obafemi Awolowo University, Ile Ife, Nigeria.

*Corresponding author
Email: ayoadebenjamin@gmail.com

Article Information

https://doi.org/10.69798/k4806493

Published Online: April 1, 2024

Academic Editors:
Olawale Adejuwon, PhD
Abiodun Egbetokun, PhD

Additional Information Peer Review: Publisher thanks Sectional Editor and other anonymous reviewers for their contribution to the peer review of this work.

Publisher's note: Koozakar remains neutral about jurisdictional claims in published maps and institutional affiliation.

Reproduced with permission. A prior edition of "Technology Management and the Challenges of Sustainable Development: A Festschrift for Matthew Ilori" with ISBN: 978-978-136-101-2 was published by Obafemi Awolowo University Press, Ile-Ife Nigeria

Abstract

The Yorùbá ethnic group has diverse language speakers across the world, and translating the language to other widely spoken languages must be emphasized. This study aims to develop an English to Yorùbá machine translation system which can translate the English verb phrase text to its Yorùbá equivalent. Words from both languages (Source Language and Target Language) were collected for the verb phrase group in the home domain. Lexical translation was done by assigning values of the matching word in the dictionary. The syntax of the two languages was realized using Context Free Grammar, the rewrite rules were validated with finite state automata. Human evaluation method was used and expert opinion scored. The evaluation shows the system performed better than that of sampled Google translation with over 70% of the responses matching that of the system's output.

Keywords: *Machine translation; Low resource NLP, Rule-based, CFG, Automata.*

1.0. Introduction

The advancement in Natural language Processing (NLP) can be attributed to recent improvements in the strategy and techniques of large data collection, archiving, analysis, and visualization. NLP began in the '50s as machine translation (MT) intended to aid code-breaking during World War II. Although the translations were not successful, these early stages of MT were necessary stepping stones on the way to more sophisticated technologies (Zhang, 2018; Quinn, 2017).

Machine Translation, therefore, is a more classical term, that is classified as a subfield of artificial intelligence that specifically uses computer software for automatic text translation in the absence of a human translator or where humans sparingly participate in the translation process. Machine Translation is highly essential in areas of human needs, while it is gradually taking over from the commonly used human translator, its ever-improving translation techniques promises to be a faster and cheaper alternative. If communication is an integral part of human coexistence then language translation is the engine for its success.

This field of interest is driven by the rapid rise in the area of global information exchange, of which language has been a common barrier in global villages. Hence, translation is an important factor where there exists a language barrier or extreme differences. According to Awoyale and Bamba (2015) in West Africa today, the continent's sub-region alone is home to native speakers of over 500 languages and one of the major African languages is Yorùbá (where there are 19,380,800 native speakers and between 45 and 55 million speakers across the world). Despite the population of speakers, Yoruba is still considered as a low resource language (for which few language resources exist), making it very difficult for the development of more advanced models such as the Neural Machine model that requires large volumes of data.

With the number of speakers, translating the language to other widely spoken languages was not initially emphasized. However, recent linguistic researchers are taking up the challenges by giving more attention (as compared to the high-resource language of the Western World) to endangered languages. The area of interest is machine translation (MT). MT is defined by Kishore *et al.* (2002) as the use of a computer to translate a message, typically text or speech, from one natural language to another. At its basic level, machine translation performs simple substitution of words in one natural language for words in another (Ruchikal and Gupta, 2014). MT systems also are often needed for translating literary works from any language into native languages. This breaks language barriers by making available rich sources of literature to people across the world (Ruchikal and Gupta, 2014; Tsegay and Azath, 2020).

Essentially, there are two major types of translation, which can be categorized as the partial translation and the full translation (Novianti, 2012). Full translation is when every part of the source language text is replaced by the target language text material, while in partial translation, some part of the source language text is translated. In any type of translation that could be implemented, the translator's task is to represent the meaning of source text in the clearest and most acceptable form The technique of translation is always secondary to the understanding of the source text (Aresta *et al.* 2018). A phrase is a part of a clause or sentence (usually a single grammatical unit) with no complete sense. A verb phrase on the other hand usually consists of a verb and a preposition phrase or noun phrase. There have been significant works on English to Yorùbá (E-Y) phrase translation systems such as noun phrase translation (Abiola *et al*., 2014). This research work focuses on the E-Y verb phrase translation. The aim of this project is to develop English to Yorùbá machine translation system which can translate verb phrase text in English to its Yorùbá equivalent.

The objectives of this study are to:

a) collect words from both languages (SL and TL) for verb phrase groups;
b) design the translation process model of the two languages;
c) implement the model in (2); and
d) evaluate the model implemented in (3).

Different models on language translation systems already exist but non treated verb phrases. This research handles verb phrases as its contribution to knowledge.

2.0. Literature Review

2.1 Yorùbá: A brief introduction

The Yorùbá are today one of the three main ethnic groups that make up Nigeria. Yorùbá people are a large ethno-linguistic group or ethnic nation of Africa, and the majority of them speak Yorùbá; a tonal language spoken natively by about thirty million people in Nigeria and in the neighboring countries of the Republic of Benin and Togo (Project Solutonz, 2019). Yoruba has 17 consonant phonemes, /b,f,m,t,d,s,l,r,ʤ,ʃ,j,k,g,k͡p,g͡b,w,h/ (Jolaade, 2016). The language is spoken with three tones, "high" "mid "and "low". According to Eludiora (2014), the high tone is indicated by an acute accent á, é, ẹ́, í, ó, ọ́ and ú while the mid-tone is not marked and the low tone is marked with a grave accent (à, è, ẹ̀, ì, ò, ọ̀ and ù).

I. Yoruba sentence

The Yoruba language clearly follows the SVO (Subject-Verb-Object structure) sentential word order. For example, "Adé pa ewúré" meaning "Ade killed a goat", "Adé" as the Subject, "killed" as Verb (transitive) and goat (Object). The basic structure is as shown:

S → NP (Aux) VP NP

where S is the sentence, NP is noun phrase and VP is verb phrase, N is noun and V is the verb.

NP → N,

VP → V

NB: The NP serves as the Object as well as Subject.

In a case when the verb is intransitive (does not require an object, it comes with a structure

S → NP (Aux) VP

For example, "Adé sún" meaning "Ade slept", "Adé" as the Subject, "slept" as Verb (intransitive). where S is the sentence, NP is noun phrase and VP is verb phrase, N is noun and V is the verb.

S → NP (Aux) VP

NP → N

VP →V

(a) Yoruba transitive verb phrase

A transitive verb phrase consists of a transitive verb and an object or a modifier. For example - "rí kìnìún kan" which means "see a lion". The verb "rí" comes with a Noun-phrase "kìnìún" as a complement accompanied with an indefinite article "a". The phrase structure for such can be represented as;

VP →V NP DET

Where V is the verb, DET is the article and NP is Noun-phrase.

(b) Yoruba intransitive verb phrase

This verb phrase includes an intransitive verb but does not include a Noun Phrase as an object to complement the verb. For example, the phrase" fò kia "which means "jump quickly", the Verb "fò" is the Verb-phrase (VP) and followed by an adverb "kia". In this case the intransitive Verb can be followed by an adjunct/modifier.

VP →V AdvP

Where V is verb, AdvP is an adverbial phrase and Adv is an Adverb,

VP →V

AdvP →Adv

However, it is important to note that, Noun Phrase in Yoruba is quite different from English as the determiners and adjectives, follow the noun, making Yoruba noun phrase head initial. Yoruba has poor morphological processes, such as inflection. Instead, it uses syntax to convey the grammatical meaning. However, Unlike English, Yoruba can express its inflection in nouns and verbs with the help of auxiliaries and other words. A common practice is when a native speaker adds a word before the noun instead of inflecting it.

2.2. Machine Translation

Machine Translation (MT) is a mode of translation in which computer applications are delegated the task of accepting source language (usually text and audio), process the source language according to the instructions and data provided, generating a target language equivalent of the source language and unaltered in meaning as required by the user.

Compared to human translation technique, MT also provides

a) a substitution of one language for another language,
b) study of how to form and structure words,
c) syntax which is the rules over the sequence of words combination i.e., how words are to be combined to form phrases and sentences,
d) semantics which is the branch of linguistics that deals with the meaning of words.

Human translators interpret and analyse all conditions within the text to understand how each word influences the context of the text. An expert (with the knowledge of the syntax and the semantics of the language of interest) is required in source and target language to achieve good result. The Auto translator explicitly simulates the expert knowledge of the structure and rules of the languages with a computational model. Auto-translation has challenges, for instance, automated translators might find it difficult to interpret some context based on the type of system. Also, there exist challenges in the dataset collection such as biases which affect the quality of the trained model. When implemented correctly it promises to improve economies of scale when translating in domains suited to MT.

2.2.1. Rule-based machine translation (RBMT)

The rule-based machine translation systems are built on dictionaries and linguistic rules and depend on the quality and volume of translated text (both the TL and the SL) in the database; say large corporal. This method is applicable when rules are developed manually over time by experts in both languages. The rules are then hand-coded, while the system is designed to find a matching translation and concatenate the text to give a target text. RBMT is a good approach for MT engines of language with lots of grammar (Eludiora, 2014) in which Yorùbá language is one. It is also a good practice for low resource language. Users can improve translation quality by adding terminology into the corpus and improving on the rules generated.

2.2.2. Statistical machine translation (SMT)

In the mid-20th century, Warren Weaver came up with the proposed theory of “cryptanalysis” (Quinn, 2017). Given enough data, patterns generated should be applicable for other newer translations. The Statistical MT usually makes use of very large bilingual corpora for efficiency (Benyamin *et al*., 2018). Finding this type of volume can be specific to some notable languages (the application of SMT is rare to achieve in a low resourced language). The bilingual corporal gets trained until it learns the language’s pattern. Such a pattern is expected to be efficiently reproducible. Brown *et al.* (1990) showed phrase-based SMT models to be often defined with the log-linear framework as shown in the equation below.

$$P(t|s) = \frac{exp(\sum_{m=1}^{M} \lambda_m h_m(t,s)}{\sum_i exp(\sum_{m=1}^{M} \lambda_m h_m(t,s)} \quad (1)$$

where $h_m(t, s)$ is a feature function and λ_m is the weight.

2.2.3. Neural machine translation

The Neural Machine translator (NMT) is based on Deep Neural Network, an architecture designed to imitate the working principle of neurons of the human brain. LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Unit) have proven to be efficient over common neural network algorithm such as ANN as they learn on long term dependencies quickly. This method can be used to map source sequence (input) to target sequence (output) via an encoder and decoder, LSTM and GRU being the basic building block. The encoder extracts a fixed-length representation from a variable-length input sentence, and the decoder generates a correct translation from this representation (Cho *et al.,* 2014). Such complex system would be time consuming to learn new language pairs. As the computational demand of training NMT becomes more widely available and more research is performed, these gaps (accuracy and time optimization) will hopefully be filled in**.**

2.3. Related Works

Abiola *et al* (2014) worked on the computational model of the English to Yoruba (E-Y) noun-phrase translation system. Their project is resolving the restriction of computers to only those who understand the English language in Nigeria which consequently demoting the development of indigenous language. Rule-based translation model (CGG) was the approach used. The model was tested on 160 randomly selected noun phrases from daily news and was used to test the translation. They look forward to using an improved machine language technique in their future work.

Eludiora (2014), implemented and designed English to the Yorùbá Machine Translation System. The authors are aiding the use of the Yorùbá language on computer systems. Transfer Rule-Based, the rewrite rule was verified using NL Toolkits (and implemented using python language).

3.0. Methodology

The translation process was modelled using a Unified Modelling Language (UML) as shown in Fig. 1. A rule-based approach was used in the system's architectural design; using principles and rules developed based on the two languages. These rules guide the translation from the source language (SL) to the target language as shown in the class activity diagram in Fig. 2. The architecture of the system was designed based on the system's capability to tokenize input text (source language). That is, English language, the token is then re-arranged according to the rewrite rules. After which it has searched the dictionary for lexical translation of the token; and then gives the equivalent translation in the target language (Yorùbá language). The output of the system is displayed by the graphical user interface.

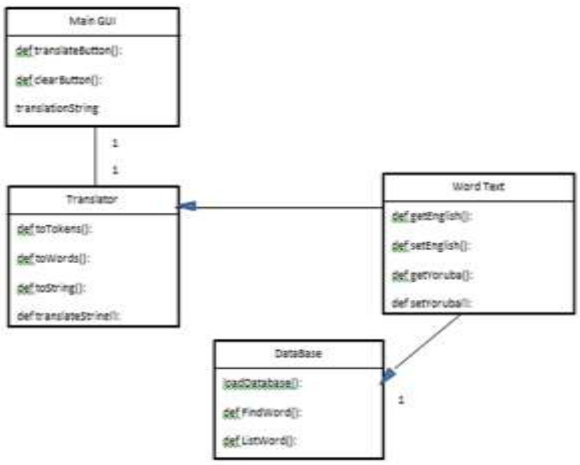

Figure 1: The Unified Modelling Language (UML) Diagram

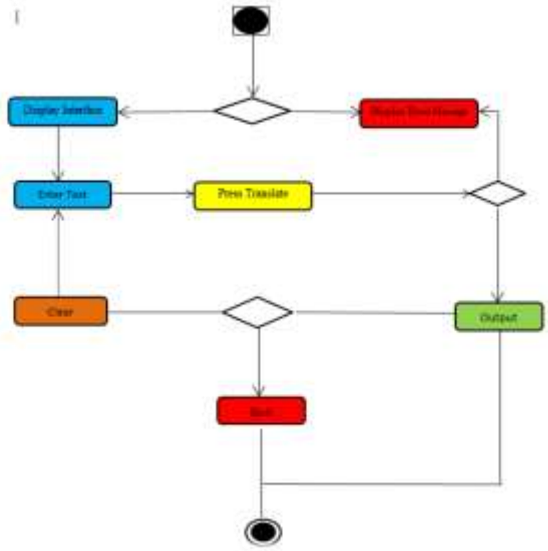

Figure 2: Showing the Class Activity Diagram

Collection and storage

The data collected are five parts of speech (nouns, pronouns, adjectives, verbs, prepositions) and their respective Yorùbá equivalent stored in a python dictionary syntax. The data collected is limited to the home domain i.e., the words used and translated are familiar words that we often use at home. Data were collected from various sources which includes; internet, books, journals, and newspapers as well as the native, since there exists no ready-made resources system library (Yorùbá is one of the many under-resourced African languages). The collected words were grouped and represented using a dictionary for retrieval. The sample of the data is shown in Fig. 3.

```
nouns = {'father': 'bàbá', 'mother':'iya','boy':'ọmọdọkúnrin',

det  =  {'the': 'náà', 'a': 'kan'}

verbs = {'gave':'fún','give':'fún','wash': 'fọ','washed':'fọ',
         'kill':'pa','fight':'já','are': 'wa', 'sat': 'jókò',
preps = {'in': 'nínu', 'to': 'sí', 'for': 'fún', 'with': 'pẹ̀lu

adjectives = {'cold': 'tútù', 'small': 'kékeré', 'big': 'nlá',
```

Figure 3: Sampled dictionary for data retrieval

3.1. Translator engine

The development process was divided into stages, the Graphical User Interface (GUI) and the Translator Engine. The translator engine was designed (database plus rules) using a monolithic architecture, that is, the components are developed as a single unit. The GUI was designed in three layers. The first layer accepts input from the user while the other two layers are used to output the result. The lexical translation is done by a search through the diction for a match after the sentence has been tokenized. The values of the corresponding text are returned, rearranged, and combined using these rules. The concatenated words are sent back to the GUI in order to display the grammatically correct Yorùbá equivalent.

3.2. Grammar

The type of grammar used in this work is context free. The verb phrase analysis depends upon knowing which theory will be obtained in the context. Using the formal language theory, suppose $\beta = (V, \Sigma, \rho, \gamma)$ is a 4-tuple context-free grammar, where V is a set of non-terminal symbols, Σ is terminal symbol, ρ is production rule and γ is the start symbol which is an element of V. Let $L(\beta)$ be the function of language generated by β over the finite alphabet Σ; A typical rule is $A \rightarrow \omega$ where A is a symbol of V and ω is a string element in the language. David (2020) shows the language generated by β is $L(\beta)$ over the alphabet Σ.

$$L(\beta) = \{\omega \in \Sigma \mid \gamma\} * \Rightarrow * \beta\ \omega \qquad (2)$$

This contains terminal symbols with start symbol γ and applying the suitable production rules ρ. The base sentence was represented as a typical simple sentence which consists of the non-terminal noun-phrase (NP) and verb phrase (VP) where S is a single nonterminal symbol. Where VP is a verb phrase, V is the verb, N is a noun, NP is a noun phrase, ADJ is an adjective, DET is determinant, P is a preposition, PP is a prepositional phrase.

Start Rule

S → NP VP

NP →N

VP → V

This grammar above takes any noun and verb and shows how to combine them to form a sentence. This work is focused on verb phrase, as such, a need to improve on the verb phrase rule a bit more. Prepositional phrases can be handled with the rule as well, by combining our free structure rules for noun phrases and verb phrases.

a) Source language production rule

The start rule structure was used to derive a simple but more complete phrase structure. Now given the Verb phrase as a single nonterminal left-hand symbol VP, the 6 right-hand rules (terminal and nonterminal shown in fig 4) were parsed;

VP → V,
VP→ V NP,
VP → V ADJ N PP,
VP→ V DET ADJ N PP,
VP→ V DET ADJ N P NP,
VP → V DET ADJ N P DET N.

From the above rules, according to the scope of this work, all non-terminal symbols are:

Verb Phrase
VP → V NP,
VP → V NP PP,
VP →V

Noun Phrase
NP → N,
NP → DET N,
NP →DET ADJ N

Prepositional Phrase
PP → P NP

Some verb phrases and structures are shown below;

V ⇒ | eat |
V NP ⇒ | eat | | the meat |
V PP ⇒ | eat | |on the table|
V NP PP ⇒ | eat | |the meat| |on the table|
While the terminal symbols are:
V ⇒ |ate|,
N ⇒ |meat|, |table|,
ADJ ⇒ |fresh|,
DET ⇒ |a| |the|,
P ⇒ |on|

b) Target language production rule

From the start rule, given the verb phrase as a single non terminal left hand symbol VP. Generally, the phrase structure Rule for all the VPs in Yoruba would be VP→ V (NP) (PP) (AdvP). In this work, 4 right-hand production rules (terminal and non-terminal) were derived for Yorùbá verb phrase as given below; The verb phrase parse tree for the SL production rule is as shown in Fig. 4.

VP → V,

VP → V NP,
VP →V N ADJ,
VP → V N ADJ DET PP.

The terminal and non-terminal symbols are maintained but there are changes in the language, While the terminals of the Yorùbá production rule remains the same with the source language (English) production rules, there exists a slight variation in some of its structure.

V ⇒ | je |
V NP ⇒ | je | | eran naa |
V PP ⇒ |je| |lori tabili|
V NP PP ⇒ | je | |eran tutu naa| |lori tabili|

The terminals of the verb phrase representation in Yorùbá language are as listed,

V →| je |,
N → |eran|, |tabili|
ADJ → |tutu|,
DET → |naa|,
PP → |lori|

Note that the rules derived in this work are only within the scope of this work.
The Yorùbá verb phrase parse tree for the TL production rule is as shown in fig 5.

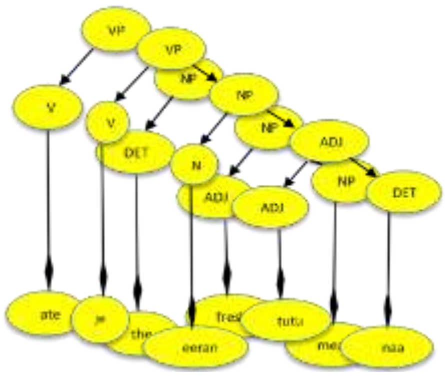

Figure 5: VP parsed tree for target

3.3. Finite State Automata (FSA)

The finite automaton is a mathematical model used to compute how both systems changed states based on the input supplied, the SL FSA model was simulated using JFLAP in fig 6 and and the TL in fig 7.

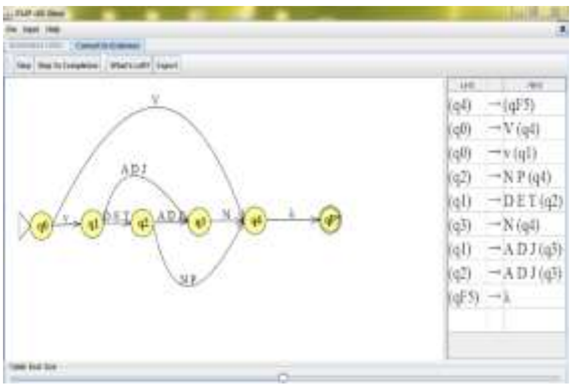

Figure 6: Showing the FSA model of Source

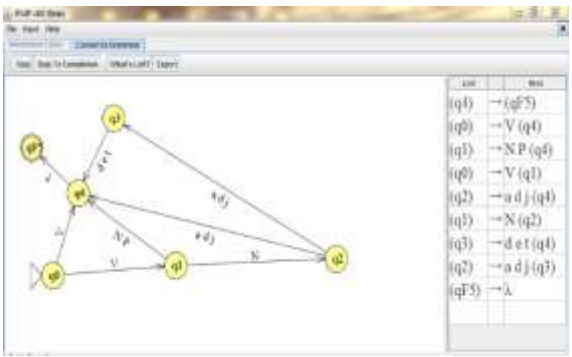

Figure 7: Showing the FSA model of Target

4.0. System Implementation and Evaluation

4.1 System implementation

Python programming language was used for the software coding and the interface of the machine is designed using a module named Tkinter[1]. The translation process is based on the grammar built in the program code which follows the rewrite rules. The graphical user interface in fig 8 gives a user-friendly interaction. The output section is divided into two parts: the first part contains a text box that has the direct diction-based translation in the Yorùbá language and the second text box which contains rule-based translated words. Fig 9 shows a quick the step-by-step methods used for the word for word and rule-based output.

4.2 Evaluation

Figure 8: Translation Sample

The human evaluation method was used for the evaluation of this system. Although this method was time-consuming, it is very extensive. The system output and google translator output were compared with the opinion of the respondent (language expert) using random examples from the data collected. There were 70 responses, the opinion shows that 69% of respondent's translations tally with the E-Y VP translation, 29% follows the google translation while 1% had a totally different opinion as shown in Fig. 10. The responses were collected using Microsoft excel online form and the visualization was done with excel chart.

5.0. Recommendations and Conclusion

The translation conveys the meaning and undertone of the translation when applicable. The system translates with appropriate tones and under-dots. This machine translation can be used as a means of teaching students. It is not a closed model and is practically easy to learn. Therefore, a lot of improvements can be made on the part of the phrases that make up the sentence. There is room to make the system's engine more robust by adding to the rewrite rules. And finally, the system can be further improved by adding to the database, this will improve the response of the system given any relevant query.

---

[1] https://github.com/BenAji/VerbPhrase

Create widget: the widget method links the user interface to the engine using the get () method.

Initialize the empty list: a list that holds text having the same part of speech (POS).

Initialize empty string: initialize an empty string to hold the converted/concatenated text and another to hold Word for Word (w4w) translate.

Initialize the dictionary: dictionary stack as database.

> ***def intoTokens***: returns a token from an input text.
>
> ***def group***: grouping the tokens to their respective part of speech by iterating through the tokenized text and appending to the list created.
>
> ***def convert***: iterating through the list if a word has a match in the initiated POS list then append the target language equivalent. Iterating through the initialized list for each word using len(list)+1 and starting from the last appended value.
>
> ***def w4wconvert***: word for word(w4w) method iterate through the list, if a word has a match in the initiated POS list then append the target language equivalent. Return the dictionary value of tokenized text, and then concatenate the strings.

Output: prints the output of the convert () and w4w convert method to the widget using the get().

Figure 9. rule-based MT Pseudocode

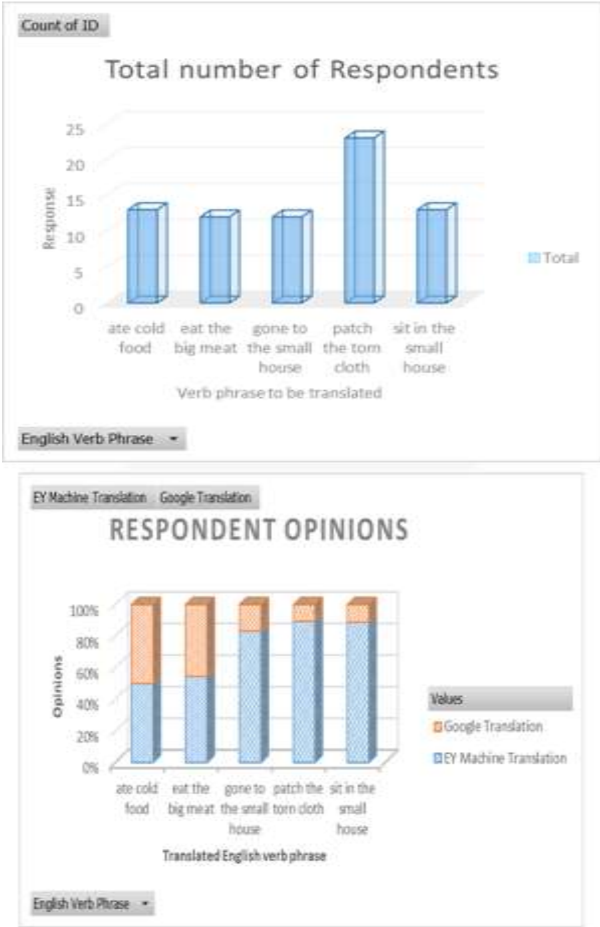

Figure 10: Chart Showing the Responses

References

Abiola O.B., Adetunmbi A.O., Fasiku A. I. & Olatunji K. A. (2014). A Web-Based English to Yoruba noun-phrases Machine Translation System. *International Journal of English and Literature*,5(3),71-78. http//dx.doi.org/10.5897/IJEL2013.0472.

Ahmadnia, B, Haffari, G and Serrano, J. (2018). Statistical Machine Translation of Bilingually Low-Resource Scenarios: A Round-Tripping Approach. *IEEE 5th International Congress on Information Science and Technology, pp.*261-265. doi:10.1109/CIST.2018.8596614

Aresta, R. (2018). The Influence of Translation Techniques on the Accuracy and Acceptability of Translated Utterances that Flout the Maxim of Quality. *Journal Humaniora*. 30. 176. https://10.22146/jh.v30i2.33645.

Awoyale, Y. and Bamba, M. West African Language Data Sheet. 2015. [Available Online]. https://www.ldc.upenn.edu/sites/www.ldc.upenn.edu/files/west-african-languages.pdf

Brown, P. F., Cocke, J., Della Pietra, S.A., Della Pietra, V., Jelinek, F., Lafferty, J.D., Mercer, R. L. and Roossin, P.S. (1990). "A Statistical Approach to Machine Translation," Computational Linguistics. 16(2), June 1990, https://www.aclweb.org/anthology/J90-2002

Cho, K., Merrienboer, B., Bahdanau, D., and Bengio. Y. (2014). On the Properties of Neural Machine Translation: Encoder-Decoder Approaches. *Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation (SSST-8)*. https://arxiv.org/abs/1409.1259v2

Eck. D. J. (2020). Context-Free Grammars. *EngineLibretexts*. [Available Online]

Eludiora, S (2014). Development of English to Yorùbá Machine Translation System, unpublished Ph.D. Thesis, Submitted to the Postgraduate College, Obafemi Awolowo University, Ile-Ife, Nigeria, 2014, 241pp

Kishore, P., Salim, R., Todd, W. and Jing, Z-W (2002). BLEU: A Method for Automatic Evaluation of Machine Translation. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, Pages 311–318, https://doi.org/10.3115/1073083.1073135

Novianti, E. (2012). An Analysis of the Translation Strategies of Idiomatic Expressions; *in Lewis Carroll's Alice In Wonderland in its Translation by Khairi Rumantati*, Universitas Negeri, https://eprints.uny.ac.id/9241/ [Accessed 23 January 2021]

Okanlawon, J. (2016). An Analysis of the Yoruba Language with English. *Phonetics, Phonology, Morphology and Syntax. North Eastern University.* [Available Online]

Project Solutionz. (2019). Role of Propagating Yoruba Culture. https://projectsolutionz.com.ng/role-of-propagating-yoruba-culture

Quinn, D. (2017). The Cryptological Origins of Machine Translation, from al-Kindi to Weaver. amodern. 8, pp 1–20. http://amodern.net/article/cryptological-origins-machine-translation/

Sinhal, R, Gupta, K. (2014). Machine Translation Approaches and Design Aspects. *IOSR Journal of Computer Engineering*. 16. 22-25. 10.9790/0661-16122225.

Tsegay, K, Azath, M. (2002). Statistical Machine Translation for English to Tigrigna Translation, *International Journal of Scientific & Technology Research*. 9(1), 2095-2099

Xinwen, Z. (2018). The Evolution of Natural Language Processing and Its Impact on AI, Forbes.[AvailableOnline].https://www.forbes.com/sites/forbestechcouncil/2018/11/06/the-evolution-of-natural-language-processing-and-its-impact-on-ai/#1fb55aeb1119

A Appendices

Full EY corpus used can be accessed here[2]

Noun

Parts of the body

arm Apa
back Eyin
cheeks Ereke or eeke
chest Aya
ear Eti

Objects

bathroom Baluwe
bed beedi
bedroom Yara Ibusun
ceiling Orile

Food/Meal

bread Buredi
breakfast Onje aaro
butter Bota

Yorùbá Numbers (Cardinal and Ordinal)

One eyokan
first okan
two meji

Adjectives

Colors

Black dudu
Grey awo-resuresu

[2] https://github.com/BenAji/VerbPhrase/blob/main/EY-Corpus.docx