

Modeling Predictors of Accuracy in Web-Translated Newspaper Stories

Emmanuel Ifeduba, Redeemer's University, Nigeria*

 <https://orcid.org/0000-0002-7121-3279>

ABSTRACT

Recently, newspapers published in Africa began to adopt web translation applications to make their businesses more competitive. However, studies indicate that web translations of even major languages are often inaccurate and generally gloss over how this affects African languages. And the predictors of translation inaccuracy seem to be inadequately interrogated. This study, therefore, investigates the extent to which Google Translate accurately translates English to eight African languages and the relationship between translation accuracy and perceived journalistic errors, orthography, and technological limitations of the translating machine. Through document analysis of six newspaper stories, the study ascertained that the meaning of over 45% of the text was either lost or unclear. Statistical analysis shows that perceived journalistic errors, inadequate orthography, and technological limitations significantly predict translation inaccuracy, suggesting that improvement in these variables would improve the accuracy of web-translated news.

KEYWORDS

African Languages, Google Translate, Journalistic Error, Orthography, Technological Innovation, Web Translation

INTRODUCTION

The primary aim of web translation services has been to help internet-driven media to disseminate content to a wider audience by removing the language barrier which has been a drawback to massive cross cultural dissemination of information (Hub, 2016). Thus, by virtue of web translation innovations, people from different parts of the world can now read the same content and access the same services and products on websites using multiple language translation services (Manasse, 2017). This, understandably, must be why the number of online translation services has increased tremendously from a handful in 2000 to about thirty in the past ten years. Currently, such services available online include Google Translate, Bing Translator, Translate.com, SDL Free Translation, DeepL Translator and so many other services ((Munday, 2001; Lingo, 2016). Whereas some of the services can only be used for a fee, numerous others are free to readers. A list of twenty-two free online translation services is presented in Table 1.

DOI: 10.4018/IJTIAL.304074

*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

Table 1. Free online translation services

SN	Translator	SN	Translator
1	Google Translate	12	Yandex
2	Bing Translator	13	Reverso
3	SDL Free Translation	14	Collins Online Translator
4	Permondo	15	Online-translator.com
5	Eprevodilac	16	Prompt Online Translator
6	Onlinedoctranslator	17	Systranet.com
7	Translatesonline.com	18	Tradukka
8	PONS	19	Daytranslations.com
9	Elanlanguages.com	20	Linguee
10	Worldlingo	21	Babelfish
11	Applied Language	22	Windows live Translator

The more the free translation services, the more the number of newspapers adopting the innovation. However, studies have shown that these translation services translate major languages such as English, Spanish, Portuguese and French almost perfectly, but exhibit some shortcomings in translating less popular languages and technical subjects (Parisa, 2015). Consequently, there have been studies interrogating the accuracy of web translations with regard to major European languages, Arabic and Chinese, but only a passing mention is usually made of African languages in a few of the studies (Patil and Davies, 2014). Notwithstanding, African newspapers have begun to adopt this innovation, especially Google Translator, with a view to reaching a wider audience within the continent and in the diaspora. However, little is known as to how accurately Google translates most African Languages or how well Google translation serves readers who would rather be served in their native African languages (Muller, 2009).

Statement of the Problem: Previous research has shown that only a few studies have interrogated web translation of newspaper reports; and coverage of African languages in such research have been passing mentions or conjectures lacking empirical merit (Fredholm, 2014). And it appears that there is hardly any attempt as yet to examine how well such translations serve newspaper readers in Africa. Furthermore, though previous research by Huang (2011) pointed out the inadequacies of African language lexicography as a factor which might affect translation accuracy, the correlates of inaccurate translation are yet to receive adequate attention in extant literature. This study, therefore, examines the accuracy of Google translation of newspaper reports to African languages with a view to highlighting the shortcomings and perceived correlates towards improving translation services and user experiences in Africa. The study attempts to accomplish this objective by providing answers to the following questions.

To what extent is Google Translator able to accurately translate Afrikaans, Swahili, Shona, Igbo, Yoruba, Arabic, Amharic and Hausa languages? What aspects of the source language (English) are difficult for Google Translator to translate to African languages? What is the relationship between translation accuracy and perceived journalistic errors, inadequate orthography and technological limitations of the translating machine?

BACKGROUND LITERATURE

- **Web Translation:** Web translation is regarded as a subfield of computational linguistics that investigates the use of computer software to translate speech from one natural language to another

by performing a simple substitution of words in one language for words in another. The broad goals of web translation include dissemination, assimilation, information exchange and access, and newspapers surely are in position to benefit from dissemination, information exchange and access, especially as they struggle to find their feet in the digital environment. And the cost-free advantage presented by Google Translate seems to suit their revenue diversification purposes (Ifeduba and Olatunji, 2019).

Google translate application, launched in 2006 with two languages, was based on statistical machine translation, but in 2017, Google moved away from phrase-based machine translation to neural machine translation, meaning that it now translates whole sentences at a time, instead of pieces of a sentence. It is currently probably the best known online language translation service provider, and can process upwards of 100 billion words per day, or around 4.2 billion words per hour (Leddy, 2019). Presently, it offers full support for translation between 100 different languages. Translations are generated on the basis of statistical models whose parameters are derived from the analysis of bilingual text corpora (Patil and Davies, 2014, McGuire, 2018).

A study evaluated the accuracy of google translate in medical communication. Ten medical phrases were translated to 26 languages selected from three continents including Africa. In total, 110 out of 260 phrases were inaccurately translated and 45% of the medical terms translated to two African languages were inaccurate (Patil and Davies, 2014). Another study assessed the accuracy of Google translate to allow data extraction from trials published in non-English languages and found that relative to English, extractions of translated Spanish articles were most accurate compared with other translated languages on the one hand. On the other hand, translated Chinese articles yielded the highest percentage of items that were incorrectly extracted more than half the time (Balk, et. al., 2013).

Lin, et al. (2010) stated that there is a huge gap between human and machine translations. On the one hand, machine translations always have limitations in quality and, therefore, are not often used for translating sensitive documents for which accuracy is of essence. As a consequence, the various market segments may in future be dominated by either human translation or machine translation. Which product offering dominates in any particular market segment will depend on the unique characteristics and demands of that segment (Lazzari, 2006). The mass media segment is likely to depend on machine translation due to its peculiarities. However, the fact that the current machine translation systems do not pay detailed attention to discourse and context in source language analysis, like sentence based systems, may continue to constitute impediments to understanding. Notwithstanding, they are likely to grow in popularity due to their low-cost advantages (Zhang, 2009).

The Internet drives that cost advantage and creates endless communication traffic between different language groups, turning translation into a bridge that connects speakers of different languages. And when instant translations are needed, human translators are not able to supply them fast enough, implying that the need for machine translation should grow even though Google Translate can misinterpret complex structures and provide inaccurate translations (Butler, 2011). Though Van-Rensburg et. al (2012), argued that people expect Google Translate to deliver accurate translations, not taking into consideration that such machine translation applications do not have the world knowledge and language capabilities of human translators, it makes sense that the depth and breadth of its inadequacies should be explored, understood and addressed.

When machine translation tools became available online, there were several reasons for rejection and low adoption of machine translation software (Ghasemi and Hashemian, 2016). One of the reasons is that, the early commercial systems required labor-intensive and time consuming process of data transfer from existing collections to new systems. But the cost advantage and a burgeoning culture of mobility continue to increase the importance of and demand for web translation applications.

Abdelaal and Alazzawie (2020) cited the work of Handschuh (2013) on German-English translation using four different online machine translations (Google Translate, SYSTRAN, Bing and Babylon) and found that there are errors found in the target text output produced by the machine

translations. It was observed that original meanings were often not retained when longer texts were translated. Keshavarz (2012) studied Italian-English translation using the Google translate but the findings showed that the direction of translation did not affect the quality of the translations rendered by the Google Translate. In the same vein, Schairer (1999) investigated the efficiency of three Spanish-English translation programs: Spanish Scholar, Spanish Assistant, and Spanish Amigo. Participants were asked to evaluate the accuracy and acceptability of the translations produced by these programs and made a comparison with human translations. The study also investigated the efficiency of the post-editing process, striving to ascertain whether original human translations are more effective than the machine translation. English-Spanish bilinguals evaluated 23 sentences that were translated into Spanish from English and scored each sentence twice on a scale of 1 to 5 for correctness and comprehensibility. The study concluded that even when users followed pre-editing instructions, the quality of output received poor ratings. Thus, machine translations were not evaluated as “successful” by the participants (Schairer, 1999).

A similar result was reported in a study attempting to answer the question: Which languages are Google Translate best at translating? Benjamin (2018) stated that English, Spanish and Portuguese are among the languages which the Google application translates with a high degree of accuracy. In line with this observation, a 2019 study indicates that Google translate has improved its accuracy in the last eight years by 34% with regard to English to German Portuguese, Spanish, Danish, Greek,, Afrikaans, Polish, Hungarian, Finnish, and Chinese (Aiken, 2019).

- **Web Translation of Newspaper Stories:** Another study stated that machine translation of news headlines may not be as accurate as expected due to the fact that the sentences are usually fragmentary and abbreviations and acronyms of proper names are used frequently. The study also observed that headlines which come at the top of a news items seem to be more difficult for machines to translate since the context information useful to disambiguate abbreviations, names and acronyms is often not available (Ono, 2003).

In the same vein, scholars observed that translating newspapers presents peculiar challenges mainly because newspapers are made up of different sections and different types of articles often presented with different styles of writing, suggesting that what works in one section of the same paper or article may not work for others. It argued that style, terminology and cultural contexts also affect the accuracy of machine translation of newspaper articles (Jordan, 2015). Similarly, Valdeon (2020) examined the use of the concepts of framing, gatekeeping and convergence in the study of newspaper translation as well as the use of mixed method approaches, and the variety that seems to characterise journalistic translation research.

Observing that European newspapers were banding together to provide more thorough coverage in 24 official languages spoken in Europe, Lichterman (2016) explained that re-using another newspaper’s reports often means that publishers need to translate stories and adapt them to local audiences, observing that translating content presents challenges, increases costs and consumes time. Consequently, different news organizations are trying different approaches to translation, including centralized translation employed by the seven LENA papers— Germany’s *Die Welt*, Belgium’s *Le Soir*, Italy’s *La Repubblica*, Spain’s *El País*, Switzerland’s *Tages-Anzeiger* and *Tribune de Genève* as well as France’s *Le Figaro*.

- **Correlates of Translation Accuracy:** Online translation machines rely heavily on existing electronic dictionaries, and African language dictionaries are increasing online annually. Large-scale commercial production of electronic dictionaries has also boomed since the mid-1990s. Today, dictionaries on CD-ROM typically come in the back pocket of their hardcopy counterparts, while the number of dictionaries on the Internet already runs into tens of thousands. A study focused on Southern African languages found that there are nearly one hundred and twenty

different African language dictionaries were online by 2010. However, the sizes of the current African-language internet dictionaries were found to be generally small, and the contents not often of a high quality (Huang, 2011). The implication is that this could pose a limiting factor on accurate web translation of African languages. Furthermore, it is necessary for the translating machine to pay attention to what is described as the transitivity system, mood system, modality system and theme system. These embody ideational meaning, interpersonal meaning, and textual meaning which machines often miss or mix up (Si and Wang, 2021). Details of Sub-Saharan African languages with web dictionaries are presented in Table 2.

The more technical a text appears the more likely the inaccuracies in its translation. Technical languages such as medical language and specialized writings such as journalistic writing may, therefore, pose additional problems to translation machines. With regard to medical communication, a study indicates that Google provided accurate translation for simple sentences, but inaccuracies increased when the original English sentences became more complex and sophisticated (Chen, et al, 2016; Setiawati, et al, 2020). For African languages, struggling with under-reportage, and striving to expand its reach, the need for translation cannot be over-emphasized irrespective of the shortcomings of web translation (Chikaire and Ezeru, 2021).

As indicated by Napitupulu (2017), lexico-semantic errors, tense errors, preposition error, word order error, distribution and use of verb group errors as well as active and passive voice errors can all affect translation accuracy, implying that journalistic errors made in the process of hurrying to

Table 2. Distribution of Sub-Saharan African language internet dictionaries by country

S/N	Country	Languages/Dictionaries
1	Angola	1
2	Botswana	3
3	Burundi	2
4	Congo	2
5	DRC	6
6	Gabon	2
7	Kenya	14
8	Malawi	7
9	Mozambique	5
10	Namibia	2
11	Rwanda	2
12	South Africa	4
13	Tanzania	81
14	Uganda	10
15	Zambia	5
16	Zimbabwe	5
17	Nigeria	5
18	Ghana	3
	Total	159

Adapted from Huang, 2011

meet deadlines would affect translation accuracy one way or another. In the same vein, Abdelaal & Alazzawie (2020) stated that omissions or lexical errors and semantic errors arising from inappropriate choice of words were the most common errors in some Google translations. They observed that inappropriate lexical choice could arise from the homophonic nature of some source- text words, and these words are sometimes misinterpreted by the translating machine. In other words, three factors—translation machine, orthography and journalistic errors can all contribute to the accuracy level of web translations.

- **Theoretical Framework:** The evaluation of accuracy is anchored on El Shiekh and Saleh's translation evaluation framework. The framework measures whether a translation is easily understood, fluent and smooth, conveys the literary subtleties of the original, conveys the meaning of the original text, reconstitutes the given text into the target language accurately, captures the style or atmosphere of the original text. Studies in support of this approach argue that it is necessary for translators to factor -in the mood system, modality system, transitivity system, and the theme system that embody ideational meaning, interpersonal meaning, and textual meaning (Si, 2021). This framework is supported with the theory of supervening social necessity (El Shiekh and Saleh, 2011).

Brian Winston, the proponent of the theory of supervening social necessity, argues that a supervening social necessity, such as the need for instant translation, creates a need for a particular technology, but the law of the suppression of radical potential means that the new technology would be integrated into the *status quo* (in this case, human translation) as opposed to radically disrupting it. This theoretical perspective fits into the picture created by the reviewed literature in the sense that the new technology is being actively integrated into the *status quo* irrespective of its inadequacies. It also aligns with the observation that the radical potential of web translation technology has not been suppressed by its development for specific and narrowly defined applications such as online newspaper readership across languages (Ashraf, 2018).

PURPOSE AND METHODS

- **Sample Selection:** Six newspaper stories (two from each) were purposively selected from three African newspapers published in English language. The newspapers, *Vanguard*, *Complete Sports* and *Champion*, were online and had installed the Google translation application at the time of the preliminary survey. Business, politics, entertainment and sport stories were purposively selected to ensure that various aspects of language orthography were covered, and because these are popular contents likely to attract readers from within and outside the immediate environments of the newspapers. Eight most widely spoken African languages (Afrikaans, Amharic, Arabic, Hausa, Igbo, Shona, Swahili and Yoruba) were selected because they adequately represent the African population.
- **Quantitative Analysis:** The translations were analysed using a framework proposed by El Shiekh and Saleh (2011). It measures whether a translation successfully answers some questions including: Is it easily understood? Is it fluent and smooth? Does it convey the literary subtleties of the original? Does it convey the meaning of the original text? Does it reconstitute the given text into the target language accurately? Does it capture the style or atmosphere of the original text? And one simple indicator of machine's inability to provide a positive answer to all of these questions is appearance of the source language in a translated text. Thus, occurrence of English words in translated texts was adjudged as evidence of inaccuracy. Each story published in English language was translated into each of the African languages and the number of English

language words that could not be translated to each language was counted. The frequencies and the percentages were computed to answer research question one.

- **Qualitative Analysis:** Qualitative data were used to answer research question two. English words which could not be translated and by the machine as incorrect were copied out and two native speakers of each language were asked to confirm the correctness of the words, phrases and sentences.
- **Perception Analysis:** To answer research question three, people who understand English and speak at least one of the languages were asked in a questionnaire to state their perception of the sources of inaccuracy in Google translated stories attached to the questionnaire. A total of 191 responses were retrieved and analysed using the Statistical Package for the Social Sciences, SPSS. Pearson correlations were computed to measure relationships and assess predictability.

DATA PRESENTATION AND ANALYSIS

- **Quantitative Analysis:** Six newspaper stories containing 1306 words were translated, and the data indicate that Arabic (3.56%), Amharic (6.10%) and Swahili (7.8%) had the least number of English words whereas Shona (20.6%) Yoruba (19%) Igbo (18.5%) and Afrikaans (17%) had the highest share of the 393 words not translated. Details are presented in Table 3.

In terms of verbiage, the total number of English words translated was 1,306. The app translated this to Shona with the least number of words (1334) followed by Amharic, 1486 words, Arabic, 1594 words and Swahili, 1658 words. Verbiage was highest in Hausa with 1982 words, followed by Yoruba 1945 words, Igbo 1828 words and Afrikaans, 1827 words. Verbiage pattern in the translations suggest that the languages with more advanced orthography (Arabic, Amharic and Swahili) accomplished the task in fewer words than the others, except Shona whose performance could not be explained by its level of orthography advancement.

- **Perceived Correlates of Translation Inaccuracies:** Respondents that speak at least one of the languages were asked to indicate the extent of their agreement with statements about their perception of journalistic errors, inadequate orthography and limitations of the translation technology as correlates of the inadequacies in the translations. Their responses indicate that between 60% and 70% of the 191 respondents thought that inaccuracies arose from all three

Table 3. Extent of inaccuracies (words not translated)

SN	Language	Total Story Length in Words (English)	Total Story Length (Translation)	Words Not Translated	Percent
1	Swahili	1, 306	1658	31	7.88
2	Arabic	1, 306	1594	14	3.56
3	Hausa	1, 306	1982	28	7.12
4	Shona	1, 306	1334	81	20.61
5	Igbo	1, 306	1828	73	18.57
6	Yoruba	1, 306	1945	75	19.08
7	Amharic	1, 306	1486	24	6.10
8	Afrikaans	1, 306	1827	67	17.04
	TOTAL			393	100%

possible sources, with a higher number (73%) indicating that technology is a more likely determinant of inaccuracy. Details are presented in Table 4.

Pearson correlations were computed to assess the relationship between the performance of the translation machine and respondents' perception of such performance as emanating from journalistic errors, orthographic limitations of target languages and technological inadequacies of Google Translator. The results for all the languages indicate that there is a significant relationship between the extent of errors in the translations and respondents' perception of shortcomings in the three variables: journalistic errors, orthography and Google translate technology. With N=187, Pearson r ranged between .190 for Arabic and .813 for Shona while p-value was .000 apiece. This means that reductions in journalistic errors, orthography limitations and technological limitations would lead to corresponding reductions in translation inaccuracies all other things being equal.

Further analysis was conducted to assess the predictive value of a model based on the three variables and the results indicate that the model is significant in predicting an increase in Google translation accuracy. The R square value of the model was .977, meaning that 97.7% of the variance in the accuracy of web translation could be explained by perceived journalistic error, orthography and technology. Regression details are presented in Table 5.

- **Predictors:** Journalistic Error, Orthography and Technology
- **Qualitative Analysis:** For each of the languages, many words could not be translated from the source (English) to the target languages. Though it appears that some could be accounted for by machine programming (programmed to recognize all capitalizations within a sentence as names), some other words beginning with small letters within sentences could not be translated. Some examples of the words not translated are presented in Table 6.
- **Inaccurately Translated Sentences and Phrases:** Over 45% of the translations were inefficiently done, producing either meaningless or partially meaningful outcomes. An example is a headline which read: "Yakubu loses bid to retrieve forfeited \$9.8m" which Google translated to: "Yakubu

Table 4. Perceived correlates of translation inaccuracies

SN	Statements	Agreement	Neutral	Disagreement
1	Errors in Google translation of English to African languages originate mainly from journalistic errors in the newspaper stories	131 (68.5%)	8 (4.1%)	52 (27.2%)
3	Errors in Google translation of English to African languages originate mainly from the limitations of the orthography of my native language	133 (69.6%)	7 (3.6%)	51(26.7%)
4	Errors in Google translation of English to African languages originate mainly from the inadequacies of the translation machine	140 (73.2%)	10 (5.25)	41(21.4%)

Table 5. Analysis of variance between level of accuracy and perceived journalistic error, orthography and technology

Model		Sum of Squares	DF	Mean Square	F	Sig.	R. Square
1	Regression	115.652	2	57.826	4018.167	.000	.977
	Residual	2.706	188	.014			
	Total	118.358	190				

Table 6. Examples of words and sentences not translated

SN	Language	Examples of Words and Sentences not Translated
1	Afrikaans	Argument, High School, video, singles, begin, tweets, superstar, studies, better, week, handbook, later, was, is, below, united, word, my league, champion, respondent
2	Amharic	Ex-, pm, buzz, lamentation, super and eagles, invited, conference
3	Arabic	E-mail, billboard, Music, sports, Village Square, “Anichebe doubtful for Arsenal as Moyes eyes...”
4	Hausa	Mr, May, Billboard, Tweeted, united, premier, super, eagles, week, latest, notoriously, spree, celebrity, artists, rapper, award, hip hop, online, my, an
5	Igbo	Federal, email, fans, tweets, dash, rapper, premier, Champions league, trophies, striker, “cup nke nations”
6	Shona	Federal high court, knowledge, February, dead serious, “Let me make those payments mangwana tozoona kuti ndasarirwa, nemari here” singles, semi-chokupedzisira, super, Africa cup of nations, qualifiers, chief executive officer, tuition, quartet
7	Swahili	Metamorphosed, tweet, alitweet, super, eagles
8	Yoruba	Federal, metamorphosed, tweeted, united, cauldron, league, sports village, crucial, “Africa cup of nations”

efunahụ ike iweghachite \$ 9.8m furu efu” in Igbo. And this translation actually means “power lost Yakubu to recover forfeited \$9.8 million.” But a properly translated version should read, “Mbo Yakubu gbara iweghachite nde dola 9.8 furu efu enweghi isi.”

The headline of the second story which reads: “Nicki Minaj to pay university tuition for dozens of fans” is translated to Yoruba as “Nicki Minaj lati sanwo ile-eko giga fun opolopo awon egeb”. This google translation does not capture the idea expressed in the original language properly. Two expressions “to pay” and “fans” are given wrong translation in the target language (Yoruba). Contextually, from the introduction, the idea is that Nicki Minaj promised to pay fans. The right expression for supporters or fans in Yoruba is *alatilehin*. The headline, therefore, should either be “Nicki Minaj yoo sanwo ile-eko giga fun opolopo awon alatilehin re” or “Nicki Minaj seleri lati sanwo ile-eko giga fun opolopo awon alatilehin re”. The word “egeb” found in the google translation does not exist in Yoruba.

- Perceived Correlates of Inaccuracies:** In line with the quantitative findings indicating that all three factors contributed to inaccurate translations, all the news items indicated that the source language (English) contained journalistic errors that contributed to the translation inaccuracies. For instance, a punctuation error in one of the headlines resulted to translation inaccuracy in all the translated versions. The headline reads: Udeze: Ndidi, Moses Crucial for Super Eagles AFCON, World Cup Hopes instead of: Ndidi, Moses Crucial for Super Eagles AFCON, World Cup Hopes—Udeze. This journalistic error of placing colon, dash and Udeze in the wrong positions confounded the machine, resulting in either inaccurate or totally meaningless translation in each of the eight target languages. In the Igbo, Hausa and Yoruba translations, for instance, this confusion extended to the introductory part of the story to the extent that the Igbo version made little or no sense. This is it:

Kpuru Super Eagles agbachitere Ifeanyi Udeze achoputala Wilfred Ndidi na Victor Moses di ka isi ihe na-aga nke oma 2018 World Cup na 2019 African Cup nke Nations eru mkpota maka Nigeria.

In the Afrikaans version, the same journalistic errors in the headline resulted in the following translation: Udeze: Ndidi, Moses Crucial For Super Eagles AFCON, Wêreldbeker-hoop. This is more of English than Afrikaans, and the person desiring to read the story in Afrikaans loses the chance since the translation failed to deliver meaning in that language. Though the Arabic version succeeded in turning the words to Arabic, meaning is highly compromised in this headline: **ي: زي دوا: ناون عل: مل اعل ساك لام ، زلج ي: ربوسل قمساح يسوم ، يدي دن**

In the same vein, the second story started this way: Hip-hop star Nicki Minaj has promised to pay university tuition for dozens of fans after a promotional contest metamorphosed over Twitter. The translation reads: Hip-hop star Nicki Minaj ya yi alkawarin biya jami'a koyarwa ga dama, magoya bayan wani promotional hamayya metamorphosed akan Twitter. Not only that it retained several English words, "Nicki Minaj ya yi" should have read "Nicki Minaj ta yi".

Other types of journalistic error found in the source language include inappropriate capitalizations, wrong punctuations, wrong spellings and inappropriate spacing. Though the responses supported the literature indicating that inaccuracies could result from inadequate orthography, it was difficult to associate all the non-translated words with this variable alone. This is because the web translating machine seems to have been configured to treat all words beginning with capital letters as names, thereby retaining them in the source language.

DISCUSSION OF FINDINGS

In line with the evaluation framework (El Shiekh and Saleh, 2011) the discussion evaluates the findings with regard to whether the translation is easily understood, fluent and smooth, conveys the literary subtleties of the original, conveys the meaning of the original text, reconstitutes the given text into the target language accurately and captures the style or atmosphere of the original text. It also relates the findings to the literature that shapes the study.

Are the translations fluent, smooth and easily understood? Generally, the translations are not smooth, especially in Shona, Igbo, Afrikaans and Yoruba. In most cases, where non-translated English words are not obstructing the meaning, some words may be translated out of context, rendering many sentences outright transliteration instead of translations. For instance when the words, idol and star, referring to celebrities are translated as *arusi* and *kpakpando* in Igbo, it conjures the idea of a man-made god and a physical star interacting with human beings.

Do the translations convey the literary subtleties of the original? Generally, the answer is No; only a few sentences from the six stories could convey the literary subtleties of the source languages. However, translations to Arabic, Amharic and Swahili were much clearer than translations to the other languages. This is consistent with the level of advancement and development of these languages and echoes the fact that the richness of a language's orthography could enhance web translation of contents written in that language. This is consistent with the findings of previous studies indicating that African-language Internet dictionaries were found to be generally small, and the contents not often of a high quality (Huang, 2011).

Examples of compromised literary subtleties include the translation of the word "breaking". The expressions "break" and "to break" in Yoruba literally denote *fo* and *fifo* respectively. But by semantic extension, they may be used for something new and fresh. However, the most appropriate Yoruba words for something hot, fresh and new are *yajoyajo* or *ajaabale*. Therefore, a translation that would capture the contextual meaning and literary subtleties of the language should read *irohin ni yajoyajo* (*yajoyajo* for short) or *irohin ajaabale* (*ajaabale* for short).

These findings support the observation by Parisa (2015) stating that web translation machines generally exhibit shortcomings with regard to translating less popular languages. In other words, it could be argued that the translation does not reconstitute the given texts into the target languages accurately neither does it capture the style or atmosphere of the original texts. But at same time, it

is unsafe to lay all the blame on the Google translating machine since orthographical inadequacies and journalistic errors could contribute to inaccuracies. The point made by Huang (2011) regarding the shortcomings of African language dictionaries online seems to complement or accentuate the shortcomings of Google translator's handling of the same African languages.

FUTURE RESEARCH DIRECTIONS

1. Experimental studies should be conducted to test-run the predictors of inaccuracy—orthography and journalistic errors—modeled in this study. This would reveal the level of improvement that improvement on these variables would add to the translation process.
2. This study, like many others, examined only Google translation. Attempts to pay attention to other online translation machines could enrich web translation research.

RECOMMENDATIONS

To address some of the concerns raised by the findings, this study recommends:

1. That the challenge of inadequate orthography in the online dictionaries should be urgently addressed by the publishers and their funding partners. Google Corporation's support towards the development of adequate orthography will no doubt facilitate it and lead to mutual benefits in the long run.
2. Newspapers that install the translation application on their websites should take extra care to write, edit and proofread their stories to align them to the application to reduce inaccuracies.
3. The only way to achieve the level of accuracy already attained by Google in the translation of major European languages is to pay more serious attention to African language lexicography with a view to leveraging on web translation innovations since the socio-economic benefits to all parties concerned cannot be over-emphasized.
4. As Google begins to shift away from the statistical approach to grammar-focused translation, success would require close collaboration and constructive engagement among the three parties—Google, newspapers and publishers of online dictionaries of African languages.

CONCLUSION

This study set out to provide answers to three questions on the extent to which Google Translator is able to accurately translate English to the selected African languages, the aspects of the source language difficult for Google Translator and the relationship between translation accuracy and respondents' perception of journalistic errors, inadequate orthography and technological limitations of the translating machine. It accomplished that objective by ascertaining that over 45% of the text was either not translated or translated inaccurately, with 393 English words appearing in the translated text. What the participants got from the translations was partial translation with little or no meaning most of the time. Aspects of the language most affected are newly developed words such as Twitter, tweet, email, and vocabulary like metamorphosis, wrongly punctuated and capitalized words and sentences containing inappropriate spacing of words and letters.

It is perceived that the combined influence of journalistic errors, inadequate orthography and the limitations of Google translate algorithms contributed to the inaccuracies almost to the same degree. The statistical analysis also shows that there is significant relationship between inaccurate web translation and these factors, thereby suggesting that improvement in journalistic writing, orthography and web translation technology predicts improved accuracy.

REFERENCES

- Aiken, M. (2019). An Updated Evaluation of Google Translate Accuracy. *Studies in Linguistics and Literature*, 3(3), 253. doi:10.22158/sll.v3n3p253
- Aiken, M., & Balan, S. (2011). An analysis of Google Translate accuracy. *Translation Journal*, 16(2). Retrieved July 8, 2019, from <https://translationjournal.net/journal/56google.htm>
- Ashraf, F. (2018). *Brian Winston and Media*. Slideshare.net. <https://www.slideshare.net/FathimaAshraf6/brian-winston>
- Balk. (2013). *Assessing the Accuracy of Google Translate to Allow Data Extraction From Trials Published in Non-English Languages*. Agency for Healthcare Research and Quality (US). Report No.12 (13)-EHC145-EF. <https://www.ncbi.nlm.nih.gov/books/NBK121304/>
- Benjamin, M. (2018). *Which languages are Google Translate best at translating?* Kamusi Project International. Retrieved July 8, 2019, from <https://www.quora.com/Which-languages-are-Google-Translate-best-at-translating>
- Casacuberta, F., & Vidal, E. (2007). Learning finite-state models for machine translation. *Machine Learning*, 66(1), 69–91. doi:10.1007/s10994-006-9612-9
- Chen, X., Acosta, S., & Barry, A. (2016). Evaluating the accuracy of Google Translate for diabetes education material. *JMIR Diabetes*, 1(1). 10.2196/diabetes.5848
- El Shiekh, A., & Saleh, M. (2011). Translation versus Transliteration of Religious Terms in Contemporary Islamic Discourse in Western Communities. *International Journal of English Linguistics*, 1(2), 141. <https://d1lwqxts1xzle7.cloudfront.net/5576628/12056-36160-1-SM.pdf>
- Ghasemi, H., & Hashemian, M. (2016). A Comparative Study of Google Translate Translations: An Error Analysis of English-to-Persian and Persian-to-English on of Religious Terms in Contemporary Islamic Translations. *English Language Teaching*, 9(3). doi:10.5539/elt.v9n3p13
- Huang, H. J. (2011). Intermediality and human vs. machine translation. *Comparative Literature and Culture*, 13(3), 1-11. www.google scholar.com
- Jordan, S. (2015). The Challenges of Translating for Newspapers. *Alpha Omega Translations*. <https://alphaomegatranslations.com/business-translation/the-challenges-of-translating-for-newspapers/>
- Kroulek, A. (2016). *11 Google Translate Facts you should know*. Retrieved July 8, 2019, from <https://www.k-international.com/blog/google-translate-facts/>
- Lazzari, G. (2006). From Human Translation to MT: Dramatic Costs Reduction and Accessibility Improvement. *Human Language Technologies for Eroup*.
- Leddy, M. (2019). *How we used Weglot to translate our website to Spanish*. Retrieved from: <https://blog.weglot.com/weglot-translate-english-website-spanish/>
- Lichterman, J. (2016). *European publishers are teaming together to translate the news to reach broader audiences*. <https://www.niemanlab.org/2016/08/european-publishers-are-teaming-together-to-translate-the-news-to-reach-broader-audiences/>
- Lin, D., Murakami, Y., Ishida, T., Murakami, Y., & Tanaka, M. (2010). Composing Human and Machine Translation Services, Language grid for improving localization process. *Proceedings of the 7th International Conference on Language Resources and Evaluation*. <https://waseda.pure.elsevier.com/en/publications/composing-human-and-machine-translation-services-language-grid-fo>
- Lingo, H. (2016). An Evaluation of the Accuracy of Online Translation Systems. *Communications of the IIMA*, 9(4), 69–84.
- Manasse, G. (2017). *The Ultimate List of Website Translation Tools, Platforms and Services*. <https://www.bablic.com/blog/the-ultimate-list-of-website-translation-tools-platforms-and-services/>
- McGuire, N. (2018). How Accurate Is Google Translate in 2018? *Argo Translation*. Retrieved July 8, 2019, from <https://www.argotrans.com/blog/accurate-google-translate-2018/>

- Muller, R. (2009). Google now speaks Afrikaans. *My broadband*. Available at: <http://mybroadband.co.za/news/internet/9395-google-now-speaks-afrikaans.html>
- Munday, J. (2001). *Introducing translation studies: Theories and applications*. Routledge.
- Napitupulu, S. (2017). Analyzing Indonesian-English Abstracts Translation In View of Translation Errors by Google Translate. *International Journal of English Language and Linguistics Research*, 5(2), 15–23. <https://www.eajournals.org/>
- Ned, A. (2020). Machine Translation: The Case of Arabic- English Translation of News Texts. *Theory and Practice in Language Studies*, 10(4), 408–418. doi:10.17507/tpls.1004.09
- Ono, K. (2003). *Translation of news headlines*. https://www.researchgate.net/publication/228819866_Translation_of_news_headlines
- Pan, H. (2016). How BLEU measures translation and why it matters. *Slator*. Retrieved July 8, 2019, from <https://slator.com/technology/how-bleu-measures-translation-and-why-it-matters/>
- Patil, S., & Davies, P. (2014). Use of Google Translate in medical communication: Evaluation of accuracy. *BMJ (Clinical Research Ed.)*, 349(2), 1–3. doi:10.1136/bmj.g7392 PMID:25512386
- Schairer, K. (1999). Improper Conduct, Castro's Cuba. *Hispania*, 82(4), 797. doi:10.2307/346372
- Setiawati, I., Yadnya, I., & Anyawibawa, N. (2020). A Comparison of Word, Above Word, and Grammatical Equivalence of Google and Human Translation Output in Medical-Surgical Nursing Book. *IJRP*, 56(1). Advance online publication. doi:10.47119/IJRP100561720201319
- Si, X., & Wang, J. (2021) Grammatical Metaphor in English-Chinese Translation. *International Journal of Translation, Interpretation, and Applied Linguistics*, 3(1), 1-15. <https://www.igi-global.com/pdf>
- Turner, K. (2016). Google Translate is getting really, really accurate. *Washington Post*. Retrieved July 8, 2019, from: <https://www.washingtonpost.com/news/innovations/wp/2016/10/03/google-translate-is-getting-really-really-accurate/>
- Valdeon, R. (2020). Journalistic translation research goes global: theoretical and methodological considerations five years on. *Studies in Translation Theory and Practice*, 28(3), 325-338. doi:10.1080/0907676X.2020.1723273
- Van-Rensburg, Snyman, O., & Lotz, H. (2012). *Google Translation Toolkit*. Available at: <http://www.translationautomation.com/technology/google-translation-toolkit.html>
- Vincent, J. (2016). *Apple boasts about sales; Google boasts about how good its AI is*. The Verge. Retrieved July 8, 2019, <https://www.theverge.com/2016/10/4/13122406/google-phone-event-stats>
- Wang, F., Humble, P., & Chen, W. (2019). *A Bibliometric Analysis of Translation Studies: A Case Study on Chinese Canon. In The Journey to the West*. LEAN Library. doi:10.1177/2158244019894268
- Wilfred, C., & Ezeru, W. (2021). The Continued Domination of Western Journalists in Global African News Telling: The Imperatives and Implications. *African Journalism Studies*, 42(1), 36–55. doi:10.1080/23743670.2021.1886961
- Zhang, J. (2009). Discourse, Context and Machine Translation Quality. *Shandong Foreign Language Teaching Journal*, 6(1). Retrieved from: <http://cnki.com.cn/>