



Machine Translators and Indigenous Knowledge in Hausa Language: A Case for Google and Facebook Translators

ABIODUN ABOSEDE OSHIN

Adekunle Ajasin University, Akungba Akoko, Ondo State, Nigeria

Abstract. The prominence of technology and its intractable effects on every sector of human life in the world has been acknowledged. In the field of translation, technological advancement has influenced the process of translation. The questions that beg for answers therefore are? How adequate are these technological translation machines? Can these machines transport successfully cultural relics and imports embedded in a given corpus to the target language? These questions provided the impetus for this study. This study examines machine translated corpora and their intelligibility with a view to ascertaining their adequacy and effects on information in Nigerian indigenous languages. Quantitative research method was adopted for this study. Text corpus in Hausa languages were subjected to Google translation machine and Facebook translation system. The study revealed that the translated corpora were fairly intelligible; machine lexical choices were ambiguous; Translation does not recognize proper nouns and so places them arbitrarily. On the basis of these findings, it is recommended, amongst others, that more Hausa words are encrypted into the translation machine. Translation machine should be improved upon to cater for the adequacies detected.

Keywords: Machine Translation, Technology, Translation, Facebook Translation, Google Translation, Hausa Indigenous Knowledge

1. Introduction

Communication is the live wire of any human system. Communication in whatever conduit – verbal, written, gesturing, facial expression or sign is language. Language is the quintessence of human existence; it makes us inimitable and makes a distinction between us other animals. Through language, man perceives and organizes his reality;

expresses and promulgates rules, morals and values that mark them off as distinctive group of people with a definite identity. Thus, language is a fulcrum of social mobilization communities leverages on towards achieving sets of developmental goals.

All languages of the world, Hausa language inclusive, are effective in catalyzing development. It is the depository of a people's identity and way of life and also, the principal network for the transmission of societal belief system and ideologies to other people and evolving generation. Culture, literature, philosophies and the entirety of a people's creativity were communicated to the new generation through information encrypted in signs/symbols/drawings/paintings on walls of caves, in carvings, sculptures, etc. Consequently, the indigenous knowledge shared is concealed from the vast majority; while only a privileged few who share the same linguistic code have access to the information.

Indigenous knowledge, according to Udoh and Anyanwu (2015:3) comprises mutually implicit and explicit knowledge that is "held and made use of by people who regard themselves as indigenous to a particular community". Indigenous knowledge systems greatly depend on indigenous languages for sustenance. Knowledge encrypted in indigenous languages are significantly important to a people; they depend on it for medicinal and herbal needs, food supply, conflict resolution and spiritual growth, inter alia. Udoh et al (2015) further argued that a people's indigenous knowledge helps in information processing, allocation of resources and help in the production methods and decisions of the community in question. In a bid to transmit knowledge in indigenous languages to other people in around the globe, the services of translators come into effect.

The ontological significance of translations has been established among different circles. Translation is “the process of translation can be described as a two-stage process of decoding the source language input (to obtain the meaning) and encoding the meaning thus obtained into outputs in the target language” (Ayogu, Adetunmbi and Ojokoh, 2018:2). Mundy cited in Ezeafulukwe (2018) sees translation as the process of “changing of an originally written text (the source text or ST) in the original verbal language (the source language or SL) into a written (the target text or TT). In a nutshell, translation has to do with passing a message expressed in one language (which could be oral, written or signed language) into another. This is done to disseminate knowledge to other people who may not be familiar with the linguistic code the knowledge was encrypted.

In the last few decades, Information Communication Technology (ICT) and its paraphernalia has played a significant role in translation. Computers now automate some or all of the process of translating from one language to another, unlike what we used to have before now. This development places new demands on translators which necessitate that they adopt new technologies to carry out their day-to-day work. In view of this, therefore, translators are expected to improve themselves in order to meet these demands. The objective of using machine translator is to assist human translators produce output that can be used directly without further processing.

Machine Translation process involves the use of computational systems to decrypt and translate the meanings in a given input sentence from one language – the source language, e.g. Igbo, to the other(s) – the target language(s), e.g. Yorùbá, Hausa, English or Chinese (Ayogu *et al*, 2015). This basically involves two processes – decoding the source language input (to obtain the meaning and encoding the meaning thus obtained into outputs in the target language. It is, however, expedient to note that Machine Translation, as observed overtime, at its basic level usually do not produce a good translation of indigenous knowledge that non-native can run with. This study, therefore, sought to determine the peculiarities that are associated with Facebook Translation and Google Translation of indigenous knowledge in Hausa language, a language spoken in the Northern part of Nigeria.

2. Theoretical Framework

The theoretical framework adopted for this study is computational linguistics. Computational linguistics

is a field between linguistics and computer science. This theory is founded on the basic fact that computers can be used to treat linguistic issues. The term “computational linguistics” which was credited to David Hay evolved in the 1950s when there was a dire need for United States to use computers to automatically translate texts from foreign languages (particularly Russian scientific journals) into English (Nzuake, 2018). It is in the light of this that computational linguistics has been considered as a branch of computer science as well as of linguistics.

Muslim (2007) highlighted two essential motivations behind computational linguistics:

1. **Theoretical:** It came from the progressive thought that adopting computational aids (sic) would cause important progress in linguistics.
2. **Technological:** It came from the sheer need to produce a technology to serve the practical needs for translation, information extraction, grammar checking---etc.

It is expedient to note that none of these projects can be achieved by the adoption of linguistic methods alone. From these points of view, two groups emerged. According to Muslim (2007), one applies their experience in computer science to linguistics while the other apply their experience in linguistics to computer science so that computers can understand everyday human language and translation. In the light of the foregoing, computational linguistics is synonymous to automatic processing of natural language, since its major task is the use of computer programs to process words and texts in natural language.

3. Conceptual Framework

3.1 Indigenous Knowledge

According to Udoh and Anyawu (2015), language is an essential aspect of a community, as well as an important index of identity. It is a conveyor of culture and warehouse of indigenous knowledge on all aspects of life like ecosystem, flora and fauna, traditional medicine, social norms, political and historical traditions. When a language goes extinct, all the accumulated knowledge and experiences of the people equally begin to go into extinction (Dandaura, 2017). UNESCO (2003) recognizes indigenous languages as expression of cultural diversity and its mutual sustainability with indigenous knowledge system and sustainable development. Indigenous languages, basically, are languages that are spoken in particular regions by the

natives of the region over a long period (UNESO 2003) while indigenous knowledge refers to the traditional or folk knowledge that is unique to a particular culture.

Indigenous knowledge includes but not limited to traditional songs, stories, legends, dreams, methods and practices of a people. When there is trouble, people run to it for medicinal and herbal needs, food supply, conflict resolution and spiritual growth, inter alia. For the benefit of others who may not share the same linguistic code to which the knowledge has been coded, machine translation becomes expedient.

3.2 Machine Translation

There is no clanger in the fact that Machine Translation is an indispensable commodity, as the world is titling towards a global village. According to Folajimi and Omonahin (2012), machine translation is extremely significance in view of the growth of international communication – both at government (European Union EU, United Nations Organization UN) and business and commerce levels (exporters need product documentation in the languages of the countries where their products are marketed). For effective communication between people of different language and geographical backgrounds, therefore, machine translation is eminent.

Machine Translation has been variously defined by different scholars. As one of the branches of applied research, it is a conglomeration of ideas and techniques from linguistics, computer science, Artificial Intelligence (AI), translation theory, and statistics. To Nzuanke and Chinaka (2018:40), machine translation which is also known as fully automated machine translation (FAMT), is a process whereby “the source language text is fed into the computer as a file, and the computer produces a translation automatically without any human intervention”. According to Hutchins and Somers (1992:3), Machine Translation (MT) is “computerised systems responsible for the production of translations from one natural language to another, with or without human assistance. From the foregoing, one may conclude that in Machine Translation, the translator has little or no inputs to make in this translation process. All that is required of him/her is to feed into the computer the knowledge in indigenous languages s/he wants to translate while the computer which has been programmed gets the job done.

4. Results and Discussion

Below, excerpts written in Hausa indigenous language were subjected to Google machine translation while some comments in Hausa language were translated using Facebook translation. The peculiarities of these translation machines were accentuated.

Indigenous knowledge in Hausa language and its parallel corpora which has been subjected to Google translation

Except 1:

Hausa Language: Ra'in Bincike

Bincike ko nazari dangane da maganin kowace al'umma abu ne da ya shafi rayuwar al'ummar gaba daya. Bahausha ya gama magana da ya ce: “lafiya uwar jiki.” Idan babu uwa, lallai babu rabin zance domin ko a dara ma “fid da uwa ake.” Yayin da aka dāura damarar nazartar maganin gargajiya na Bahausha, haƙiƙa an tunkari wani babban ginshiki ne a rayuwarsa. Za a iya tabbatar da hakan yayin da aka yi la'akari da cewa, nason magani ya mamaye al'ada da adabin Bahausha baki daya. Duk wani motsinsa, sai ya sanya magani ciki.

Google Translated Version: Search Options

Research into the treatment of any community is an issue that affects the lives of the community as a whole. “**Good morning, mother,**” said Bahausha. If there is no mother, there is no half-talk because even more so is the “mother out.” As the study of Baha'i traditional medicine was under way, it was approached as a major milestone in his life. This can be proved by considering that medicine has dominated Baha'i culture and literature as a whole. Whatever his movements, he put medicine inside.

Comments:

In the Google translated version of the corpora above, the following are noted. Firstly, the phrase “Good morning, mother” which is a parallel version of “lafiya uwar jiki” (health is wealth) in Hausa language was wrongly translated by Google. “lafiya uwar jiki” is an adage in Hausa language which could be literarily translated as Health is wealth. In Hausa cosmology, being healthy is the chief of all things. It was also noticed that “Bahausha ya gama magana da ya ce” was not translated by the Google Translation, probably because it does not recognize it. The expression was completely omitted by Google translation. Furthermore, the expression “If there is no mother, there is no half-talk because even more so is the “mother out” refers to a game of seeds played in Hausa land. The mother seed is said to be the chief among others, once the mother seed is defeated, then the player may likely not win the game. Google

translation could not appropriately contextualize this. It only interpreted the string on the surface level which almost rendered the string unintelligible. Noticeably, the expression “as the study of Baha’i traditional medicine was under way, it was approached as a major milestone in his life” contain some grammatical fillers from Hausa language and this was translated superficially. The author concludes that “whatever an Hausa man does, he mixes medicine” and this has been largely the reason an Hausa man hardly fall sick. In the translated version, Google translation puts it thus “Whatever his movements, he put medicine inside”. Google translation language choice is done indiscreetly and this could hamper the passing of a particular knowledge.

Excerpt 2:

Hausa: Ma’anar Magani

Bunza, (1990: 135) ya bayyana kirdadon samuwar kalmar magani da ya fi karɓuwa a tsakanin masana Hausa a yau. Wannan ra’ayi ya yi daidai da ɗaya daga cikin ra’ayoyin da suka shafi samuwar magani ga Bahaushe, wato ra’ayin gwadawa. Ra’ayin na nuna cewa, kalmar magani ta samu ne a lokacin da Bahaushe ke ƙoƙarin samar wa kansa magungunan cutuka daban-daban. Yayin da cuta ta samu, sai Bahaushe ya tattaro ganyaye da saiwa da sassake-sassake domin gwadawa, har sai ya samu wanda zai samar masa da waraka. Zai dauki nau’in sassake ko saiwa ko ganye (da makamantansu), sai kuma ya ce:

Hausa ► English

An tarki wata sara a Kano ta shan ruwan gishiri da siga yayin sahur ko buɗa baki. Abin ma nema yake ya zama wani salon wayewa. Da an sha ruwa sai ka ga mutane na siyan ruwan roba suna ɗura ORS.

Likitoci dai sun fara gargadin mutane da su kiyaye. Zallar ruwa kawai zai fi gamsar da lafiyar jiki fiye da ruwan gishirin da ake ta sha.

Barka da shan ruwa.

A sara has been trapped in Kano by drinking salt water and cigarette during sahur or opening its mouth. It's about to be a form of civilization. When it rains you see people buying plastic water and carrying ORS.

Doctors have started warning people to be careful. Just water can be more satisfying to the body than the salt water that is being drank.

Happy rainfall.

Comments:

The first sentence was badly translated by Google. Some words, however, are noticed to be strange to Facebook Translation. Hausa word such as “sara” which means what is in the vogue was not recognized by Facebook translation. The translation error is so severe that someone in an outer circle would be left in quandary. The statement should be “It has become a new development in Kano for people to drink salt and sugar water (ORS) during the breaking of a fast. It is also expedient to note that Facebook translation infixed the word “cigarette” to its

“Ma gwada wannan mu gani.” Wato dai za a gwada a gani ko za a dace?

Google Translated Version: The Meaning of Solution

Bunza, (1990: 135) describes the hypothesis of the emergence of the most widely accepted term of medicine among Hausa scholars today. This idea agrees with one of the concepts of medicine in Baha’i, namely the concept of experimentation. The theory goes that the word medicine came about when Baha’u’llah was trying to find a cure for various ailments. When the disease struck, the Baha’is collected leaves, roots and carvings to test, until they found a cure. He would take some kind of carving or root or leaf (and the like), and then say: “Let’s try this and see.” That is to say, it will be tested to see if it will be appropriate?

Comments:

The subtitle – Magani means medicine and not solution as translated by Google. As noticed overtime, the nomenclature “Hausa man” takes different spelling. In the first excerpt, it was coded as “Bahaushe” however, Google translation outputs have consistently switch this with “Baha’u’llah” and “Bahai” respectively. In line 5, Google translation repeatedly made wrong lexical choices through choosing “carvings” over barks of wood. This is observed in line 5 and 6 respectively.

Excerpt 1: Facebook Translation

outputs which was not in the original corpus. However, the error could be traced to the translation of “siga” which means sugar in Hausa language. The Hausa phrase, *dan sha ruwa*” which Facebook Translation has transcribed as “when it **rains**” is wrong. “*Dan sha ruwa*” means, when they (people) break their fast ... It is also important to note that Facebook translation appears to make no lexical choices as words are chosen without due considerations to the context they occur. The selection of “when” and “carrying” respectively is entirely word for the string we just examined. The last phrase which was translated as Happy Rainfall by Google Translation is wrong. It should be Happy Fasting Breaking!

Except 2: Hausa corpora

Facebook Translation

Abin da yake faruwa tsakanin Ukraine da Russia da sauran kasashen Turai, izina ce babba ga daidaiƙunmu. Duk mai tunzura ka ka yi rigima da mutane a social media, ka tabbatar da idan ta kwabe ma zai iya tsaya ma. Kar ka damu da cewar kai ke da gaskiya ko akasin haka. Mutane suna da wayon tsiya, kowa zaman lafiya yake nema. Kar ka dinga rubutun neman fitina ka ga ana ta ma like da kalaman yabo ka yi zaton goyon bayanka ake.

Ga shi nan dai kunu ya tsinke wa Ukraine amma ana ta kawo mata uzirin rashin kasacewarta daya daga cikin NATO.

Kai ma za ta iya kasancewa da kai. Idan giyar 4G ta debe ka ka zagi wani babban mutum ya ci kwalarka a wata jahar daban, sai dai mu karade ko'ina da [#hashtag](#). Idan an ce mu dau mataki, sai mu ce ai don kai ba dan Kano ba ne, ba mu da damar shiga.

See Translation

What's happening between Ukraine, Russia and other European countries, is a big deal to our altitude. Whoever wants you to argue with people on social media, make sure that if it's too late he can stand for you. Don't worry if you're right or wrong. People are so smart, everyone wants peace. Don't be posting for trouble and get likes and compliments and assuming that they support you.

Here it is, Kunu has broken against Ukraine but she is being excused for the loss of one of NATO.

It could be you too. If 4G beer makes you insult a big guy to eat your brain in a different state, we just add everywhere with a [#hashtag](#). If we are asked to take a step, we will say because you are not from Kano, we have no chance to enter.

Comments:

The second sentence in the first paragraph which reads “whoever wants you to argue with people on social media, make sure that “if it’s too late he can stand for you” was translated at the surface level. The phrase “if it is too late” explains a situation where the purported supporters can advocate for one in the case where such arguments lands one into trouble. In the beginning of second paragraph, the translated Facebook version was completely unintelligible and meaning may be hampered. As noticed from the outputs, the author created an analogy of what is happening between Ukrain and Russia using “Kunu”, an indigenous food in Hausa land which when not monitored may go awry. The implied meaning of the expression “Kunu has broken against Ukrain but she is being excused for the loss of one of NATO” is that the case before Ukraine has become deteriorated because she is not a member of NATO. Again, the last paragraph was erroneously by Facebook Translator. So many words were mistranslated by

Facebook translator and this rendered the paragraph useless as one cannot make any meaning out of it. A Hausa word, such as “kwalarka” which is a shirt collar and “kwakwalwa” means brain. The translator only has “kwakwalwa” in its repertoire; hence, it translated “kwalarka” as brain. This is seen in this excerpt “if you drink 4G beer and you insult a big guy to eat your **brain** in another state”. Noticeably, Facebook translator algorithm could not detect the meaning of “kwalarka”, so it replaces it with brain which it has in its word economy. Implied in the paragraph, was that insulting someone of influence lands one into trouble and those him/her to support others may not be able to defend him/her than to make an hashtag posts on social media. Again, Facebook translator indiscriminately makes odd lexical choices which may not represent what the author has in mind. For example, the sentence “we would just add everywhere with hashtag” does not reflect the through intention of the author. Perhaps,

the word “add” and the structure of the output made mess of the discourse in view.

5. Conclusion

The significance importance of Machine Translator has been acknowledged in all circles. As the world tilts towards global village, there is increasing need to breaking the various language barriers which have confined us from growing as we would desire. This study revealed that Google Translator can handle complex corpus, however, cannot translate adages, maxim, clichés. Google translator does not recognize context and often, omit any expression it does not recognize. In the same vein, Facebook Translator cannot handle complex corpus. Hausa expressions in adages, maxims, clichés, and proverbs amongst others were miserably translated that the reader may be totally confused as to what the discourse dwell on. Noticeably, from the translation outputs, Google Translator and Facebook Translator operate on different algorithms. For Facebook translator, the web crawler searches for words in its repertoire first, if the word is not in its bank, it replaces that word with another word that has similar orthography.

6. Recommendations

On the basis of the findings of this study, the study recommends that:

- Google and Facebook translators improve on their services
- Google and Facebook translators feed their machines with Hausa adages, maxims, clichés and proverbs etc
- Google and Facebook translators consider how to integrate context into their operating system
- Google and Facebook translators reconfigure their operating system to translate words beyond the surface level

References

- Ayogu I. I. & Adetunmbi A. O. (2018) Developing Statistical Machine Translation System for English and Nigerian Languages, *Asian Journal of Research in Computer Science*1(4): 1-8, 2018; Article No.AJRCOS.44217
- Emmanuel S. Dandaura (2017). Safeguarding Indigenous Nigerian Languages from Extinction: The Strategic Role of Indigenous Language Newspapers, a Paper Presented at the Two-Day Workshop on Indigenous

Language Newspapers and National Development, Organized by the National Institute for Cultural Orientation (NICO), Federal Ministry of Information and Culture, Abuja

- Ezeafulukwe (2018) Translation: A Panacea to Nigeria’s Social Problems, Apostolate of Language and Linguistics: A Festschrift in Honour of Professor Emmanuel Okonkwo,
- Folajimi O., & Omonayin I., (2012). Using Statistical Machine Translation (SMT) as a Language Translation Tool for Understanding Yoruba Language, 2nd Intl’ Conf.Comp., Energy, Net., Robotics and Telecom
- Hutchins, J. W. and Somers, H. L. (1992). An Introduction to Machine Translation. London: Academic Press.
- Muslim E. (2007). An Introduction to Computational Linguistics: Advantages and Disadvantages, *Journal of College of B .ED., NO.51* 92
- Nzuanke S., & Chinaka U., (2018). Technology and Translation: Areas of Convergence and Divergence between Machine Translation and Computer-Assisted Translation, *Journal of Languages, Linguistics and Literary Studies (JOLLS)* Volume 5. June 2018 ISSN: 2536-6300
- Udoh I., & Anyanwu O., (2015). Protection and Promotion of Indigenous Nigerian Languages as Utility Vessels of Indigenous Knowledge Systems (IKS), *Awka Journal of Linguistics and Languages (AJILL)* Vol. 9.