International Journal of Computer Applications (0975 - 8887)
Volume 185 - No.33, September 2023

An Exploratory Study of Stacked Multilingual SMT Systems for Low Resource Languages

Ikechukwu Ignatius Ayogu
Department of Computer Science
Federal University of Technology
Owerri, Imo State, Nigeria

ABSTRACT

The indigenous capacity for the development of computational linguistics tools for Nigerian languages is yet low compared to what has been achieved in other multi-ethno-linguistic nations such as India. Effective communication among Nigerian citizens of different tongues, and who are unable to use English has been continuously hampered. Thus the need to inter-translate Nigerian languages has become increasingly urgent. Though machine translation (MT) research has achieved state-of-the-art for English and some few privileged languages of the world, the lack of datasets for many Nigerian languages further increases the difficulty of developing MT systems for them. This paper proposes a model for rapidly developing MT system for a new language in a multilingual setup. The overall aim of this research is to establish a scalable platform for the continuous development of MT systems for Nigerian languages using English language as a pivot language. For ease of adaptation and inclusion of a new language, purely data-driven approaches that carefully avoids absolute dependence on the availability of linguistic expertise is adopted. This paper presents a multilingual translation system for English, Igbo, and Yorùbá language mix. Using a research dataset, an overall best BLEU score of 35.62 was obtained for the English-Igbo system, 32.10 for English-Yorùbá system, and 21.03 for Igbo-Yorùbá. These results are encouraging, given the size of the training corpora used.

General Terms

1st General Term, 2nd General Term

Keywords

Multilingual Machine Translation; SMT; Parallel corpora; Low resource languages; Nigerian languages

1. INTRODUCTION

The rich multi-ethnic and socio-cultural diversity of the Nigerian society makes the existence of language barrier inevitable, especially between Nigerians of different ethnic nationalities who lack communication skills in English, Nigeria's official language [1, 22, 21] . One of the primary purposes for which early machine translation (MT) researches were conceived was to eliminate communication barriers among people of different tongues. MT in the modern era has attained much more relevance than simply bridging communication gap between citizens [23]. The roles of MT in national development is clear and evident: it is a tool for national integration and socio-economic well-being; it is extremely useful in security with great potentials to aid counter terrorism efforts; it plays key roles in linguistic and cultural revitalization; it can be applied over social network platforms for language promotion and covert intelligence gathering; it has capacity to improve potentials in business networking and tourism [2, 13, 14, 15, 3, 24].

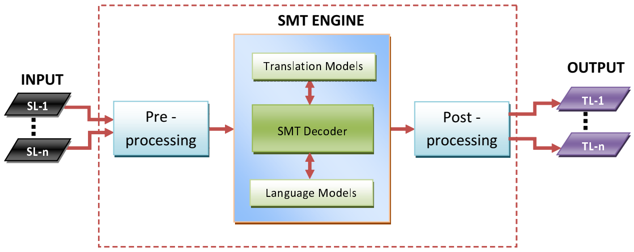

Fig. 1: Structural components of the proposed MT System

Sadly, indigenous capacity for the development and use translation and other language technology tools are still quite low. This research was therefore conceived fill this gap. At present, a prototype MT systems for English-Igbo, English-Yorùbá, and Igbo-Yorùbá translation has been developed.

The multilinual approach is empoyed because it has a potential to fast-track the development of MT for Nigerian languages and English language using scalable, language independent data-driven framework, described in section **2**. This research is focused on translating betweenn Nigerian languages. The objective is to develop a framework for rapid development and deployment of translation tools for Nigerian languages using data-driven, language-independent techniques.

2. THE PROPOSED ARCHITECTURE

The architecture being proposed is a stacked multilingual SMT-based translation engine which maintains a link between the optimal parameters for each of the language pairs in the mix. The model is trained and optimized for each language pair separately. During translation, the appropriate model is selected based on the input language. As shown in Figure 1, the input and output can be from any pair of languages that has been covered. The translation engine is equipped with sub-systems that has been pre-trained to handle preprocessing, post-processing and translation tasks.

3. THE SMT APPROACH

In the Phrase-based SMT approach, input sequences are grouped into word sequences that are technically refered to as 'phrases'.

The notion of phrase as used here does not imply grammatical phrases; it is a mere treatment of word sequences which are translated as a unit. Prior to the discovery of the phrase-based approach, translation probabilities were conditioned on words [11, 10, 12]. Research has however, shown that conditioning translation probabilities on sequences of words produces a better result. In the phrase-based model approach, the source language words are first grouped into sequences of words: s_1,s_2, s_3,..., s_l, then each phrase, s_i, is translated to the corresponding target phrase, t_i; the target phrase is reordered, if necessary, using a language model of the target language. The pieces of the target phrases are then combined into a sentencen using the distortion probabilities. The phrase-based model is powered by two core probabilistic model components - the translation probability, $\phi(s_i|t_i)$ and the distortion probability $d(a_i - b_{i-1})$. a_i = the start index of the source phrase generated by the ith target phrase; b_i = the end position of the source phrase produced by the $i-1th$ target phrase t_{i-1}.

The distortion is a function that measure the relative distance between the positions of a phrase in the languages that is meant to moderate the model's decision for large distortion by imposing heavy penalty on large distortions [8]. From these, the translation model of a phrase-based SMT is therefore represented as in Equation

$$p(s|t) = \prod_{i=1}^{l} \phi(s_i, t_i) d(a_i - b_{i-1}) \tag{1}$$

Decoding in phrase-based SMT is a search problem which is focused on parameter estimation of the translation probabilities. In the process of decoding, the translation probabilities, $\Phi(si, ti)$, is estimated. This set of parameter is learned from phrase alignment probabilities. Alignments can be one-to-one as in IBM model 1, many-to-one or even zero-to-one as in IBM model 3 but phrase-level alignments cannot be handled by the IBM models 1-4 or HMM word alignment algorithms since words are treated independently by these models [8]. Early augmentations to these models for phrase-level alignment can be found in Och and Ney [19], Och [16], Och and Ney [18] and Koehn [11].

4. MATERIALS AND METHODOLOGY

4.1 Materials

The raw data were obtained for Igbo and Yorùbá languages, together with their corresponding English language versions were obtained from jw.org, taking several weeks. The raw corpus was then prepared, following the procedures specified in [9].

Sentence alignment was performed using a re-implementation of the Gale and Church character-based alignment algorithm [6]. This gave rise to the parallel corpora; a sample parallel paragraph is shown in Table 1.

4.2 Methodology

This paper adopts the phrase-based statistical machine translation (Pb-SMT) approach first described by Koehn et al in 2003. It is an empirical, data-driven, language-independent methodology. Developed at IBM Watson research center in the late 1980s [4], SMT and its extensions quickly grew into a global first choice MT development technique and although attention is currrently focused on deep learning-driven approaches, Pb-SMT and its extension are still relevant [2, 13, 14, 15, 3, 24]. The Pb-SMT model extends the noisy-channel model-inspired word based models for SMT. The extensions allow the Pb-SMT method to capture more context, thus overcoming the shortcomings of the word-based model [11]. The Pb-SMT modeling centers around two key component factors: a language model (LM), with probability distribution $p(t)$ and a translation model (TM), with a probability distribution $p(s|t)$. Translation is performed through a search for the target sentence t using the Bayesian inference model in equation 2.

Table 1. : A sample of parallel corpora

English	**Igbo**	**Yorùbá**
Do you know what a leper is? A leper is a person who has a sickness called leprosy. That sickness can even cause some of the person's flesh to fall off. When Jesus lived on earth, lepers had to live away from other people. And if a leper saw another person coming, he had to call out to warn that person to stay away from him. This was done so that other people would not get too close and maybe get the leper's sickness. Jesus was very kind to lepers.	Ǹjẹ́ o mọ ẹni tí à ń pè ní adẹ́tẹ̀? adẹ́tẹ̀ ni ẹni tó ní àìsàn kan tó ń jẹ́ ẹ̀tẹ̀. Àìsàn yẹn tiẹ̀ lè mú kí apá kan ara èèyàn gé kúrò. Ní ìgbà àtijọ́ tí Jésù wà lórí ilẹ̀ ayé, àwọn adẹ́tẹ̀ kì í gbé pẹ̀lú àwọn èèyàn ní àárín ìlú, wọ́n máa ń gbé lọ́tọ̀ ni. Bí adẹ́tẹ̀ bá sì rí èèyàn kan tò ń bọ̀ lọ́dọ́ rẹ̀, ó ní láti tètè sọ fún ẹni náà pé kí ó dúró sọ́hùn-ún kí ó má ṣe dé ọ̀dọ̀ òun. Wọ́n máa ń ṣe èyí torí kí àwọn èèyàn má ṣe sún mọ́ wọn kị́ àìsàn ètè náà má bàa ràn wọ́n. Jésù máa ń ṣàánú àwọn adẹ́tẹ̀ gan-an.	Ị ma ihe bụ́ ekpenta? Ekpenta bụ ọrịa nke pụrụ ọbụna ime ka anụ ahụ́ mmadụ na-adapụ adapụ. N'oge Jizọs biri n'ụwa, ndị ekpenta na-ebi ebe dịpụrụ adịpụ site n'ebe ndị ọzọ bi. Ọ bụrụkwa na onye ekpenta ahụ ka mmadụ na-abịa, ọ ga na-eti mkpu iji dọọ onye ahụ aka ná ntị ka ọ ghara ịbịa ya nso. A na-eme nke a ka ndị ọzọ wee ghara ịbịaru ha nso nke ukwuu ma eleghịkwa anya bute ọrịa onye ekpenta ahụ. Jizọs nwere obiọma dị ukwuu n'ebe ndị ekpenta nọ.

$$\underset{t}{\text{argmax}}\ p(t|s) = argmax_t \frac{p(s|t)p(t)}{p(s)} \tag{2}$$

The $p(s)$ term in equation 2 is constant for the best translation and hence is ignored to allow for model tractability. The likelihood $p(s|t)$ in equation 2 is obtained by splitting s into I phrases, $\bar{s}_i$, each corresponding to the target phrases, $\bar{t_i}$; $i = 1, 2, \ldots, I$. Combining TM with a distortion parameter d which controls the limit of re-ordering, equation 3 is obtained:

$$p(\bar{s}_1^I|\bar{t}_1^I) \;=\; \prod_{i=1}^{I} \phi(\bar{s}_i|\bar{t}_i) d(a_i - b_{i-1} - 1) \tag{3}$$

where ϕ is the translation model function, a and b are the start of phrase and end of phrase i being translated respectively. Optimal distortion limits of were experimentally determined to be 6 for English-Igbo, 6 for Englis-Yorùbá and 5 for Igbo-Yorùbá systems. The interpolated trigram language model (over both part-of-speech and surface word forms) was used for all the experiments carried in this research. For this configuration, the LM probabilities were estimated using the maximum likelihood technique. Combining the TM, LM and re-ordering model with weights λ_ϕ, λ_d, λ_{LM} respectively, derives the phrase-based SMT model of the best translation, t_{best}:

$$t_{best} \;=\; \underset{t}{\text{argmax}} \prod_{i=1}^{I} \phi(\bar{s}_i|\bar{t}_i) d(a_i - b_{i-1} - 1) \prod_{i=1}^{|t|} p(t_i|t_{i-1}, t_{i-2}) \tag{4}$$

In the implementations, the model components were weighted according to the responses of the system, given a component following equation 4.

$$t_{best} \;=\; \underset{t}{\text{argmax}} \prod_{i=1}^{I} \phi(\bar{s}_i|\bar{t}_i)^{\lambda_\phi} d(start_i - end_i - 1)^{\lambda_d} \prod_{i=1}^{|t|} p(t|t_i...t_{i-1})^{\lambda_{LM}} \tag{5}$$

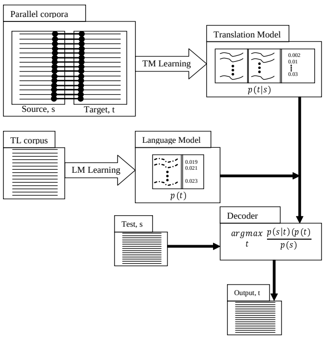

Fig. 2: The key processes in the phrase-based SMT approach

	Train	Tune	Test
$en \rightarrow ig$	en:ig	en:ig	en:ig
NoS	32k:32k	3.9k:3.9k	787:787
NoW	672k:695k	67k:74k	16k:17k
Voc	15k:17k	4.97k:5k	1.95k:1.99k
ASL	21:21.7	20.4:21.9	16.7:18.5
$en \rightarrow yo$	en:yo	en:yo	en:yo
NoS	36.9k:36.9k	4k:4k	1.2k:1.2k
NoW	622k:802k	67k:82k	17k:22k
Voc	15.6k:11.5k	4.8k:3.8k	2.1k:1.7k
ASL	16.9:21.7	16.8:20.3	13.8:17.3
$ig \rightarrow yo$	ig:yo	ig:yo	ig:yo
NoS	32k:32k	4k:4k	800:800
NoW	683k:801k	79k:89k	15k:17k
Voc	17.0k:11.3k	5.2k:4.2k	2.1k:1.7k
ASL	21.4:25.0	19.9:22.4	18.8:21.8

Table 2. : Data used for training, tuning and testing.

5. EXPERIMENTS

5.1 Baseline Systems

This paper used the standard procedure for training, tuning and testing for a typical PB-SMT [11] and Huck and Birch [7]. Data pre-processing was done painstakingly to ensure that obvious non-correspondences in the parallel corpus were taken out. The parallel corpus used in this research has not been described in any other work; it was created as part of this research. The compositions of the corpora were purely based on the researcher's decisions and judgements. Specifically, Huck and Birch [7] approach was adopted for the creation of word alignments and the subsequent extraction of bilingual phrases.
The language model was trained using IRSTLM [5], utilizing smoothing techniques based on modified Kneser-Ney smoothing. Translation options were scored using Good-Turing method. To reduce model errors, and maximize BLEU [20] scores, feature weights (model parameters) were tuned using MERT [17] on the development data, relying on a 100-best list of candidate translations. The details of the data set for training, tuning and testing is presented in Table 2[1]. The basic features used in the experiments are log-probabilities, for both phrase translations and lexical probability translations, word penalty, phrase penalty, distortion penalty. Phrases of length 5 were extracted and a distortion limit of 6 was set. The BLEU measure was used to evaluate the performance of the systems.

The preliminary study reported in this paper investigated the performance of the baseline systems: $en \rightarrow ig$, $en \rightarrow yo$ and $ig \rightarrow yo$ languages that was built. In all the experiments, language models (LM) and translation models (TM) are trained using identical settings. The data used were adjusted to close ranges in terms of number of sentences (Table 1). Table 2 shows the performances of the three systems in terms of BLEU score. The statistics on the translated corpora is presented in Table 3. Furthermore, error analyses of the translated outputs were carried out with the aim of studying the failures of the systems in order to understand the causes of failures for the purposes of designing guided improvements to the baseline systems. The BLEU scores attained by the $en \rightarrow ig$ and $en \rightarrow yo$ systems are comparable; $ig \rightarrow yo$ system has a rather low BLEU score compared to the first two systems, this is despite the fact that the two languages share common characteristics for being in the same language family. The target of this ongoing work is to improve on the quality and quantity of parallel data, enrich the baseline system with extra features from linguistic sources and include advance features of the Moses system.

[1]NoS - Number of Sentences, NoW - Number of Words, Voc - Vocabulary size, ASL - Average Sentence Length (words)

System	BLEU
$en \rightarrow ig$	35.62
$en \rightarrow yo$	32.31
$ig \rightarrow yo$	21.03

Table 3. : The BLEU scores for the experimental systems.

	$en \rightarrow ig$	$en \rightarrow yo$	$ig \rightarrow yo$
	in:out	in:out	in:out
NoS	787:787	1282:1282	800:800
NoW	16k:17.9k	17.7k:30.7k	15k:24.5k
Voc	1948:1654	2117:1416	2053:1459
ASL	20.38:22.83	13.84:24.01	18.83:30.69

Table 4. : Statistics on the test (in) vs translated (out) corpora.

6. SYNTACTIC DIFFERENCES BETWEEN THE THREE LANGUAGES

English, Igbo, and Yoruba languages share an SVO word ordering. However, English language differ from these two Nigerian languages by being left-branching.English language is significantly different from these two Nigerian languages, nothwithstanding similarity in word order. There are more adjectives in English language than any of Igbo and Yoruba languages, and this creates a deficiency that these languages try to address through the use of descriptive constructs, and complementation. English is stable with respect to the positioning of adjectives and determiners relative to the position of the head noun, Igbo is fluid; much more than it is for

Yoruba. English is predominantly preposed while Igbo and Yoruba languages fluctuate between pre-positioning and post-positioning of adjectives, demonstratives, determiners quantifiers, and genitives depending on the construction [22, 21, 1]. Morphological distinctions are also exist among these languages. Morphological processes are centred around the verb for Igbo and Yoruba languages, unlike English language. Plurality of count nouns is marked by the addition a morph, +s in English, for instance; remakeably, Igbo and Yoruba do not have a way to mark plurality of count nouns directly. Rather this is achieved by post-positioning of ordinal or cardinal numbers, depending on the construction, after the noun. Plurals are formed in Yoruba and Igbo by placing some appropriate plural marker before the noun. The possessive s marker is absent in Igbo and Yoruba languages. Other subtle differences exists but the discussion shall be limited to these which are of most important relevance to the presentation in this paper.

7. RESULTS AND DISCUSSION

Preliminary assessments of the prototype MT systems on an experimental test set indicates a good prospects for the research. A BLEU (Papineni et al, 2002) score of 30.73, 30.04, and 19.49 were respectively attained by the best English-Igbo (S-1), English-Yorùbá (S-2) and Igbo-Yorùbá (S-3) experimental systems. These performances are not excellent per se, but they represent a good achievement, given that the research work had just started and currently at an early stage. In Tables 6 and 7 , the performance of the English-Igbo and the English-Yorùbá system perform on the sentences in Table 5 are shown.

Table 5. : Sample Test Sentence

SN	Test Sentence
1.	also, read the book with your mother or father.
2.	but then the Bible explains how we can live forever in God's new world.
3.	he said: Keep praying that your flight may not occur in wintertime.
4.	so do you understand what the sign means?
5.	let's see what the Bible says happened.
6.	now, remember, Jesus said that what happened to those people is a lesson for us today.

Table 6. : Sample translation outputs: S-1

Reference Translation	Translation by S-1
1: Ọzọkwa, gị na nne gị ma ọ bụ nna gị na-agụkọ akwụkwọ a ọnụ. 2: Ma, Bible na-akọkwa otú anyị pụ rụ isi dị ndụ ebighị ebi n'ime ụwa ọhụrụ Chineke. 3: Ọ sịrị: "Nọgidenụ na-ekpe ekpere ka mgbapụ unu ghara iwere ọnọdụ n'oge oyi."	1: Ịgụ akwụkwọ ahụ, ya na nne gị ma ọ bụ nna gị. - (0.45) 2: Ma, Bible na-akọwa otú anyị pụrụ isi ndụ ruo mgbe ebighị ebi n'ụwa ọhụrụ Chineke. - (0.25) 3: Ọ sịrdi : "Nọgidenụ na-ekpe ekpere ka mgbapụ unu ghara ime n'oge oyi." - (0.73)

Table 7. : Sample translation outputs: S-2

Reference Translation	Translation by S-2
4: Nítorí náà, ǹjẹ́ o mọ ohun tí àmì yẹn túmọ̀ sí? 5: Jẹ́ kí á wo ohun tí Bíbélì sọ pé ó ṣẹlẹ̀. 6: Wàyí o, rántí pé Jésù sọ pé ohun tó ṣẹlẹ̀ sí àwọn èèyàn náà jẹ́ ẹ̀kọ́ fún wa lónìí.	4: Nítorí náà, ǹjẹ́ o mọ ohun tí àmì túmọ̀ sí? - (0.81) 5: Jẹ́ kí á wo ohun tí Bíbélì sọ ṣẹlẹ̀. - (0.71) 6: Wàyí o, rántí o, Jésù sọ pé ohun tó ṣẹlẹ̀ sí 'awọn ènìyàn náà ńkọ́ fún wa lónìí. - (0.54)

7.1 Brief Comments on the Performances

Test case 1 is a typical ambiguous sentence. This ambiguity arises from the prepositional phrase attachment marked by 'with' in the input sentence. S-1 tries to translate this very short segment with noticeable flaws. System S-1 though with a sentence-level BLEU score of 0.45, has a poor word order. System S-1 also completely throws away the word 'also'. The high BLEU score for system S-1 is a manifestation of the typical weaknesses of the BLEU metric at the sentence and subsentence levels; the output barely makes sense. Test case 2 presents some interesting scenarios. First it can be seen that the two systems produce quality translations of the input using different alternatives forms that are perfect translations of the input and each other. The low BLEU scores are not surprising given that n-gram matches of the two translations as compared to the reference translation would be mostly zeros. System S-1 makes an error of omission, omitting *dị* in-between *isi* and *ndụ*. In test case 3, a case of an error due to confusion of senses committed by the human translators being carried into the models of system S-1 is noted. Errors in reference translations affects the BLEU score points falsely.

In case 3, System S-2 failed to account for 'the' on the source side. In the case 4, the sequence 'says happened' is clearly problematic for the systems - it commits errors of omission. Case 5 shows an occurrence of mis-selection due to tense; the form ńkọ́ (the act of continuous learning) rather than ẹ̀kọ́ (the nominal).

8. CONCLUSION AND FUTURE WORK

This research paper described a scalable stacked model for performning machine translation on low resourced languages. It is built on the phrase-based machine translation methodology and facilitates speedy development of machine translation engines by allowing the inclusion of more langauages into the stack as the data become available. The results of a multilingual translation experiments carried out for Yorùbá, Igbo and Engish languages show that the approach is promising. Future improvements are targeted improving the performance of the system by incorporating neural machine translation engine in an ensemble framework to leverage the power of representational learning.

9. REFERENCES

[1] B. Adelabu. A contrastive analysis of adjectives in english and yoruba. *International Journal of Educational Research*, 2(4):509–524, 2014.

[2] M. Artetxe, G. Labaka, and E. Agirre. Unsupervised statistical machine translation. *rXiv preprint arXiv:1809.01272*, 2(4):509–524, 2018.

[3] L. Benkova, D. Munkova, L. Benko, and M. Munk. Evaluation of english–slovak neural and statistical machine translation. *Applied Sciences*, 7(11):2948–2, 2021.

[4] P. Brown, J. Cocken, S. Della-Pietra, V. Della-Pietra, F. Jelinek, R. Mercer, and P. Roossin. A statistical approach to language translation. in:. *COLING 1988: Proceedings of the 12th International Conference on Computational Linguistics, Budapest, Hungary*, pages 71–76, 1988.

[5] M. Federico, N. Bertoldi, and M. Cettolo. Irstlm: an open source toolkit for handling large scale language models. In *Interspeech*, pages 1618–1621, 2008.

[6] W. A. Gale and K. W. Church. A program for aligning sentences in bilingual corpora. *Computational Linguistics*, 1(19):75–102, 1993.

[7] M. Huck and A. Birch. The edinburgh machine translation systems for iwslt 2015. in:. *In Proceedings of the International Workshop on Spoken Language Translation. Da Nang, Vietnam*, pages 31–38, 2015.

[8] D. Jurafsky. *Speech and Language Processing. 2nd Edition*. Prentice Hall, 2008.

[9] P. Koehn. Europal: A parallel corpus for statistical machine translation. *In Proceedings of the tenth machine translation summit, Phuket, Thailand.*, 2005.

[10] P. Koehn, H. Hoang, A. Birch, C. Callison-Burch, M. Federico, N. Bertoldi, B.e Cowan, W. Shen, C. Moran, R. Zens, C. Dyer, O. Bojar, A. Constantin, and E. Herbst. Moses: Open source toolkit for statistical machine translation. in:. *ACL-2007: Procedings of the 45th Annual Meeting of the Association for Computational Liguistics, Prague, Czech Republic*, pages 177–180, 2007.

[11] P. Koehn, F. J. Och, and D. Marcu. Statistical phrase-based translation. in:. *HLT-NAACL 2003: Proceedings of the Joint Human Language Technology Conference and the Annual Meeting of the North American Chapter of the Association for Computational Linguistics, Edmonton, AB, Canada*, pages 127–133, 2003.

[12] Philip Koehn. *Statistical Machine Translation*, volume 1. Prentice-Hall, Englewood Cliffs, NJ, 2010.

[13] B. Marie, H. Sun, R. Wang, K. Chen, A. Fujita, M. Utiyama, and E. Sumita. Nict's unsupervised neural and statistical machine translation systems for the wmt19 news translation task. In *In Proceedings of the Fourth Conference on Machine Translation*, pages 294–301, 2019.

[14] M. Martindale, M. Carpuat, K. Duh, and P. McNamee. Identifying fluently inadequate output in neural and statistical machine translation. In *In Proceedings of Machine Translation Summit XVII: Research Track*, pages 233–243, 2019.

[15] M. S. Maucec and J. Brest. Slavic languages in phrase-based statistical machine translation: a survey. *Artificial intelligence review*, 1(51):77–117, 2019.

[16] F. J. Och. *Statistical machine translation: from single-word models to alignment templates*. PhD thesis, Bibliothek der RWTH Aachen, 2002.

[17] F. J. Och. Minimum error rate training in statistical machine translation. in:. *ACL-2003: 41st annual meeting of the association for computational linguistics, Sapporo, Japan*, pages 160–167, 2003.

[18] F. J. Och and H. Ney. A comparison of alignment models for statistical machine translation. in:. *ACL-00:Proceedings of the 18th conference on Computational linguistics-Volume 2*, pages 1086–109, 2000.

[19] F. J. Och and H. Ney. Discriminative training and maximum entropy models for statistical machine translation. in:. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, Philadelphia*, pages 295–302, 2002.

[20] K. Papineni, S. Roukus, T. Ward, and W. J. Zhu. Bleu: a method for automatic evaluation of machine translation. in:. *ACL-2002: 40th Annual meeting of the association for computational linguistics, Philadelphia*, pages 311–318, 2002.

[21] A. U. Umeodinka. Corpus-based contrastive analysis of igbo and english adjectives. *journal of igbo language and linguistics*, (3):54–62, 2011.

[22] B. C. Uzoigwe. A contrastive analysis of igbo and english determiner phrases. *journal of igbo language and linguistics*, (3):73–83, 2011.

[23] L. N. Vieira, M. O'Hagan, and C. O'Sullivan. Understanding the societal impacts of machine translation: a critical review of the literature on medical and legal use cases. *Information, Communication and Society*, 24(11):1515–1532, 2021.

[24] H. Wang, H. Wu, Z. He, L. Huang, and K. W. Church. Progress in machine translation. *Engineering*, 7(11):2948–2, 2021.