big data and cognitive computing

MDPI

Article

Twi Machine Translation

Frederick Gyasi and Tim Schlippe *

IU International University of Applied Sciences, 99084 Erfurt, Germany
* Correspondence: tim.schlippe@iu.org

Abstract: French is a strategically and economically important language in the regions where the African language Twi is spoken. However, only a very small proportion of Twi speakers in Ghana speak French. The development of a Twi–French parallel corpus and corresponding machine translation applications would provide various advantages, including stimulating trade and job creation, supporting the Ghanaian diaspora in French-speaking nations, assisting French-speaking tourists and immigrants seeking medical care in Ghana, and facilitating numerous downstream natural language processing tasks. Since there are hardly any machine translation systems or parallel corpora between Twi and French that cover a modern and versatile vocabulary, our goal was to extend a modern Twi–English corpus with French and develop machine translation systems between Twi and French: Consequently, in this paper, we present our Twi–French corpus of 10,708 parallel sentences. Furthermore, we describe our machine translation experiments with this corpus. We investigated direct machine translation and cascading systems that use English as a pivot language. Our best Twi–French system is a direct state-of-the-art transformer-based machine translation system that achieves a BLEU score of 0.76. Our best French–Twi system, which is a cascading system that uses English as a pivot language, results in a BLEU score of 0.81. Both systems are fine tuned with our corpus, and our French–Twi system even slightly outperforms Google Translate on our test set by 7% relative.

Keywords: Akuapem Twi; machine translation; low-resource language; Twi–French

check for updates

Citation: Gyasi, F.; Schlippe, T. Twi Machine Translation. *Big Data Cogn. Comput.* **2023**, *7*, 114. https://doi.org/10.3390/bdcc7020114

Academic Editor: Alberto Fernandez Hilario

Received: 4 May 2023
Revised: 16 May 2023
Accepted: 29 May 2023
Published: 8 June 2023

1. Introduction

In recent years, machine translation systems have played a key role in communication by removing language barriers [1]. Google's Neural Machine Translation System [2], for instance, is a multilingual machine translation system, which handles translations of over 100 language pairs. The need for machine translation services has expanded in recent years due to the massive interchange of information across different regions using multiple regional languages [3]. Companies operating in numerous countries throughout the world use machine translation services for a variety of purposes, including internal and external communication, client interaction on a global scale, and more [4]. Moreover, people all around the world are now able to communicate in a variety of languages on social media platforms because of machine translation systems [5]. Furthermore, machine translation has shown considerable potential in terms of revolutionizing foreign language teaching and other applications in the field of education [6–9], and research also demonstrates that machine translation has increased international trading [10].

Despite these numerous benefits, machine translation is not available or has not produced desirable results in several indigenous African languages [11], as compared to the state-of-the-art results achieved with high-resource languages, such as English, Spanish, and French [12–15]. High-resource languages are those that have a large volume of digitalized text [16]. One major reason for the shortfall of machine translation in most African languages is the lack of parallel corpora for these languages [17]. Parallel corpora are required for training machine translation systems, and the performance of statistical machine translation and neural machine translation systems is directly impacted by the number of parallel sentence pairs available for training [18]. Languages with a large volume of parallel corpora available, such as English and French, are spoken globally and have established themselves in many regions as a mode of education and other activities related to communication [19]. Having machine translation services from an indigenous African language into one of these globally recognized languages has become an essential tool not only for improving communication between the rest of the world and language speakers, but also for assisting the rest of the world in learning about the people's culture [20].

Ghana is a country in West Africa with over 75 indigenous languages [21]. The Akan Twi language is the most widely spoken language, with about 80% of the country speaking it as their first or second language [21]. Despite Ghana sharing common borders with Togo, Burkina Faso and Côte d'Ivoire, whose official languages are French [22], only an estimated 1–5% of Ghanaians can speak or comprehend some French [23]. French is a global language with an estimated 300 million speakers and "official language of 32 states and governments" [24]. With 59% of global French speakers in Africa and as a major global language for trading [25], the demand to learn French among Ghanaians is increasing rapidly. The demand even led the Ghanaian parliament to approve French as a second official language [26]. A machine translation system between Ghana's most widely spoken language and French is certainly a valuable resource for bridging the gap.

However, most of the Twi natural language processing resources that can be used to build machine translation systems are classified as *noisy* and *religiously biased* [21]. There are already machine translation systems that allow translations from local Ghanaian languages to other languages. For example, Khaya, a neural machine translation system by the GhanaNLP group and Algorine, allows machine translation from Ghanaian languages, such as Twi, Ewe, and Ga, to English [27]. A recent addition of Twi to Google's Neural Machine Translation System enables translations from Twi to over 100 languages [28]. Furthermore, pre-trained Twi machine translation models are provided by the Language Technology Research Group at the University of Helsinki, including a model for Twi and French machine translation [29], which, however, require parallel corpora to be fine tuned.

Consequently, since there are still hardly any open-source machine translation systems or parallel corpora between Twi and French that cover a modern and versatile vocabulary, our goal was to extend the modern English–Akuapem Twi corpus of [21] and develop machine translation systems between Twi and French. A Twi–French parallel corpus and the corresponding machine translation applications will offer various advantages, including stimulating trade and job creation. Moreover, it will support the global Ghanaian diaspora in French-speaking nations, as they will be able to acquire Twi and Ghanian culture [21]. The system will also help French-speaking tourists and immigrants seeking medical care in Ghana [30]. Additionally, the French–Twi parallel corpus can be used for numerous downstream natural language processing tasks, including named entity recognition and part-of-speech tagging with the appropriate annotations.

Our contributions are as follows:

- We are the first to introduce non-commercial machine translation systems for Twi–French and French–Twi.
- We created a parallel Twi–French corpus by extending an existing Twi–English corpus.
- For our language pairs, we investigated direct machine translation and cascading systems that use English as a pivot language.
- We compared our systems with the commercial system of Google Translate and managed to slightly outperform Google Translate with our best French–Twi system.
- To contribute to the improvement of low-resource languages, we share our code and our corpus with the research community (https://github.com/gyasifred/TW-FR-MT).

In the following section, we will give a brief insight into the linguistic categorization and the peculiarities of Akuapem Twi. In Section 2, we will describe related work regarding existing parallel text corpora and machine translation systems for Twi. Section 4 will present our parallel Twi–French–English corpus. Our Twi–French and French–Twi machine translation experiments will be described in Section 5. In Section 6, we will briefly discuss the performance of our machine translation output. We will conclude our work in Section 7 and suggest further steps.

2. The Language Twi

As visualized in Figure 1, Twi is a collection of dialects which belongs to Akan. Akan is the language of the Akan ethnic group in Ghana [31] and the principle language of Ghana [32]. The Akan dialects include Agona, Akuapem, Akwamu, Asante, Akyem, Assin, Bono, Fante, Kwahu, Wassa, Sefwi, Anyi, and Guan [33]. These dialects are divided into two categories: Fante and Twi (which includes all non-Fante dialects) [33]. Since not all dialects understand each other, Akuapem Twi serves as a pivot language and is used for education purposes in schools. The Akan Orthography Committee (AOC) developed a unified Akan orthography in 1978, based mainly on Akuapem Twi [34]. Consequently, those Twi dialects that do not have their own orthography use Akuapem Twi (marked in red) [35]. Only Asante and Bono have separate writing systems (marked in orange). Fante does not belong to the Twi dialects and has its own orthography (marked in orange).

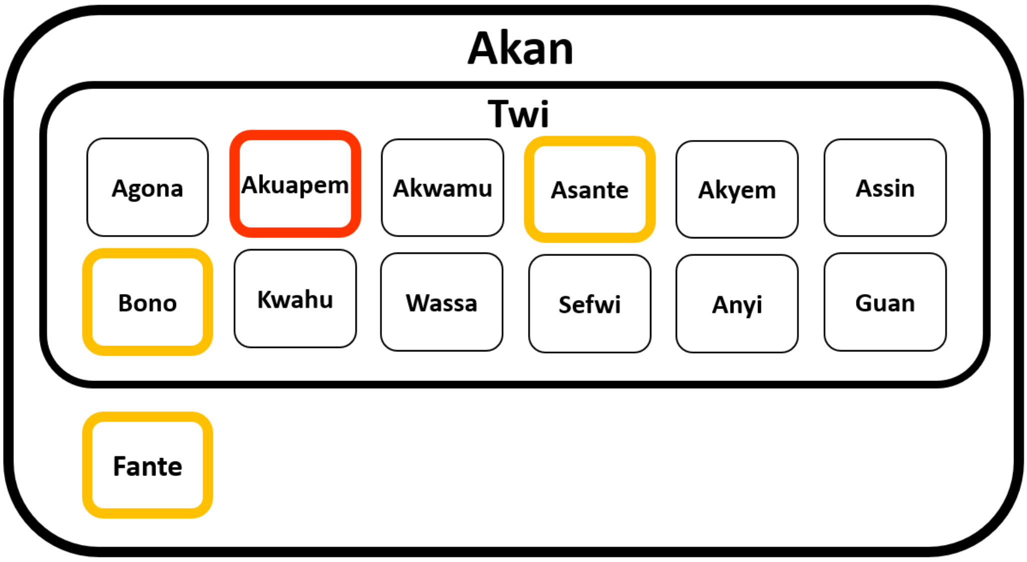

Figure 1. Akan dialects.

To allow all Twi-speaking people to have access to French and since a unified orthography exists, we decided to collect a corpus and build machine translation systems in Akuapem Twi. For simplicity and since all Twi-speaking people would be able to operate with Akuapem Twi machine translation, we refer to Akuapem Twi as *Twi* in this paper.

Twi is a tonal language with distinct semantic connotations for its high, mid, and low tones [21]. The Twi alphabet consists of 22 letters made up of 15 consonants and 7 vowels [36]. Additionally, though mostly in loanwords, the letters C, J, V, and Z are used. Twi has ten diphthongs. Many Twi words have multiple meanings and can be used interchangeably in the same context [37]. For example, the word sequence “me papa” means “good mood” or “my dad”, depending on the context. In contrast to other languages, removing stop words may affect the entire meaning of a Twi sentence. For instance, the Twi word “na” could represent the word “and”, the word “then”, the phrase “and then”, or the word “mother”. Consequently, we do not remove Twi stop words in the pre-processing pipeline of our machine translation systems.

3. Related Work

In this section, we will look at existing Twi text corpora and machine translation systems.

3.1. Parallel Corpora for Twi

The authors of [37] analyzed the use of the Twi Bible dataset, Jehovah's Witness data, Wikipedia, and the JW300 Twi corpus [38]. However, they classified Jehovah's Witness data, Wikipedia and the JW300 Twi corpus as *noisy*, i.e., not optimal for the use in machine translation tasks due to spellings, Twi sentences formulated in a non-natural way, mixtures of dialects, etc. Furthermore, [21] report that the JW300 Twi corpus and the Bible are religiously biased datasets.

Despite the "noise" and the religious bias, JW300 is a huge corpus "of over 300 languages with around 100 thousand parallel sentences per language pair on average" covering a wide range of topics [38]. In addition to other language pairs, such as English–French and French–English, parallel JW300 corpora for English–Twi, French–Twi, Twi–French, Finnish–Twi, Twi–Finnish, Swedish–Twi, Twi–Swedish, Spanish–Twi, and Twi–Spanish were provided in the OPUS repository (https://opus.nlpl.eu/Opus-MT (accessed on 16 April 2023)) [29]. Due to copyright issues, the corpus is not accessible at the moment. However, the Language Technology Research Group at the University of Helsinki (https://huggingface.co/Helsinki-NLP (accessed on 16 April 2023)) provides large machine translation Twi models called *OPUS-MT models*, which were trained on the JW300 data when they were still available in the OPUS repository plus other text data.

The authors of [39] present the Twieng corpus, a small English–Twi parallel corpus with 5,419 sentences. The corpus is based on online news portals, Twi literature, the Ghanaian Parliamentary Hansard, the Twi–English Bible, and Social Media crowdsourcing and has a focus on socio-cultural, educational and legal issues.

The TypeCraft Akan corpus provides 80,000 parallel translations in Twi and English [40,41]. However, only the release 1.0 of the TypeCraft Akan Corpus with 669 monolingual sentences is publicly available (https://typecraft.org/w/index.php?title=The_TypeCraft_Akan_Corpus (accessed on 16 April 2023)).

The LORELEI (Low Resource Languages for Emergent Incidents) Akan Representative Language Pack (https://catalog.ldc.upenn.edu/LDC2021T02 (accessed on 16 April 2023)) contains almost 3.3 million Akan words of monolingual text translated into English, as well as 115,000 Akan words translated from English data [42–44]. However, access to this corpus is not freely available, not even for researchers.

The authors of [21] offer a corpus of 25,421 English–Twi sentence pairs whose English sentences were curated from tatoeba.org. For our work, we selected this dataset since it is large enough for machine translation, and the language is more modern compared to the other corpora.

3.2. Twi Machine Translation Systems

The authors of [45] used an English–Twi parallel Bible corpus to train transformer-based neural machine translation systems [46], statistical machine translation systems developed with the Moses toolkit [47] and sequence-to-sequence recurrent neural network-based machine translation systems. The transformer-based neural machine translation systems outperformed the statistical machine translation system and the recurrent neural network-based system in both directions.

Furthermore, GhanaNLP, an open-source initiative and the company Algorine, introduced the Khaya Android app which executes neural machine translation in Ghanaian languages, such as Twi, Ewe, and Ga to English [27]. There are also other Twi machine translation systems available online, but their technology is not clearly explained.

As mentioned in Section 3.1, the Language Technology Research Group at the University of Helsinki provides the *OPUS-MT models* (https://opus.nlpl.eu/Opus-MT (accessed on 16 April 2023)) for English–Twi, French–Twi, Twi–French, Finnish–Twi, Twi–Finnish, Swedish–Twi, Twi–Swedish, Spanish–Twi, and Twi–Spanish in addition to other language pairs at their GitHub repository [29]. English–French and French–English models are also available, which can be used in pivot machine translation systems. The *OPUS-MT models* are transformer models which were pre-trained on datasets at the OPUS repository [29] with the help of the MarianMT toolkit [48]. The company John Snow Labs uses these pre-trained models and provides fine-tuned machine translation models with the Spark NLP package [49]. Additionally, [21] fine-tuned the English–Twi *OPUS-MT model* for their English–Twi machine translation experiments.

Since they have proven to be successful as base models, we also used the Twi–French, French–Twi, English–Twi, English–French, and French–English *OPUS-MT models* provided by [29] as base models in our machine translation systems. We fine-tuned these models with the training and validation sets of our corpus of 10,708 parallel sentences, which we introduce in the next section. Since Google's Neural Machine Translation System recently included Twi [28], we also compare our results to the Google Translate output.

4. Our Parallel Twi–French–English Corpus

Our parallel Twi–French–English corpus is based on a subset of the open-source English–Akuapem Twi corpus created by [21]. The benefit of this corpus is that it is large enough for machine translation, the language is more modern compared to the other corpora, and it was used to build a machine translation system by [21].

We randomly extracted 10,708 sentence pairs from this corpus and let professional translators create the French translation of these sentences. Table 1 summarizes the number of sentences, number of running words (*word tokens*) and number of unique words (*unique words*) in the resulting corpus. To contribute to the improvement of low-resource languages, we share the corpus with the research community on GitHub.

Table 1. Statistics of our parallel Twi–French–English corpus.

	#Sentences	#Word Tokens	#Unique Words
Twi	10,708	70,400	7,966
French	10,708	70,257	10,160
English	10,708	67,677	8,239

To train, tune and evaluate our machine translation systems, we split the corpus into a training set (80%), validation set (10%) and test set (10%) as shown in Table 2. Whereas the validation set was used to find the optimal model parameters in each training epoch with high BLEU scores, we used the test set to evaluate the final system.

Table 2. Distribution of training set, validation set and test set.

	#Sentences	%
Train	8,566	80
Validation	1,071	10
Test	1,071	10

5. Twi Machine Translation

In this section, we will first present our three evaluation metrics. Then, we will describe the setup and the training procedure of our systems. Finally, we will report the systems' performances.

5.1. Evaluation Metrics

As stated in [50], "a(A)lthough people refer to "the" BLEU score, BLEU is in fact a parameterized metric whose values can vary wildly with changes to these parameters". Since in the related literature different implementations of BLEU are used, we report our results with three metrics. In all metrics, there are one or more *references*, i.e., human-translated versions of a sentence, as well as a *hypothesis*, i.e., a translation generated by the machine translation system. The *hypothesis* is compared to the *reference*. In all cases, higher scores reflect better translations.

5.1.1. BLEU

BLEU (Bilingual Evaluation Understudy) [51] is a precision metric which usually compares word-level n-grams [50] of the *hypothesis* sentence and one or more *reference* sentences. To obtain sentences for a useful and comparable BLEU evaluation, often text processing (e.g., normalization, tokenization, compound splitting, and the removal of case) is required for both the *reference* and *hypothesis*. BLEU scores are dependent on the translated language pair and on the settings of the parameters used to compute the BLEU score. Since no results are yet reported for Twi–French and French–Twi machine translation, we expect scores in a similar range to English–Twi BLEU scores, e.g., in [21].

Compared to the accuracy—which is used for evaluating many other machine learning tasks—BLEU is designed to be a more nuanced and sophisticated measure of translation quality. By using n-grams, it takes into account not only word-by-word matches but also higher-level aspects of the translation and sentence structure. Thus, BLEU can better capture the overall quality of a machine translation system, whereas accuracy provides only a binary assessment of correctness.

We benefited from the *corpus-bleu* function (https://www.nltk.org/_modules/nltk/translate/bleu_score.html (accessed on 16 April 2023)) in the Natural Language Toolkit (NLTK) [52] to compute the BLEU scores. To be comparable to related work, we used a regular 4-gram BLEU implementation with the parameters *smoothing_function = 7*, *auto_reweigh = False*, and *weights = (0.25, 0.25, 0.25, 0.25)*.

5.1.2. AzunreBLEU

Since our corpus is based on [21]'s English–Akuapem Twi corpus, our goal was to investigate if our Twi–French and French–Twi machine translation systems are in the same performance range as their English–Twi machine translation system.

Consequently, we adapted the *corpus-bleu* function in the NLTK [52] with the parameter values reported by [21]. Indicating that the focus is on "adequacy" instead on "fluency" in the translations, they use the following parameters: *smoothing_function = 7*, *auto_reweigh = True*, and *weights = (0.58, 0.0, 0.0, 0.0)*. We refer to this BLEU score variant setting as AzunreBLEU.

5.1.3. SacreBLEU

To have a score that is comparable to related work despite different tokenization and normalization schemes, [50] proposed SacreBLEU. Compared to BLEU, in SacreBLEU, the *hypothesis* is compared to one or more *references* as well. However, SacreBLEU expects detokenized hypotheses, applying its own metric-internal pre-processing [50].

To compute the SacreBLEU scores, we benefited from the *sacrebleu* library (https://github.com/mjpost/sacrebleu (accessed on 16 April 2023)) provided by [50]. We used the regular 4-gram implementation with the following parameters: *smooth-method = add-k*, *force = False*, *lowercase = True*, *tokenize = intl*, and *bleu_effective_order = True*.

5.2. Systems' Setup

Some researchers propose cascading cross-lingual natural language processing approaches to solve the problems of low-resource languages by benefiting from models of rich-resource languages, such as English [53–57]. Consequently, as visualized in Figure 2, we investigate direct machine translation and cascading systems that use English (*EN*) as a pivot language for the Twi–French (*TW–FR*) and French–Twi (*FR–TW*) translations:

- In the direct machine translation systems, the texts in the source language is directly translated to the target language using a <source language>–<target language> model.
- In the cascading systems, the source language text is first translated to *EN* and then translated into the target language using a <source language>–*EN* model and an *EN*–<target language> model.

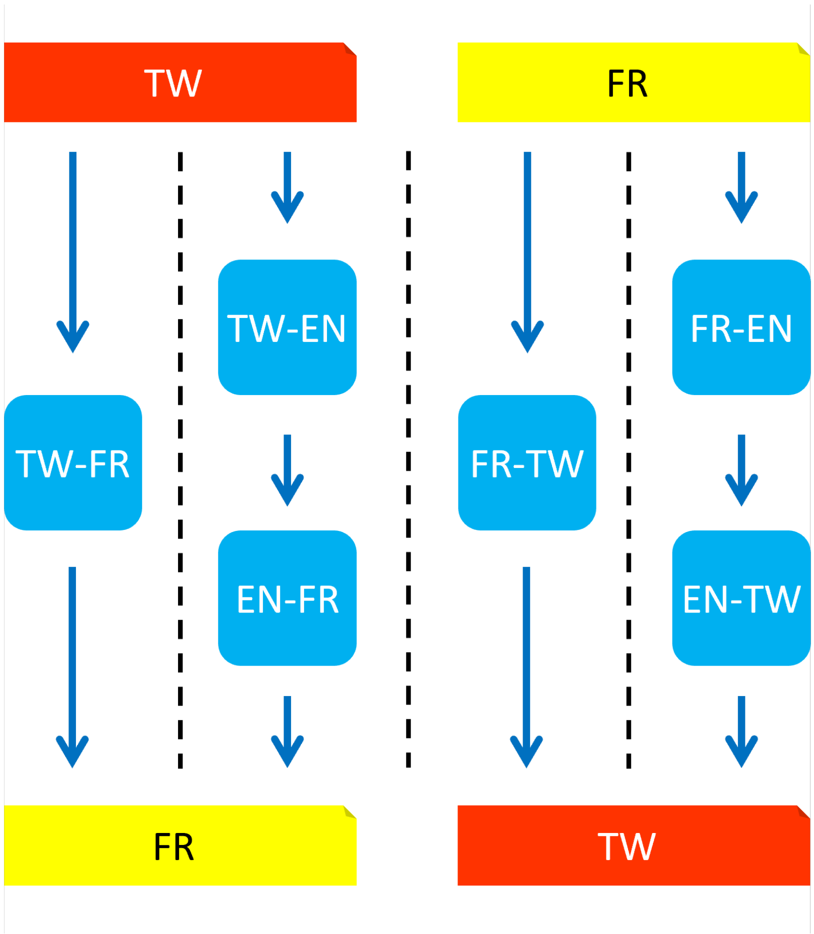

Figure 2. Systems' overview.

As mentioned in Section 3.2, the pre-trained transformer-based *OPUS-MT models* provided by [29] proved to be successful. Therefore, we used the Twi–French, French–Twi, English–Twi, English–French, and French–English models (1) directly for the downstream task. (2) We fine-tuned these models using our 8566 parallel training sentences and 1071 parallel sentences from the validation set. Since a Twi–English *OPUS-MT model* is not available in the OPUS repository, we used the weights of the Twi–French *OPUS-MT model* as the initial weights.

To process our texts with the *OPUS-MT models*, we applied the tokenization provided with the *OPUS-MT* pre- and post-processing scripts. This tokenization is based on *SentencePiece* [29,58], a language-independent sub-word tokenization algorithm developed by Google. This tokenization is typically employed in neural network-based text generation systems, where the size of the vocabulary is predetermined prior to the neural model training.

All six models (*OPUS_tw-fr*, *OPUS_tw-en*, *OPUS_en-fr*, *OPUS_fr-tw*, *OPUS_fr-en*, *OPUS_en-tw*) were fine-tuned with a batch size of 8 using the Adam optimizer [59] with a learning rate of 2×10^{-5}. For each machine translation system, we determined the number of epochs which gave the best *SacreBLEU* scores (as defined in [50]) on the validation set. While for *OPUS_tw-fr*, *OPUS_en-fr*, *OPUS_fr-tw*, *OPUS_en-tw* 16 epochs were sufficient, 32 epochs were required for *OPUS_tw-en* and 24 epochs for *OPUS_fr-en*.

Since Google's Neural Machine Translation System recently included Twi [28], we additionally compare our results to the Google Translate output. For all implementations, we used Google Colab (https://colab.research.google.com (accessed on 16 April 2023)).

5.3. Results

Table 3 summarizes the systems' BLEU, AzunreBLEU, and SacreBLEU scores. We see that systems which are fine-tuned with the particular language pair in our collected corpus—indicated with the suffix "*_tuned*" and the entry "*our corpus*"—significantly outperform the *OPUS-MT models* which are not fine-tuned.

Table 3. Machine translation systems' performance results: Twi–French and French–Twi.

Model	Fine-Tuning	BLEU	AzunreBLEU	SacreBLEU
Google_tw-fr	—	**0.41**	**0.79**	**0.44**
Google_tw-en → *Google_en-fr*	—	**0.41**	**0.79**	**0.44**
OPUS_tw-fr	—	0.22	0.60	0.21
OPUS_tw-fr_tuned	our corpus	**0.36**	**0.76**	**0.37**
OPUS_tw-en_tuned → *OPUS_en-fr_tuned*	our corpus	0.30	0.67	0.31
Google_fr-tw	—	0.37	0.76	0.34
Google_fr-en → *Google_en-tw*	—	0.37	0.76	0.34
OPUS_fr-tw	—	0.21	0.62	0.14
OPUS_fr-tw_tuned	our corpus	0.40	0.77	0.39
OPUS_fr-en_tuned → ***OPUS_en-tw_tuned***	our corpus	**0.44**	**0.81**	**0.42**

As AzunreBLEU was used as an evaluation metric in related work [21], Figure 3 visualizes a direct comparison between the AzunreBLEU scores of the Google Translate systems and our systems. Our best Twi–French system is *OPUS_tw-fr_tuned*, the direct state-of-the-art transformer-based machine translation system, which achieves an AzunreBLEU score of 0.76. Our best French–Twi system *OPUS_fr-en_tuned* → *OPUS_en-tw_tuned*, which is a cascading system that uses English as a pivot language, results in an AzunreBLEU score of 0.81. This demonstrates that we have results comparable to [21], who reported an AzunreBLEU of 0.72 in their English–Twi machine translation system.

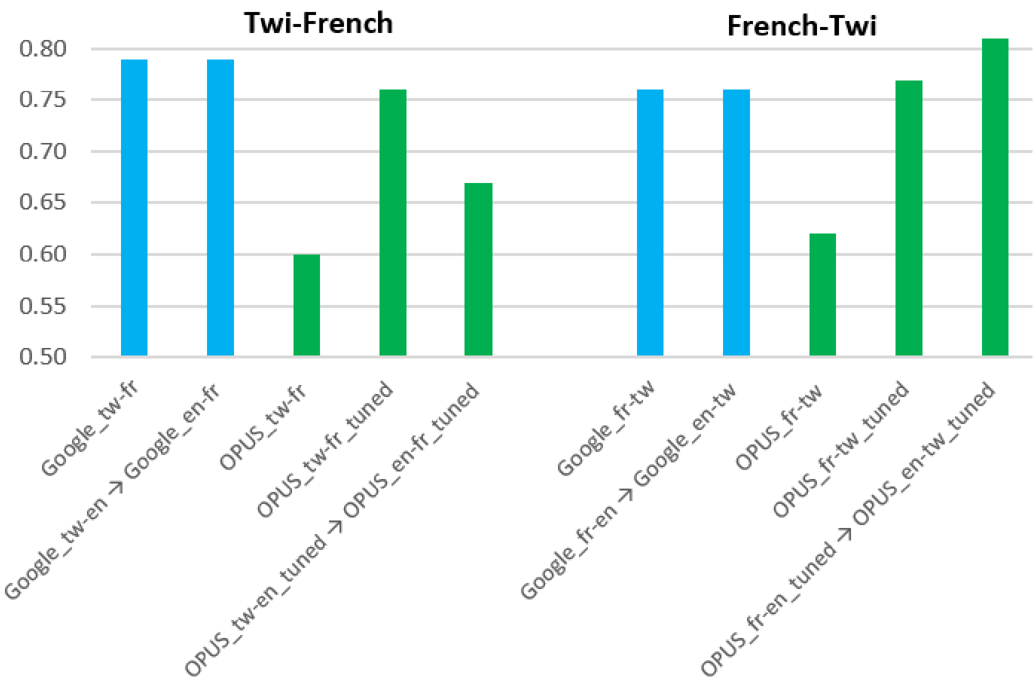

Figure 3. AzunreBLEU scores of Google Translate and our machine translation systems.

Table 4 lists the relative improvement of our best Twi–French and French–Twi systems that are fine-tuned with our corpus compared to the pre-trained *OPUS-MT models* and Google Translate on our test set: We see that leveraging our corpus to fine-tune the models has a huge impact on the systems' performance in all three evaluation metrics.

Table 4. Difference of our best systems compared to pre-trained OPUS models and Google Translate.

	BLEU	AzunreBLEU	SacreBLEU
Δ to *Google_tw_fr*	−12.2%	−3.8%	−15.9%
Δ to *OPUS_tw_fr*	63.6%	26.7%	76.2%
Δ to *Google_fr_tw*	18.9%	6.6%	23.5%
Δ to *OPUS_tw_fr*	125.0%	30.6%	200.0%

While we are not able to obtain better results than Google Translate in our Twi–French machine translation task—we suspect that the Google models were trained with significantly more data—we are able to outperform Google Translate slightly in the other direction with 6.6% relatively higher AzunreBLEU scores, 23.5% relatively higher SacreBLEU scores, and 18.9% relatively higher BLEU scores.

6. Discussion

As explained in Section 4, our parallel Twi–French–English corpus is based on a subset of the open-source English–Akuapem Twi corpus created by [21]. The benefit of this corpus is that it is large enough for machine translation, and the language is more modern compared to the other corpora. It contains vocabulary that is useful in everyday conversations. Therefore, our systems perform well in such general conversations, which was our goal in this work. However, since the systems were only fine-tuned with this corpus, they have difficulty with translating technical sentences perfectly.

Figure 4 shows an example where our best French–Twi system *OPUS_fr-en_tuned* → *OPUS_en-tw_tuned* is not able to translate the English words for "artificial intelligence" and "machine translation" into Twi. As we can see, Google Translate has no problems with these technical terms, even though in this example the Google Translate translation is not perfect either. To cover technical and other domains better, our system would have to be fine-tuned with text data of the corresponding domains.

French:

Grâce à l'intelligence artificielle, il est désormais possible de surpasser de nombreux systèmes de traduction automatique traditionnels.

(English: Thanks to artificial intelligence, it is now possible to outperform many traditional machine translation systems.)

Correct translation:

Aseda nkɔ ma nyansa a wɔde ayɛ nneɛma, afei de, wobetumi ayɛ adwuma asen mfiri nkyerɛase nhyehyɛe pii a wɔde di dwuma wɔ amanne kwan so no.

Our best system:

Esiane sɛ nnipa nyansa da adi nti, seesei wobetumi ayɛ afiri a wɔde kyerɛ nhoma ase pii wɔ ɔkwan a etu mpɔn so.

(English: Thanks to the emergence of human intelligence, many translation machines can now be developed efficiently.)

Google Translate:

Esiane nyansa a wɔde ayɛ nneɛma nti, seesei wobetumi ayɛ adwuma asen mfiri nkyerɛase nhyehyɛe pii a wɔde di dwuma wɔ amanne kwan so no.

(English: Thanks to artificial intelligence, they can now outperform many traditional machine translation systems.)

Figure 4. Example of our systems' shortcomings in technical domains.

7. Conclusions and Future Work

French is a strategically and economically important language in the regions where the African language Twi is spoken. However, only a very small proportion of Twi speakers in Ghana speak French. Since there are hardly any machine translation systems or parallel corpora between Twi and French that cover a modern and versatile vocabulary, our goal was to extend the modern Twi–English corpus of [21]. We randomly extracted 10,708 sentence pairs from this corpus and let professional translators create the French translation of these sentences.

Furthermore, we developed machine translation systems between Twi and French. Since in the related literature, different implementations of machine translation evaluation metrics are used, we reported our results with the three metrics BLEU, AzunreBLEU and SacreBLEU. We investigated direct machine translation and cascading systems that use English as a pivot language for Twi–French and French–Twi translations. Our best Twi–French system is *OPUS_tw-fr_tuned*, the direct state-of-the-art transformer-based machine translation system which achieves an AzunreBLEU score of 0.76. Our best French–Twi system *OPUS_fr-en_tuned* $\rightarrow$ *OPUS_en-tw_tuned*, which is a cascading system that uses English as a pivot language, results in an AzunreBLEU score of 0.81.

After we collect a corpus with rather simple and short sentences, we plan to extend our corpus with more complex sentences of further domains. Since, to date, no language-specific algorithms for pre-processing have been investigated, future work may deal with the peculiarities of the Twi language. Additionally, to allow everyone to access our machine translation systems, our goal is to provide a web interface. Furthermore, we plan to extend the machine translation corpus with more African low-resourced languages.

Author Contributions: Conceptualization, methodology, software, validation, resources, writing, visualization: F.G. and T.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: To contribute to the improvement of low-resource languages, we share our code and our corpus with the research community: https://github.com/gyasifred/TW-FR-MT

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Steigerwald, E.; Ramírez-Castañeda, V.; Brandt, D.Y.C.; Báldi, A.; Shapiro, J.T.; Bowker, L.; Tarvin, R.D. Overcoming Language Barriers in Academia: Machine Translation Tools and a Vision for a Multilingual Future. *BioScience* **2022**, *72*, 988–998. [CrossRef] [PubMed]
2. Wu, Y.; Schuster, M.; Chen, Z.; Le, Q.V.; Norouzi, M.; Macherey, W.; Krikun, M.; Cao, Y.; Gao, Q.; Macherey, K.; et al. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. *arXiv* **2016**, arXiv:1609.08144.
3. Garg, A.; Agarwal, M. Machine Translation: A Literature Review. *arXiv* **2019**, arXiv:1901.01122.
4. What Is Machine Translation? Available online: https://aws.amazon.com/what-is/machine-translation (accessed on 16 April 2023).
5. Sabtan, Y.; Hussein, M.; Ethelb, H.; Omar, A. An Evaluation of the Accuracy of the Machine Translation Systems of Social Media Language. *Int. J. Adv. Comput. Sci. Appl.* **2021**, *12*, 406–415. [CrossRef]
6. Urlaub, P.; Dessein, E. Machine Translation and Foreign Language Education. *Front. Artif. Intell.* **2022**, *5*, 936111. [CrossRef] [PubMed]
7. Schlippe, T.; Sawatzki, J. Cross-Lingual Automatic Short Answer Grading. In Proceedings of the 2nd International Conference on Artificial Intelligence in Education Technology (AIET), Wuhan, China, 18–20 June 2021.
8. Schlippe, T.; Sawatzki, J. AI-Based Multilingual Interactive Exam Preparation. In Proceedings of the Learning Ideas Conference 2021 (14th annual conference), ALICE - Special Conference Track on Adaptive Learning via Interactive, Collaborative and Emotional Approaches, New York, NY, USA, 14–18 June 2021.
9. Schlippe, T.; Eichinger, K. Multilingual Text Simplification and its Performance on Social Sciences Coursebooks. In Proceedings of the 4th International Conference on Artificial Intelligence in Education Technology (AIET), Berlin, Germany, 30 June–2 July 2023.
10. Brynjolfsson, E.; Hui, X.; Liu, M. Does Machine Translation Affect International Trade? Evidence from a Large Digital Platform. *Manag. Sci.* **2019**, *65*, 5449–5460. [CrossRef]
11. Emezue, C.C.; Dossou, B.F.P. MMTAfrica: Multilingual Machine Translation for African Languages. In Proceedings of the Sixth Conference on Machine Translation, Online, 10–11 November 2021; pp. 398–411.
12. Cho, K.; van Merrienboer, B.; Bahdanau, D.; Bengio, Y. On the Properties of Neural Machine Translation: Encoder-Decoder Approaches. In Proceedings of the SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation, Doha, Qatar, 25 October 2014.
13. Sutskever, I.; Vinyals, O.; Le, Q.V. Sequence to Sequence Learning with Neural Networks. In Proceedings of the NIPS'14 27th International Conference on Neural Information Processing Systems, Montréal, QC, Canada, 8–13 December 2014; Volume 2, pp. 3104–3112.
14. Bahdanau, D.; Cho, K.; Bengio, Y. Neural Machine Translation by Jointly Learning to Align and Translate. *arXiv* **2014**, arXiv:1409.0473.
15. Kalchbrenner, N.; Espeholt, L.; Simonyan, K.; van den Oord, A.; Graves, A.; Kavukcuoglu, K. Neural Machine Translation in Linear Time. *arXiv* **2016**, arXiv:1610.10099.

16. Yang, J.; Yin, Y.; Ma, S.; Zhang, D.; Li, Z.; Wei, F. High-resource Language-specific Training for Multilingual Neural Machine Translation. In Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, (IJCAI-22), Vienna, Austria, 23–29 July 2022; pp. 4461–4467. [CrossRef]
17. Haddow, B.; Bawden, R.; Miceli Barone, A.V.; Helcl, J.; Birch, A. Survey of Low-Resource Machine Translation. *Comput. Linguist.* **2022**, *48*, 673–732. [CrossRef]
18. Ranathunga, S.; Lee, E.S.A.; Skenduli, M.P.; Shekhar, R.; Alam, M.; Kaur, R. Neural Machine Translation for Low-Resource Languages: A Survey. *ACM Comput. Surv.* **2021**, *55*, 1–37. [CrossRef]
19. Meuret, D.; Duru-Bellat, M. English and French Modes of Regulation of the Education System: A Comparison. *Comp. Educ.* **2003**, *39*, 463–477. [CrossRef]
20. Cashman, K. Masakhane: Using AI to Bring African Languages Into the Global Conversation. 2023. Available online: https://en.reset.org/masakhane-using-ai-bring-african-languages-global-conversation-02042020 (accessed on 16 April 2023).
21. Azunre, P.; Osei, S.; Salomey, A.A.; Adu-Gyamfi, L.A.; Moore, S.; Adabankah, B.; Opoku, B.; Asare-Nyarko, C.; Nyarko, S.; Amoaba, C.; et al. English-Twi Parallel Corpus for Machine Translation. *arXiv* **2021**, arXiv:2103.15625.
22. Ghana. 2023. Available online: https://thecommonwealth.org/our-member-countries/ghana (accessed on 16 April 2023).
23. Fournier-Passard, Y. Ghana Relies on French Language to Influence West Africa. 2019. Available online: http://www.echosdughana.com/2019/07/08/ghana-relies-on-french-language-to-influence-west-africa (accessed on 16 April 2023).
24. The French Language in Figures. 2022. Available online: https://www.diplomatie.gouv.fr/en/french-foreign-policy/francophony-and-the-french-language/the-french-language-in-figures (accessed on 16 April 2023).
25. Qui Parle Français Dans Le Monde. 2023. Available online: http://observatoire.francophonie.org/qui-parle-francais-dans-le-monde (accessed on 16 April 2023).
26. Parliament Support Choice of French as 2nd Language. 2019. Available online: https://www.parliament.gh/news?CO=40 (accessed on 16 April 2023).
27. Khaya Translator App—Android. 2021. Available online: https://ghananlp.org/project/khaya-android (accessed on 16 April 2023).
28. Caswell, I. Google Translate Learns 24 New Languages. 2022. Available online: https://blog.google/products/translate/24-new-languages (accessed on 16 April 2023).
29. Tiedemann, J.; Thottingal, S. OPUS-MT—Building Open Translation Services for the World. In Proceedings of the 22nd Annual Conference of the European Association for Machine Translation, Lisboa, Portugal, 3–5 November 2020; pp. 479–480.
30. Chachu, S. Implications of Language Barriers for Access to Healthcare: The Case of Francophone Migrants in Ghana. *Legon J. Humanit.* **2022**, *32*, e122. [CrossRef]
31. Akan (Twi) at Rutgers. 2022. Available online: https://www.amesall.rutgers.edu/languages/128-akan-twi (accessed on 16 April 2023).
32. Akan Twi. 2023. Available online: https://celt.indiana.edu/portal/Akan%20Twi/index.html (accessed on 16 April 2023).
33. Osam, E.K. An Introduction to the Verbal and Multi-Verbal System of Akan. In Proceedings of the Workshop on Multi-verb Constructions, Trondheim, Norway, 2003.
34. Kouadio, N.J. *A Unified Orthography for the Akan Languages of Ghana and Ivory Coast: General Unified Spelling Rules*; Monograph Series; Centre for Advanced Studies of African Society, CASAS: Cape Town, South Africa, 2003; Volume 20.
35. Schachter, P.; Fromkin, V. *A Phonology of Akan: Akuapem, Asante, Fante*; Working Papers in Phonetics; University of California: Los Angeles, CA, USA, 1979.
36. The African Linguists Network Blog. Language Guide. Available online: https://alnresources.wordpress.com/african-culture-and-language (accessed on 16 April 2023).
37. Alabi, J.O.; Amponsah-Kaakyire, K.; Adelani, D.I.; España-Bonet, C. Massive vs. Curated Embeddings for Low-Resourced Languages: The Case of Yorùbá and Twi. In Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020), Marseilles, France, 13–15 May 2020.
38. Agić, Ž.; Vulić, I. JW300: A Wide-Coverage Parallel Corpus for Low-Resource Languages. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, 28 July–2 August 2019; pp. 3204–3210. [CrossRef]
39. Afram, G.K.; Weyori, B.A.; Adekoya, F.A. TWIENG: A Multi-Domain Twi-English Parallel Corpus for Machine Translation of Twi, a Low-Resource African Language. *Preprints* **2022**. [CrossRef]
40. Beermann, D.; Hellan, L. Enhancing Grammar and Valence Resources for Akan and Ga. In *West African Languages. Linguistic Theory and Communication*; WUW: Warszawa, Poland, 2020; pp. 166–188. [CrossRef]
41. Beermann, D.; Hellan, L.; Mihaylov, P.; Struck, A. Developing a Twi (Asante) Dictionary from Akan Interlinear Glossed Texts. In Proceedings of the 1st Joint Workshop on Spoken Language Technologies for Under-resourced languages (SLTU) and Collaboration and Computing for Under-Resourced Languages (CCURL), Marseille, France, 13–15 May 2020; pp. 294–297.
42. Strassel, S.; Tracey, J. LORELEI Language Packs: Data, Tools, and Resources for Technology Development in Low Resource Languages. In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), Portorož, Slovenia, 23–28 May 2016; pp. 3273–3280.
43. Christianson, C.; Duncan, J.; Onyshkevych, B. Overview of the DARPA LORELEI Program. *Mach. Transl.* **2018**, *32*, 3–9. [CrossRef]
44. Tracey, J.; Strassel, S.; Graff, D.; Wright, J.; Chen, S.; Ryant, N.; Kulick, S.; Griffitt, K.; Delgado, D.; Arrigo, M. *LORELEI Akan Representative Language Pack*; Web Download; Linguistic Data Consortium: Philadelphia, PA, USA, 15 January 2021. [CrossRef]
45. Adjeisah, M.; Liua, G.; Nortey, R.N.; Song, J. English↔Twi Parallel-Aligned Bible corpus for Encoder-Decoder based machine translation. *Acad. J. Sci. Res.* **2020**, *8*, 371–382.

46. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention is All you Need. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; Volume 30.
47. Hoang, H.; Koehn, P. Design of the Moses Decoder for Statistical Machine Translation. In Proceedings of the Software Engineering, Testing, and Quality Assurance for Natural Language Processing, Columbus, OH, USA, 20 June 2008; pp. 58–65.
48. Junczys-Dowmunt, M.; Grundkiewicz, R.; Dwojak, T.; Hoang, H.; Heafield, K.; Neckermann, T.; Seide, F.; Germann, U.; Aji, A.F.; Bogoychev, N.; et al. Marian: Fast Neural Machine Translation in C++. In Proceedings of the ACL 2018, System Demonstrations, Melbourne, Australia, 15–20 July 2018; pp. 116–121. [CrossRef]
49. Spark NLP. 2023. Available online: https://nlp.johnsnowlabs.com (accessed on 16 April 2023).
50. Post, M. A Call for Clarity in Reporting BLEU Scores. *arXiv* **2018**, arXiv:1804.08771.
51. Papineni, K.; Roukos, S.; Ward, T.; Zhu, W.J. Bleu: A Method for Automatic Evaluation of Machine Translation. In Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL 2002), Philadelphia, PA, USA, 6–12 July 2002.
52. Loper, E.; Bird, S. NLTK: The Natural Language Toolkit. In Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL 2002), Philadelphia, PA, USA, 6–12 July 2002; pp. 63–70. [CrossRef]
53. Lin, Z.; Jin, X.; Xu, X.; Wang, Y.; Tan, S.; Cheng, X. Make It Possible: Multilingual Sentiment Analysis Without Much Prior Knowledge. In Proceedings of the IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), Warsaw, Poland, 11–14 August 2014; Volume 2, pp. 79–86. [CrossRef]
54. Balahur, A.; Turchi, M. Comparative Experiments using Supervised Learning and Machine Translation for Multilingual Sentiment Analysis. *Comput. Speech Lang.* **2014**, *28*, 56–75. [CrossRef]
55. Vilares, D.; Alonso Pardo, M.; Gómez-Rodríguez, C. Supervised Sentiment Analysis in Multilingual Environments. *Inf. Process. Manag.* **2017**, *53*, 595–607. [CrossRef]
56. Can, E.F.; Ezen-Can, A.; Can, F. Multilingual Sentiment Analysis: An RNN-Based Framework for Limited Data. In Proceedings of the ACM SIGIR 2018 Workshop on Learning from Limited or Noisy Data, Ann Arbor, MI, USA, 8–12 July 2018.
57. Rakhmanov, O.; Schlippe, T. Sentiment Analysis for Hausa: Classifying Students' Comments. In Proceedings of the 1st Annual Meeting of the ELRA/ISCA Special Interest Group on Under-Resourced Languages (SIGUL 2022), Marseille, France, 24–25 June 2022.
58. Kudo, T.; Richardson, J. SentencePiece: A Simple and Language Independent Subword Tokenizer and Detokenizer for Neural Text Processing. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP): System Demonstrations, Brussels, Belgium, 7–11 December 2018; pp. 66–71. [CrossRef]
59. Kingma, D.P.; Ba, J. *Adam: A Method for Stochastic Optimization*; In Proceedings of the 3rd International Conference for Learning Representations, San Diego, CA, USA, 7–9 May 2015. [CrossRef]