Indonesian Journal of Computer Science
ISSN 2549-7286 (*online*)
Jln. Khatib Sulaiman Dalam No. 1, Padang, Indonesia
Website: ijcs.stmikindonesia.ac.id | E-mail: ijcs@stmikindonesia.ac.id

Crosslingual Transfer Learning for Arabic Story Ending Generation

Arwa Alhussain, Aqil M. Azmi
ahussain@ksu.edu.sa, aqil@ksu.edu.sa
Department of Computer Science, College of Computer Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia

Article Information

Submitted : 18 Mar 2024
Reviewed: 24 Mar 2024
Accepted : 1 Apr 2024

Keywords

Arabic; Story ending generation; crosslingual transfer learning

Abstract

In the field of natural language processing, the task of generating story endings (SEG) requires not only a deep understanding of the narrative context but also the ability to formulate coherent conclusions. This study delves into the use of crosslingual transfer learning to address the challenges posed by the scarcity of Arabic data in SEG, proposing the utilization of extensive English story corpora as a solution. We evaluated the efficacy of multilingual models, such as mBART, mT5, and mT0, in generating Arabic story endings, assessing their performance in both zero-shot and few-shot scenarios. Despite the linguistic complexities of Arabic and the inherent challenges of crosslingual transfer, our findings demonstrate the potential of these multilingual models to transcend linguistic barriers, significantly contributing to the domain of natural language processing across different languages. This research has significant implications for generating creative text and improving multilingual natural language processing in resource-limited language contexts.

A. Introduction

Storytelling is a key way people share information and connect with each other. We often remember and understand things better when they are told as stories. But when it comes to creating stories by themselves, computers are not very good at it yet. The area of making stories with computers, termed computational narrative, is getting more attention nowadays thanks to emerging technologies. This is important because it can make it easier for people to interact with smart systems. When computers can tell stories, it helps them communicate better with humans, and also helps advance in how computers process and understand human language.

Automatic story generation offers significant advantages across various applications. In entertainment, it enables the production of numerous stories with minimal effort. In education, it allows for the customization of stories to meet learners' specific needs. Additionally, in the gaming industry, interactive storytelling enhances the appeal and engagement of games [1], to name a few.

Story ending generation (SEG) tackles the problem of creating the final part of a story that is incomplete. This task requires a computer to understand what has happened in the story so far and to come up with a suitable ending. SEG is challenging because the computer needs to deeply understand language to figure out the story's setting, characters, and what they do. It also has to know about everyday knowledge and how actions are linked to feelings. The ending it makes should fit smoothly with the rest of the story, without any mismatch or conflict with earlier parts of the story.

SEG attracted the interest of researchers in the natural language processing (NLP) field with the introduction of ROCStories (cs.rochester.edu/nlp/rocstories/), a large corpus of commonsense stories. This story dataset is structured such that each entry consists of exactly five sentences: the initial four set the context, while the fifth offers the ending “to be completed” (see Table 1 for an example).

Table 1. Example of a full five-sentence story from ROCStories

Story title:	**The Hurricane**
Morgan and her family lived in Florida. They heard a hurricane was coming. They decided to evacuate to a relative's house. They arrived and learned from the news that it was a terrible storm. They felt lucky they had evacuated when they did.	

The authors behind ROCStories proposed the “Story Cloze Test”, a framework designed to assess commonsense reasoning in the context of story understanding, generation, and script learning [2]. It involves a system being tasked with selecting an appropriate ending for a story composed of four sentences.

Since then, research in SEG has focused on using transformers and graph-based methods for context analysis, alongside incorporating adversarial training and external commonsense knowledge for more coherent and realistic endings. Interested readers may consult [1], for further details.

Arabic is spoken by over 300 million people as their first language and serves as the liturgical language for approximately 1.8 billion Muslims globally [3]. It's the official language in 26 countries and is also one of the official languages used by the United Nations. However, Arabic does not have much support in the field of computer-based language generation. This study is the first to look into solving SEG problem in Arabic. The complex structure of Arabic makes it hard for computers to understand and generate text in Arabic. To work on SEG in Arabic, computers need to really understand the story and be able to create endings that make sense and fit the story. This research explores the using of cross-lingual transfer learning to generate story endings in Arabic.

The remainder of the paper is organized as follows. Section B covers the characteristics of the Arabic language and the challenges it poses to SEGs. We then look at cross-lingual transfer learning and how it may be used for handling Arabic SEG in section C. The details of our approach is presented in Section D. In section E, we present and discuss the results of the experiments. Finally, we conclude in Section F.

B. Characteristics of Arabic Langauge and the Challenges Posed

Writing story endings in Arabic poses several challenges, primarily due to the language's unique characteristics and the complexities involved in understanding and generating narrative content. Some of these challenges include:

- *Morphological Complexity.* Morphology, the study of word structure, is fundamental to linguistics. Arabic morphology is notably complex yet systematic, characterized by derivation and agglutination [4]. Derivation, in particular, is a potent process that contributes to Arabic's vast vocabulary, making it one of the richest languages in terms of lexical diversity. This complexity makes it difficult for computational models to accurately parse and understand words, affecting the generation of coherent story endings.
- *Contextual Nuances.* Arabic language and storytelling are deeply embedded with cultural and contextual nuances. Capturing the contextual nuances in Arabic storytelling involves understanding the cultural, historical, and social intricacies inherent in the Arabic language and narrative traditions. This includes the appropriate use of cultural references, proverbs, and the depiction of social hierarchies and relationships. For automated story generation, it's crucial to produce endings that are linguistically precise and resonate with the cultural and contextual values of Arabic-speaking audiences, ensuring authenticity and engagement. For instance, the hierarchical and relational aspects of Arabic societies, such as respect for elders and honor, often play a crucial role in narratives and need to be accurately reflected in story endings to maintain authenticity.
- *Dialectal Variations.* Arabic has many dialects, which can significantly differ from the Modern Standard Arabic (MSA) used in formal writing and media. Ensuring that the generated story endings align with the specific dialect or language style of the preceding story can be challenging. Arabic dialects differ by region, and there are no available dictionaries for their vocabulary or written rules for the words that are specific to those dialects [5][6]. Making it a challenging task. However, in this work we will be dealing with MSA.
- *Limited Computational Resources*. Compared to languages such as English, there are fewer computational resources available for Arabic, such as large annotated datasets and advanced language models. This limitation hampers the development of effective SEG models for Arabic. Several studies have been adversely affected by the scarcity of resources in the Arabic language,

including research on text simplification [7], answering why-questions [8], and the development of judicial support systems [9].

- *Semantic Ambiguity*. Arabic sentences can often have multiple interpretations due to the language's semantic richness and flexible syntax. This ambiguity poses a challenge in ensuring that the generated endings have the intended meaning and align with the story's context. Consider the Arabic sentence, "أكل الرجل التمساح", literally "ate the man the crocodile". Determining who was eaten in Arabic can be challenging due to the language's allowance for both VSO (Verb-Subject-Object) and VOS sentence structures. Without contextual clues, differentiating between these meanings relies solely on the correct application of diacritical marks [10]. Unfortunately, MSA is typically written without diacritical marking [11].
- *Long-Range Dependencies*. Arabic stories, like those in other languages, can have complex plots where the ending depends on events and characters introduced early in the narrative. This creates a complex web of long-range dependencies that are crucial for a story's conclusion to feel relevant and fulfilling. For automated systems generating story endings, recognizing and integrating these dependencies is key. The system must analyze the narrative in its entirety, understanding how early elements influence the ending. This demands not only a deep linguistic and cultural understanding but also advanced computational capabilities to ensure the endings are coherent and satisfying. For instance, in a story where a boy receives a magical seed, the seed might not seem important until it grows into a tree that resolves the story's main conflict. An automated system creating the ending needs to recognize the seed's significance from its introduction to ensure the story wraps up in a way that makes sense and feels complete.

Addressing these challenges requires innovative approaches in natural language processing, including advanced models that can handle Arabic's linguistic features, as well as efforts to create and curate more comprehensive computational resources for the Arabic language.

C. Cross-Lingual Transfer Learning

Transfer learning, including its cross-lingual variant, enables the application of knowledge from well-resourced domains or languages, such as English, to enhance problem-solving in less-resourced areas. This involves training a multilingual model on a task in a language with abundant data and then applying it to perform the same task in another, less-represented language.

Multilingual language models leverage their pre-training on diverse languages to generalize across languages, even when provided with annotated data only in a single language for the task at hand. Although cross-lingual transfer learning has proven effective in tasks related to language comprehension, such as classification [12][13] and extractive question answering [14], its performance in text generation falls behind [15]. In text generation tasks, cross-lingual transfer approaches often produce incoherent text, generate content in a language other than the intended target language [16], and encounter challenges such as catastrophic forgetting [17]. Additionally, performance in languages distant from English, such as Arabic, significantly lags behind those similar to English [18].

D. Research Method

The process of generating story endings can be defined as follows: when provided with a story context $X = \{x_1, x_2, \ldots, x_n\}$, comprising n sentences, the objective is to produce a single sentence Y that serves as the conclusion to the story:

$$Y^* = argmax\ P(Y \mid X)$$

The objective of this study is to evaluate the efficacy of cross-lingual story ending generation from English to Arabic through the utilization of various multilingual generative models. We specifically focus on two settings:

1. *Zero-shot cross-lingual transfer:* In this setting, models are trained on an English story corpus and then directly evaluated on their ability to generate Arabic endings, without prior training on Arabic story data.
2. *Few-shot cross-lingual transfer*: Here, we explore how the performance of the generative models improves when they are provided with a limited amount of Arabic data during training. By incorporating a small set of examples, we aim to determine whether the models can better adapt to the nuances of Arabic storytelling and produce more accurate and coherent endings.

This assessment aims to gauge the effectiveness of leveraging data available in English as a high-resource language to benefit Arabic, which has lower resources, in story ending generation. Figure 1 illustrates the methodology employed to develop both the zero-shot and few-shot models within the context of crosslingual transfer learning for the generation of Arabic story endings.

Baselines

We engaged in finetuning the following multilingual models:

- *mBART*[19]: An encoder-decoder Transformer [20] model pretrained on large-scale monolingual corpora in many languages. It stands out as one of the first multilingual models to have been pretrained on Arabic.
- *mT5* [16]: a variant of the T5 (Text-To-Text Transfer Transformer) [21] model which was trained to perform a wide range of natural language processing tasks by framing each task as a text-to-text problem. mT5 extends the capabilities of T5 by training it on a multilingual dataset, enabling it to understand and generate text in multiple languages.
- *mT0* [22]: a multitask prompted fine-tuning (MTF) variant of mT5 capable of following human instructions in 46 languages exclusively present in the pretraining corpus.

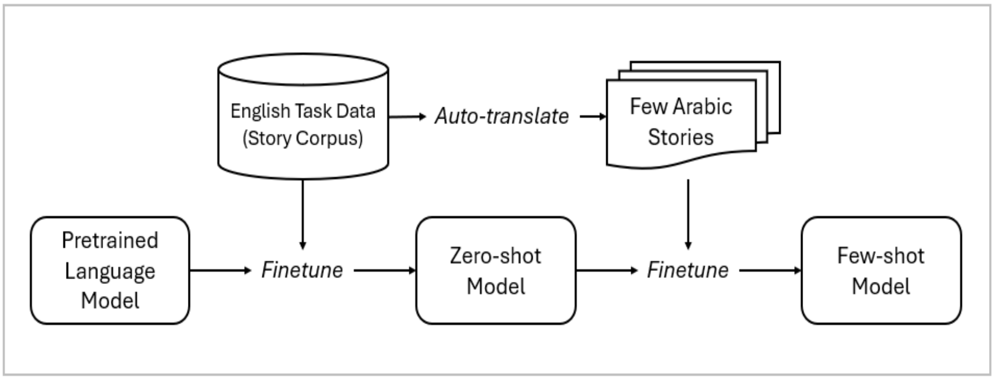

Figure 1. Training zero-shot and few-shot Arabic story ending generation models.

Data

For training our models we used The Rochester Stories Corpus (ROCStories) [2], a dataset consists of a collection of short stories designed for various language understanding tasks, particularly in the domain of story understanding and generation. This dataset contains pairs of story contexts and associated story endings. Each story context provides a brief narrative setting, while the corresponding story ending completes the narrative by providing a resolution of the story. The ROCStories dataset comprises 90,000 stories. From this corpus, we designate 5,000 stories for testing purposes. For Arabic few-shot fine-tuning, we allocate 2,000 stories, split evenly between training and validation sets, with 1,000 stories in each. Additionally, we reserve 5,000 stories for English fine-tuning validation. The remaining stories are utilized for English fine-tuning.

The designated test set and Arabic few-shot fine-tuning sets undergo automatic translation to Arabic using Google Translate [23]. However, to circumvent potential issues arising from transliterated names [7], we initially extract English person names using a Named Entity Recognition (NER) model [24] and subsequently substitute them with Arabic equivalents before translation.

E. Result and Discussion

In this section, we present the results of generating Arabic story endings employing multilingual language models within crosslingual transfer settings. Specifically, we opted for the mBART.cc25 model consisting of 24 layers and 610 million parameters. For mT5 and mT0, we employed equivalent base models featuring 24 layers and 580 million parameters. Training was conducted on an NVIDIA Tesla A100 GPU over 3 epochs, utilizing a learning rate of 2e-5 and a batch size of 16.

To assess the model's effectiveness, we adopt the evaluation methodology established in previous studies on English story ending generation [25][26][27]. We utilize BLEU [28], which evaluates n-gram overlap between a generated ending and a reference, reporting BLEU scores for n = 1 and 2 (BLEU-1 and BLEU-2). Table 2 presents the performance results of the fine-tuned models in both zero-shot and few-shot settings. As depicted in the table, mBART exhibited superior performance, followed by mT5, and finally mT0. Few-shot fine-tuning enhanced the performance of all three models. However, while the improvement from zero-shot to few-shot performance was substantial for mBART, it was relatively marginal for mT5 and mT0.

Due to the absence of prior research on Arabic story ending generation, we included the results of English story ending generation in Table 1 for the reader's convenience. Observing the table, we note that the BLEU scores in story ending generation are relatively low compared to other text generation tasks such as translation and summarization. However, while the performance of mBART-fewshot is comparable to that of English SEG results, the performance of other models lags behind. This discrepancy can be attributed to several factors. Firstly, the limited suitability of the BLEU metric for SEG. Given that story ending generation is an open-ended task not confined to a predefined "correct" ending, a high-quality ending may not necessarily receive a high BLEU score if it differs from the gold ending. Secondly, the complexity of the SEG task itself. Generating appropriate endings requires a model to not only comprehend the story context but also to possess understanding of nuanced elements such as emotions and commonsense knowledge. Thirdly, the inherent complexity of the Arabic language impacts both story context understanding and story ending generation. Lastly, the inherent challenges in crosslingual text generation contribute to the performance disparities observed.

Table 2. Performance of story ending generation models, with the first three models being English story end generators.

No	Model	*BLEU-1*	*BLEU-2*
1	SSAP [26]	26.01	10.29
2	SHGN [27]	25.60	9.40
3	MGCN-DP [25]	24.60	8.60
4	mBART-*zeroshot*	14.26	2.40
5	mT5-*zeroshot*	13.24	2.30
6	mT0-*zeroshot*	12.86	2.45
7	mBART-*fewshot*	21.18	6.35
8	mT5-*fewshot*	16.61	3.85
9	mT0-*fewshot*	13.36	3.12

We also assessed the generated endings from a linguistic perspective. Given the challenges highlighted in previous studies regarding crosslingual transfer in text generation, particularly the tendency to produce content in the wrong language, we endeavored to quantify the percentage of non-Arabic words in the generated endings. While the presence of non-Arabic content in the zero-shot models of mT5 and mT0 was negligible, mBART's zero-shot model exhibited a notable issue with 71.06% non-Arabic content. Despite this, the model was able to conclude Arabic stories with appropriate endings in terms of meaning. Following few-shot training, all three models achieved a 0% non-Arabic content score. Table 3 presents an example of a generated story ending by each model given a specific story context.

Table 3. Samples of story endings generated by the finetuned multilingual models.

Story Context	*Model*	*Generated ending*
أراد محمد أن يخوض مغامرة. طلب من أصدقائه الانضمام إليه. قرروا الذهاب في مغامرة الرحلات في الغابة. لقد خططوا لكل شيء بشكل مثالي.	mBART-*zeroshot*	They had a great time .
	mT5-*zeroshot*	لقد انتهت مغامرة رائعة.
	mT0-*zeroshot*	لقد سافروا معه لمغامرة ممتعة.
	mBART-*fewshot*	لقد استمتعوا بالرحلة
	mT5-*fewshot*	محمد قام بمغامرة رائعة.
	mT0-*fewshot*	لقد سافروا إلى الغابة لعدة أسابيع.
أراد محمد أن يكون فكاهي مرتجل. وجد مسرحًا محليًا يقدم دروسًا. تلقى محمد دروسًا للمبتدئين وحسن مهاراته. شق محمد طريقه إلى صفوف متقدمة مع فنانين متمرسين.	mBART-*zeroshot*	They gave him a job as a فكاهي مرتجل
	mT5-*zeroshot*	وفي حين يقدم دروس خاصة قدم دروس خاصة.
	mT0-*zeroshot*	حقق محمد النجاح الذي حققه على خشبة المسرح.
	mBART-*fewshot*	كان محمد أول فكاهي مرتجل على خشبة المسرح
	mT5-*fewshot*	محمد يصبح فكاهي مرتجلًا.
	mT0-*fewshot*	لقد حقق محمد النجاح في هذا المسرح.

To conduct a more comprehensive linguistic evaluation of the generated endings, we utilized the Arabic Grammar Error Detection (GED) model [29]. This model allowed us to calculate the linguistic scores of the models' outputs. As depicted in Figure 2, the linguistic scores of all three models were improved through few-shot fine-tuning. However, the impact of few-shot training on language correctness was more pronounced in mBART compared to mT5 and mT0. This discrepancy may be attributed to the fact that the linguistic scores of their zero-shot counterparts were initially high.

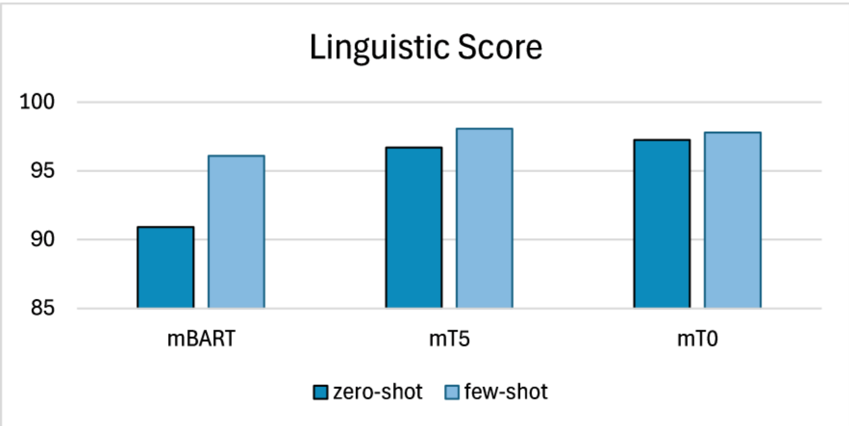

Figure 2. Linguistic scores of generated story endings.

F. Conclusion

This study covered Arabic story ending generation using multilingual language models within crosslingual transfer learning settings. Our investigation revealed promising advancements in leveraging English resources to enhance Arabic text generation capabilities. Through experimentation and evaluation, we demonstrated the efficacy of multilingual models such as mBART, mT5, and mT0 in generating

Arabic story endings, both in zero-shot and few-shot scenarios. Despite encountering challenges such as linguistic complexities and crosslingual transfer limitations, our findings show the potential of multilingual models in bridging language barriers and facilitating natural language understanding and generation tasks across diverse languages.

G. References

[1] A.I. Alhussain, and A.M. Azmi, “Automatic story generation: A survey of approaches,” *ACM Computing Surveys (CSUR)*, 2021, vol. 54, no. 5, pp. 1-38.

[2] N. Mostafazadeh, N. Chambers, X. He, D. Parikh, D. Batra, L. Vanderwende, P. Kohli, and J. Allen. “A corpus and cloze evaluation for deeper understanding of commonsense stories,” In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2016, pp. 839–849.

[3] A.M. Azmi, and A. Alsaiari, “A calligraphic based scheme to justify Arabic text improving readability and comprehension”, *Comput in Human Behavior*, 2014, vol. 39, pp. 177–186.

[4] A.M. Azmi, and E.A. Aljafari, “Universal web accessibility and the challenge to integrate informal Arabic users: a case study.” *Universal Access in the Information Society (UAIS)*, 2018, vol. 17, pp. 131–145.

[5] A.M. Azmi, and E.A. Aljafari, “Modern information retrieval in Arabic—catering to standard and colloquial Arabic users,” *J. Information Sciences*, 2014, vol. 41, no. 4, pp. 506–517.

[6] N. Alshenaifi, and A. Azmi, “Arabic dialect identification using machine learning and transformer based models: Submission to the NADI 2022 Shared Task,” *Proceedings of the The Seventh Arabic Natural Language Processing Workshop (WANLP)*, pp. 464–467.

[7] S. S. Al-Thanyyan and A. M. Azmi, “Simplification of Arabic text: A hybrid approach integrating machine translation and transformer-based lexical model,” *Journal of King Saud University - Computer and Information Sciences*, 2023, vol. 35, no. 8, article no. 101662.

[8] T. Alwaneen, and A.M. Azmi, “Stacked dynamic memory-coattention network for answering why-questions in Arabic,” *Neural Comput and Appl*, 2024, pp. 1-17. DOI: 10.1007/s00521-024-09525-0.

[9] H.A. Almuzaini, and A.M. Azmi, “TaSbeeb: A judicial decision support system based on deep learning framework,” *Journal of King Saud University - Computer and Information Sciences*, 2023, vol. 35, no. 8, article no. 101695.

[10] Z.M. Mannaa, A.M. Azmi, and H.A. Aboalsamh, “Computer-assisted i‘raab of Arabic sentences for teaching grammar to students,” *Journal of King Saud University - Computer and Information Sciences*, 2023, vol. 34, no. 10, pp. 8909–8926.

[11] A.M. Azmi, R.M. Alnefaie, and H.A. Aboalsamh, “Light diacritic restoration to disambiguate homographs in modern Arabic texts,” *Transactions on Asian and Low-Resource Language Information Processing*, 2021, vol. 3, no. 3, pp. 1-14.

[12] I. Li, P. Sen, H. Zhu, Y. Li, and D. Radev, "Improving Cross-lingual Text Classification with Zero-shot Instance-Weighting," in *Proceedings of the 6th Workshop on Representation Learning for NLP (RepL4NLP-2021)*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2021, pp. 1–7.

[13] G. Lai, B. Oguz, Y. Yang, and V. Stoyanov, "Bridging the domain gap in cross-lingual document classification," Sep. 2019.

[14] M. Artetxe, S. Ruder, and D. Yogatama, "On the Cross-lingual Transferability of Monolingual Representations," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2020, pp. 4623–4637.

[15] T. Li and K. Murray, "Why Does Zero-Shot Cross-Lingual Generation Fail? An Explanation and a Solution," in *Findings of the Association for Computational Linguistics: ACL 2023*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2023, pp. 12461–12476.

[16] L. Xue *et al.*, "mT5: A Massively Multilingual Pre-trained Text-to-Text Transformer," in *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2021, pp. 483–498.

[17] T. Vu, A. Barua, B. Lester, D. Cer, M. Iyyer, and N. Constant, "Overcoming Catastrophic Forgetting in Zero-Shot Cross-Lingual Generation," in *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2022, pp. 9279–9300.

[18] M. Zhao *et al.*, "A Closer Look at Few-Shot Crosslingual Transfer: The Choice of Shots Matters," in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2021, pp. 5751–5767.

[19] Y. Liu *et al.*, "Multilingual denoising pre-training for neural machine translation," *Trans Assoc Comput Linguist*, 2020, vol. 8, pp. 726–742.

[20] A. Vaswani *et al.*, "Attention is all you need," *Advances in Neural Information Processing Systems* (NIPS 2017), 2017, vol. 30.

[21] C. Raffel *et al.*, "Exploring the limits of transfer learning with a unified text-to-text transformer," *J Machine Learning Research*, 2020, vol. 21, pp. 1–67.

[22] N. Muennighoff *et al.*, "Crosslingual Generalization through Multitask Finetuning," 2022, *ArXiv Preprint arXiv:2211.01786*.

[23] "Google Translate." Accessed: Mar. 16, 2024. [Online]. Available: https://translate.google.com/

[24] J. Devlin, M.-W. Chang, K. Lee, K. T. Google, and A. I. Language, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *Proceedings of NAACL-HLT*, 2019, pp. 4171–4186.

[25] Q. Huang *et al.*, "Story ending generation with multi-level graph convolutional networks over dependency trees," *The Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21)*, 2021, pp. 13073–13081.

[26] Y. Liu, Q. Huang, J. Li, L. Mo, Y. Cai, and Q. Li, “SSAP: Storylines and Sentiment Aware Pre-Trained Model for Story Ending Generation,” *IEEE/ACM Trans Audio Speech Lang Process*, 2022, vol. 30, pp. 686–694.

[27] J. Wang *et al.*, “Incorporating Commonsense Knowledge into Story Ending Generation via Heterogeneous Graph Networks,” In *International Conference on Database Systems for Advanced Applications*, 2022. Cham: Springer International Publishing, pp. 85–100.

[28] K. Papineni, S. Roukos, T. Ward, and W. Zhu, “BLEU: A method for automatic evaluation of machine translation,” *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)*, 2002, pp. 311-318.

[29] B. Alhafni, G. Inoue, C. Khairallah, and N. Habash, “Advancements in Arabic Grammatical Error Detection and Correction: An Empirical Investigation,” in *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2023, pp. 6430–6448.