Brand Consistency for Multilingual E-commerce Machine Translation

Bryan Zhang bryzhang@amazon.com
Stephan Walter sstwa@amazon.com
Saurabh Chetan Birari sbbirari@amazon.com
Ozlem Eren ozleme@amazon.com
Amazon.com, USA

Abstract

In the realm of e-commerce, it is crucial to ensure consistent localization of brand terms in product information translations. With the ever-evolving e-commerce landscape, new brands and their localized versions are consistently emerging. However, these diverse brand forms and aliases present a significant challenge in machine translation (MT). This study investigates MT brand consistency problem in multilingual e-commerce and proposes practical and sustainable solutions to maintain brand consistency in various scenarios within the e-commerce industry. Through experimentation and analysis of an English-Arabic MT system, we demonstrate the effectiveness of our proposed solutions.

Keywords: multilingual e-commerce, brand consistency, machine translation, data management, Arabic language

1 Introduction

The brand consistency problem in machine translation is often viewed as a terminology enforcement issue. Previous studies (Dinu et al., 2019; Post and Vilar, 2018; Susanto et al., 2020; Wang et al., 2021; Ailem et al., 2021) have proposed terminology constraint mechanisms to tackle this problem. However, these mechanisms assume a static terminology and fail to fully address the challenges encountered in an industry setting. In reality, the localized brand entries in the terminology bank are continuously expanding, being deleted, or undergoing edits, with different aliases taking distinct forms across language pairs.

While brands from larger companies often have well-established localized forms, smaller brands frequently lack such localized versions for specific language pairs. As a result, there is a constant need for fixing or updating brand names in translations. Addressing brand-related issues may require retraining or further fine-tuning of the production machine translation engine (Kanavos and Kartsaklis, 2010; Caskey and Maskey, 2013; Luo et al., 2022), however, the turnaround time to fix these issues can be unacceptable to users and may not align with the business requirements.

To tackle these challenges, we first analyze the brand localization phenomena across various language pairs to simplify the MT brand-handling problem in multilingual e-commerce, then we propose a systematic framework to enforce brand consistency at scale in industrial e-commerce machine translation systems. Furthermore, we provide experimental results showcasing the successful application of our proposed framework in an English-Arabic machine translation setting.

Proceedings of Machine Translation Summit XIX, Vol. 2: Users Track, pages 173–182
September 4–8, 2023, Macau SAR, China.

2 Cross-lingual brand preservation and transformation

To simplify the large-scale MT brand handling problem in E-commerce, we first propose to classify language pairs into two groups based on the major pattern of how brand terms need to be localized for a given language pair. These groups are brand preservation and brand transformation language pairs.

Brand Preservation (BP) language pairs are language pairs where the brand terms from the source language typically remain unchanged in the target language. These language pairs often share similar writing systems, such as English and German. For example, brand term *Adidas* has the same form in English and German. The majority of language pairs belong to this group in the context of E-commerce.

Brand Transformation (BT) language pairs are language pairs where the brand terms in the source language need to be transformed into a different localized form in the target language. This is often the case for language pairs with different writing systems, such as English-Arabic Al'Awadhi (2014), English-Japanese, and German-Greek.

Translation:
*Appl*e - 苹果 [En-Zh]
American Standard - امريكان ستاندرد [En-Ar]
Transliteration:
Adidas - اديداس [En-Ar]
Creation:
G-Star RAW - ستار رو-G [En-Ar]
BMW - 宝马 (precious horse) [En-Zh]

Table 1: Brand Transformations and Examples

We observe brand terms are usually transformed through translation, transliteration and creation as Table 1 illustrates. Same brand terms can be transformed in different ways for different language pairs. For example: *Heineken*-ハイネケン [De-Ja] (Transliteration) and *Heineken*-喜力 [En-Zh] (Creation, meaning Happiness Power).

However, the need for brand transformations can also arise from business requirements or cultural considerations of a given region. For example, customers reading product information in **right-to-left languages** such as Arabic and Hebrew need to alter their reading direction whenever encountering a brand term in Latin-based scripts. In such cases, it is often preferred to have brands transformed in the target language to cater to the reading habits of the customers.

3 Systematic MT brand consistency enforcement for e-commerce translation

An overview of our proposed machine translation (MT) brand consistency enforcement framework is illustrated in Figure 1, the section 3.5 will discuss this framework in detail. The main components of the system are listed as following:

i Brand Mapping (BMCache)

ii Post-editing translation management system (TMS)

iii Heuristics-based approach (HEU)

iv Placeholder approach (PH)

v Data Augmentation approach (DA)

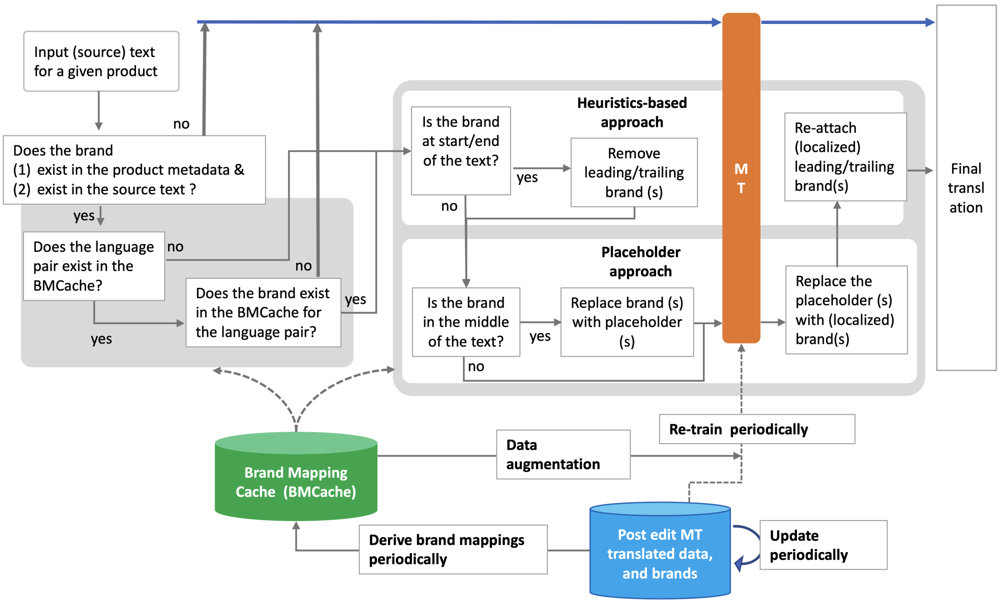

Figure 1: Systematic MT brand consistency enforcement framework for product information texts

3.1 BMCache and Post-editing TMS

The BMCache serves as a comprehensive terminology bank that stores brand terms along with their corresponding localized forms. It is continuously updated with new brand terms, localization forms, and changes derived from a post-editing Translation Management System (TMS). Although detailed discussions about the BMCache and TMS are beyond the scope of this paper, it is worth noting that in the context of e-commerce, brand terms can have multiple valid localized forms in different orthographic variations based on the target language. For instance, English brands can have multiple localized forms in Arabic due to transliteration variations. Similarly, in Japanese, the complexity of the written systems (including Romaji, Kanji, Hiragana, and Katakana)Unger (1996) can lead to multiple localized forms for an English brand, such as *Onitsuka* (Romaji) or オニツカ (Katakana).

To ensure consistency and standardization in brand mappings, a TMS should establish quality assurance guidelines that guide translators to adhere to the localized brand forms available in the BMCache. Once the localized brand forms are validated and accepted as the ground truth, they are stored in the BMCache and can be effectively utilized for communication with various stakeholders.

3.2 Data augmentation (DA) Solution

We first propose the data augmentation (DA) solution which can improve the brand handling capability of MT by incorporating localized brand forms in the target language through regular re-trainings. However, caution must be exercised when augmenting the re-training data with brand mappings from the full BMCache. It is recommended that MT practitioners only utilize brands from the BMCache that do not create ambiguity with generic word translations. For instance, in the Spanish-English language pair, the brand terms "Crema" and "Bebe" should be preserved in the English translation, however, they need to be translated as "cream" and "baby" in English as genetic words. In such cases, MT is expected to learn to disambiguate through parallel training data that contains the brands in the relevant context.

The DA strategy can also be applied periodically to update MT models at a fixed cadence. Alternatively, it can be used in an ad-hoc manner for models that require fixing due to a substantial influx of newer brands that consistently need to be localized in translations.

3.3 Heuristics-based (HEU) solution

Data sample size	Product info type	Brand position in the text		
		Start	*Middle*	*End*
∼half million	Titles	**96%**	3%	1%
∼half million	Desc.	27%	**72%**	1%
∼half million	BP	20%	**75%**	5%

Table 2: Brand term position in the product info texts: Titles, Descriptions (Desc.) and Bullet-pointts (BP)

Table 2 provides a breakdown of the positions of brand terms within e-commerce product information. It is worth noting that in 96% of cases, the brand appears at the beginning of the product titles. Exploiting this positional heuristic of brand terms, we propose a straightforward solution based on heuristics. When a brand term is located at the start and/or the end of a title, our approach involves temporarily removing the brand, translating the remaining text using MT, and subsequently reattaching the appropriate localized brand (from the BMCache) to the start or end of the title. This solution eliminates the need for MT re-training or significant engineering modifications while ensuring a quick turnaround time for brand localization fixes.

3.4 Placeholder (PH) Solution

As a robustness feature, we also propose a placeholder solution for other cases where brand terms are located at flexible positions other than the start or the end of product information texts. This is especially useful to product descriptions and bulletpoints where brand terms usually occur in the middle of texts. Based on the previous study Post et al. (2019), we propose to check through the BMCache first, serialize and replace the brand terms with `[PLACEHOLDER]` token(s). The neural machine translation system needs to be placeholder-aware through training for the following steps: (i) read the `[PLACEHOLDER]` token(s) in the source, (ii) predict the position of the `[PLACEHOLDER]` token(s), (iii) place the `[PLACEHOLDER]` token(s) at the predicted position(s) in the target translation. After the machine translation process, the `[PLACEHOLDER]` token(s) is replaced with the proper localized brand form fetched from the BMCache.

This solution adapts to the dynamic nature of the BMCache and localized brand forms across language pairs. It also provides flexibility to fix ad-hoc brand localization issues within short turnover time.

3.5 Overview of the Brand Consistency Enforcement Framework

Given the complexity of the MT brand handling problem in E-commerce and characteristics of the BP and BT language pairs, we propose a systematic MT brand term consistency enforcement framework as shown in Figure 1 to ensure brand localization consistency in translation: This approach consists of all three brand handling solutions described in the previous sections, and a universal **brand handling logic** which enables the three solutions to work independently and in combination to tackle most MT brand handling scenarios in the e-commerce industry setting.

Conditions	BT Lang Pairs					BP Lang Pairs			
(1) Input brand exists (in the product meta data) and (2) Exists in source text	Y	Y	Y	Y	N	Y	Y	Y	N
(3) The language pair exists in BMCache	Y	Y	Y	Y	Y	N	N	N	N
(4) Input brand exists in BMCache	Y	Y	Y	N	-	-	-	-	-
(5) Brand (s) exists at the start or end of source text	Y	N	Y	-	-	Y	N	Y	-
(6) Brand (s) exists in the middle of the text	N	Y	Y	-	-	N	Y	Y	-
Brand Handling Solution (s)									
Heuristics (HEU)-based solution	×		×			×		×	
Placeholder (PH)-based solution		×	×				×	×	
Augmented MT handles directly				×	×				×

Table 3: Systematic MT Brand Handling Logic

Brand handling logic: this logic table contains the conditions and solutions as s Table 3 shown, Figure 1 has more detailed illustrations. This logic table highlights the following scenarios:

Scenario 1: For both BT and BP language pairs, when there is no input brand existing in the product meta data or the input brand does not exist in the source text, augmented MT translates the input text directly and handles the brands. MT can be periodically retrained to improve the capacity to handle brand terms.

Scenario 2: If a language pair does not exist in the BMCache, it will be considered as a BP language pair, both HEU an PH-based solutions can be applied; if it exists in the BMCache, it is considered as a PT language pair. In this case, both HEU and PH-based solutions can be applied only if the input brand exists in the BMCache. Otherwise, augmented MT alone handles brands in the text since there is no established brand localized form for the target language.

Scenario 3: For both BP and PT language pairs, if there is no leading or trailing brand (s), either the augmented MT will handle the brands directly in translation or the PH solution will handle the brands if the PH solution is used for that language pair.

Flexible application solution(s): although PH-based solution can handle brand terms at flexible positions in the texts, including leading and trailing brands that the HEU-based solution focuses on, we propose to keep the two solutions separate for the following reasons: Firstly, the HEU solution can handle most brand handling cases on its own for the texts of certain domain such as product titles. Secondly, the HEU-based solution is simpler and does not require re-training, ensuring that leading and trailing brands have the proper localized form in the translation. Lastly, the two solutions are complementary, and they can be incrementally developed to meet the requirements of various e-commerce scenarios for the MT brand consistency improvement.

Scalability to multilingual MT: Our systematic approach to handling brands in MT can be potentially extended to multilingual models. The BMCache can differentiate between the BP and PT language pairs, and is compatible with both the HEU and PH-based solutions. Therefore, with multilingual MT systems, we can simply provide the input language pair as a signal to the proposed systematic brand consistency enforcement framework, then the multilingual MT can utilize the appropriate localized brand names in the translation for that language pair as an MT system of a single language pair. This approach is easily scalable to any number of languages and provides a straightforward means of maintaining brand consistency across all translated content.

4 Experiment

We experiment with English-Arabic language pair using our proposed framework for brand consistency in machine translation because it has more language-specific complexities and challenges.

4.1 Experiment Setup

We train a transformer-based (Vaswani et al., 2017) MT system that is encoder-heavy (20 encoder and 2 decoder layers) (Domhan et al., 2020) using the Sockeye MT toolkit. We use a vocabulary of 32K BPE (Sennrich et al., 2016) tokens. We optimise using ADAM Kingma and Ba (2015) and perform early-stopping based on perplexity on a held-out dev set. We train a model with in-house generic translation data and e-commerce translation data with the above specifications. The model is fine-tuned with only the in-domain e-commerce translations to create a baseline model, M_0. Below, we describe four more model variants using various components from the Figure 1.

1. M_1: We further obtain millions of the PE product data (in-domain data) and add to the original in-domain data. The newer in-domain data has approximate 13% more brand occurrences than the original, we train an updated model M_1 same way as the baseline model using the newer data. This can show us the effectiveness of simply adding more data with proper brand localized forms without additional brand consistency enforcement.

2. M_2: We further incorporate English-Arabic BMCache data to augment the in-domain data in order to improve the model's ability to learn brand mappings. The English-Arabic mapping data accounts for 5% of the indomain data. This model corresponds to the Data Augmentation approach described in Figure 1.

3. M_3: Building on M_2, we extend the M_2 with the HEU component in our framework.

4. M_4: Building on M_3, we add the PH component in our framework to show the performance of the framework to the full extent.

For each product information type, we obtain approximately half a million test cases together with brands that can match the BMCache. The number of unique brands across three product information types is in the range of 60 K to 120 K.

5 Results

	Titles	Descriptions	Bulletpoints
M_0	-	-	-
M_1	+89.9%	+70.1%	+55.3%
M_2	+96.9%	+72.5%	+57.0%
M_3	+239.4%	+105.3%	+81.8%
M_4	+247.0%	+175.1%	+155.0%

Table 4: Precision of brand terms consistently localized in translation as per entry in BMCache

Table 4 presents the precision scores of the proper localized brand forms in the translations checking against the ground truth in the BMCache. By simply adding more data with brand localized forms without additional brand consistency enforcement, the precision of proper brand localized forms in the translation of M_1 has improved across all titles, descriptions and bullet points compared with the baseline model M_0. Thereafter, we see that M_2 continues to improve when the training data is augmented with the brand data from BMCache.

Finally, after applying the heuristics (HEU) and the placeholder (PH) solutions, the results show that M_3 and M_4 model can achieve close to nearly perfect brand localization consistency in translation for product titles, and precision is also increased by a large margin for product descriptions and bulletpoints. However, we observe cases of missing placeholder tokens in descriptions and bulletpoints, especially when there are three or more placeholder tokens inserted. This issue is also highlighted in the previous study Post et al. (2019) on using placeholder features.

	Titles	Descriptions	Bulletpoints
M_0	-	-	-
M_1	**+90.6%**	+2.6%	+27.8%
M_4	+74.5%	**+6.3%**	**+29.1%**

Table 5: BLEU Scores on Test Sets

	Titles	Descriptions	Bulletpoints
M_0	-	-	-
M_1	**+41.6%**	+0.6%	+41.9%
M_4	+34.3%	**+2.1%**	**+43.0%**

Table 6: ChrF Scores on Test Sets

Table 5 and 6 presents the BLEU and chrF scores of the generic translation quality disregarding the enforcement of consistent brand translations. We see that the M_4 system that makes use of our brand consistency MT framework achieve the best results for product descriptions and bulletpoints but not the titles.

We have further conducted online A/B testing in a store with English-Arabic MT. Customers in this store typically shop in Arabic and use localized versions of brand names while browsing for branded products. Customers are presented with different versions of the product translations from the baseline model (M_0), and the update model (M_4) with three brand handling solutions.

After a 4-week A/B testing experiment, the results have shown that the translations from the MT with brand handling solutions (M_4) had a much larger positive impact on the customers' shopping experiences. This indicates the effectiveness of our approach.

6 Related work

Previous approaches to terminology in machine translation are evaluated on translations in the generic domain (Crego et al., 2016; Dinu et al., 2019; Post and Vilar, 2018; Susanto et al., 2020; Wang et al., 2021; Ailem et al., 2021; Zhang et al., 2022). Other work related to lexical constrained machine translation have focused on general named entities and cross-lingual named entities mapping (Ugawa et al., 2018; Yan et al., 2018; Alankar Jain, 2019).

However, brand handling in the e-commerce machine translation has received little attention in the literature. Previous study Guha and Heger (2014) presents various challenges of "non-standard" source language structure in e-commerce specific texts, they highlight the issue of brand names that are lexically ambiguous, and show that their e-commerce MT systems can manage to preserve brands for 90% of the translations. In the paper, we extend their description of other brand translation issues for e-commerce as presented in Table 3.

Other brand handling studies for e-commerce translations focus on the effects of efficiently standardizing and brand localization Dong et al. (2020); Jeong et al. (2019) on Arabic Benmamoun et al. (2016); Abuljadail and Badghish (2021) and Chinese e-commerce Liu et al. (2016).

7 Conclusions

This study examines brand localization in various language pairs and proposes language groups for brand preservation and transformation to simplify machine translation (MT) brand consistency problem. Then we propose a systematic approach for MT brand consistency enforcement for product info texts translation. This approach consists of a universal brand handling logic as a framework and three MT brand handling solutions which can work independently or in combination to address most MT brand handling cases across language pairs in various e-commerce scenarios. The proposed approach is successfully applied in a case study of MT for English to Arabic, with offline and online experiments showing improved effectiveness and customer experiences.

References

Mohammad Abuljadail and Saeed Badghish. 2021. Exploring type of strategies used by global brands to engage the saudi consumers more in brands' facebook pages in saudi arabia in terms of "like, share and comment". المجلة العربية للإدارة, 41(1):405–416.

Melissa Ailem, Jingshu Liu, and Raheel Qader. 2021. https://doi.org/10.18653/v1/2021.findings-acl.125 Encouraging neural machine translation to satisfy terminology constraints. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 1450–1455, Online. Association for Computational Linguistics.

Zachary Chase Lipton Alankar Jain, Bhargavi Paranjape. 2019. Entity projection via machine translation for cross-lingual ner. In *EMNLP/IJCNLP 2019 Computer Science, Mathematics*. EMNLP.

Noura Al'Awadhi. 2014. *Arabic Brand Names: To Translate or Transliterate?* Ph.D. thesis.

Mamoun Benmamoun, Rana Sobh, Nitish Singh, and Francisco Tigre Moura. 2016. Gulf arab e-business environment: Localization strategy insights. *Thunderbird International Business Review*, 58(5):439–452.

Sasha P. Caskey and Sameer Maskey. 2013. https://patents.google.com/patent/US20130007405 Translation cache prediction.

Josep Maria Crego, Jungi Kim, Guillaume Klein, Anabel Rebollo, Kathy Yang, Jean Senellart, Egor Akhanov, Patrice Brunelle, Aurélien Coquard, Yongchao Deng, Satoshi Enoue, Chiyo Geiss, Joshua Johanson, Ardas Khalsa, Raoum Khiari, Byeongil Ko, Catherine Kobus, Jean Lorieux, Leidiana Martins, Dang-Chuan Nguyen, Alexandra Priori, Thomas Riccardi, Natalia Segal, Christophe Servan, Cyril Tiquet, Bo Wang, Jin Yang, Dakun Zhang, Jing Zhou, and Peter Zoldan. 2016. http://arxiv.org/abs/1610.05540 Systran's pure neural machine translation systems. *CoRR*, abs/1610.05540.

Georgiana Dinu, Prashant Mathur, Marcello Federico, and Yaser Al-Onaizan. 2019. https://doi.org/10.18653/v1/P19-1294 Training neural machine translation to apply terminology constraints. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 3063–3068, Florence, Italy. Association for Computational Linguistics.

Tobias Domhan, Michael Denkowski, David Vilar, Xing Niu, Felix Hieber, and Kenneth Heafield. 2020. https://www.aclweb.org/anthology/2020.amta-research.10 The sockeye 2 neural machine translation toolkit at AMTA 2020. In *Proceedings of the 14th Conference of the Association for Machine Translation in the Americas (Volume 1: Research Track)*, pages 110–115, Virtual. Association for Machine Translation in the Americas.

Lily C Dong, Chunling Yu, et al. 2020. Globalization or localization: Global brand perception in emerging markets. *International Business Research*, 13(10):53–65.

Jyoti Guha and Carmen Heger. 2014. https://aclanthology.org/2014.amta-users.3 Machine translation for global e-commerce on eBay. In *Proceedings of the 11th Conference of the Association for Machine Translation in the Americas: MT Users Track*, pages 31–37, Vancouver, Canada. Association for Machine Translation in the Americas.

Insik Jeong, Jong-Ho Lee, and Eunmi Kim. 2019. Determinants of brand localization in international markets. *Service Business*, 13:75–100.

Panagiotis Kanavos and Dimitrios Kartsaklis. 2010. https://aclanthology.org/2010.jec-1.3 Integrating machine translation with translation memory: A practical approach. In *Proceedings of the Second Joint EM+/CNGL Workshop: Bringing MT to the User: Research on Integrating MT in the Translation Industry*, pages 11–20, Denver, Colorado, USA. Association for Machine Translation in the Americas.

Diederik P. Kingma and Jimmy Ba. 2015. Adam: A method for stochastic optimization. volume abs/1412.6980.

Sindy Liu, Patsy Perry, Christopher Moore, and Gary Warnaby. 2016. The standardization-localization dilemma of brand communications for luxury fashion retailers' internationalization into china. *Journal of Business Research*, 69(1):357–364.

Chen Luo, Vihan Lakshman, Anshumali Shrivastava, Tianyu Cao, Sreyashi Nag, Rahul Goutam, Hanqing Lu, Yiwei Song, and Bing Yin. 2022. https://www.amazon.science/publications/rose-robust-caches-for-amazon-product-search Rose: Robust caches for amazon product search. In *The Web Conference 2022*.

Matt Post, Shuoyang Ding, Marianna Martindale, and Winston Wu. 2019. https://www.aclweb.org/anthology/W19-6618 An exploration of placeholding in neural machine translation. In *Proceedings of Machine Translation Summit XVII Volume 1: Research Track*, pages 182–192, Dublin, Ireland. European Association for Machine Translation.

Matt Post and David Vilar. 2018. https://doi.org/10.18653/v1/N18-1119 Fast lexically constrained decoding with dynamic beam allocation for neural machine translation. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 1314–1324, New Orleans, Louisiana. Association for Computational Linguistics.

Rico Sennrich, Barry Haddow, and Alexandra Birch. 2016. https://doi.org/10.18653/v1/P16-1162 Neural machine translation of rare words with subword units. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1715–1725, Berlin, Germany. Association for Computational Linguistics.

Raymond Hendy Susanto, Shamil Chollampatt, and Liling Tan. 2020. https://doi.org/10.18653/v1/2020.acl-main.325 Lexically constrained neural machine translation with Levenshtein transformer. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 3536–3543, Online. Association for Computational Linguistics.

Arata Ugawa, Akihiro Tamura, Takashi Ninomiya, Hiroya Takamura, and Manabu Okumura. 2018. Neural machine translation incorporating named entity. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 3240–3250.

J Marshall Unger. 1996. *Literacy and script reform in occupation Japan: Reading between the lines*. Oxford University Press on Demand.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, ukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems*, 30.

Ke Wang, Shuqin Gu, Boxing Chen, Yu Zhao, Weihua Luo, and Yuqi Zhang. 2021. https://aclanthology.org/2021.wmt-1.85 TermMind: Alibaba's WMT21 machine translation using terminologies task submission. In *Proceedings of the Sixth Conference on Machine Translation*, pages 851–856, Online. Association for Computational Linguistics.

Jinghui Yan, Jiajun Zhang, JinAn Xu, and Chengqing Zong. 2018. The impact of named entity translation for neural machine translation. In *China Workshop on Machine Translation*, pages 63–73. Springer.

Wu Zhang, Tung Yeung Lam, and Mee Yee Chan. 2022. https://doi.org/10.1145/3556677.3556691 Using translation memory to improve neural machine translations. In *Proceedings of the 2022 6th International Conference on Deep Learning Technologies*, ICDLT '22, page 49–54, New York, NY, USA. Association for Computing Machinery.