JOTE Volume 6 Nomor 2 Tahun 2024 Halaman 20-30
JOURNAL ON TEACHER EDUCATION
Research & Learning in Faculty of Education
ISSN: 2686-1895 (Printed); 2686-1798 (Online)

English Syntactic Analysis in Automatic Translation Text: A Study on Sentence Structure Accuracy

Afia Nur Meiza[1], Natasya Miranda Gihar[2], Khairun Nisa[3], Najla Regar[4], Khairunnisa[5], Siti Ismahani[6]
Program Studi Tadris Bahasa Inggris, Universitas Islam Negeri Sumatera Utara
e-mail: afia0304222093@uinsu.ac.id[1], natasya0304222075@uinsu.ac.id[2], khairunnisa040505@gmail.com[3], najla0304222092@uinsu.ac.id[4], khairunnisa0304222082@uinsu.ac.id[5], sitiismahani@uinsu.ac.id[6]

Abstrak

Penelitian ini menyelidiki peran analisis sintaksis dalam meningkatkan akurasi penerjemahan mesin (MT), dengan fokus pada struktur kalimat dalam bahasa Inggris. Penelitian ini menggunakan pendekatan kualitatif, menganalisis 500 kalimat bahasa Inggris yang diterjemahkan ke dalam bahasa Indonesia dan Spanyol menggunakan tiga sistem MT: Google Translate, DeepL, dan Microsoft Translator. Penelitian ini mengidentifikasi masalah sintaksis utama seperti urutan kata, kesepakatan, dan struktur klausa, serta mengevaluasi dampaknya terhadap kualitas terjemahan. Hasil penelitian menunjukkan bahwa meskipun semua sistem menghadapi tantangan dengan struktur kalimat yang kompleks, DeepL lebih unggul dibandingkan yang lain dalam menangani dependensi jarak jauh dan mempertahankan kelancaran. Penelitian ini menekankan pentingnya analisis sintaksis dalam meningkatkan sistem MT dan menyarankan bahwa model hibrida dan multibahasa dapat menawarkan solusi untuk keterbatasan saat ini, berkontribusi pada terjemahan yang lebih baik di masa depan.

Kata Kunci: *Analisis Sintaksis, Terjemahan Otomatis, Tata Bahasa Inggris, Penguraian Sintaksis.*

Abstract

This study investigates the role of syntactic analysis in enhancing the accuracy of machine translation (MT), specifically focusing on English sentence structures. The research employs a qualitative approach, analyzing 500 English sentences translated into Indonesian and Spanish using three MT systems: Google Translate, DeepL, and Microsoft Translator. The study identifies key syntactic issues such as word order, agreement, and clause structure, and evaluates their impact on translation quality. The results show that while all systems faced challenges with complex sentence structures, DeepL outperformed others in managing long-range dependencies and maintaining fluency. The study underscores the importance of syntactic analysis in improving MT systems and suggests that hybrid and multilingual models may offer solutions to current limitations, contributing to better translations in the future.

Keywords: *Syntactic Analysis, Automatic Translation, English Grammar, Syntactic Parsing.*

INTRODUCTION

Over the past few decades, the field of automatic translation (AT) has witnessed significant progress, driven by advancements in computational linguistics, machine learning, and natural language processing (NLP). These innovations have substantially improved the fluency and precision of translations, positioning machine translation (MT) as an essential tool for overcoming language barriers in a globalized world. However, despite these technological strides, one of the persistent challenges in AT remains the accurate representation of syntactic structures during the translation process, especially when translating from English, a language known for its intricate and flexible syntax, into languages with vastly different grammatical systems.

Recent progress in computational linguistics, machine learning, and natural language processing (NLP) has significantly advanced the capabilities of machine translation. Although these technological breakthroughs have led to improvements in both the fluency and accuracy of translations, one persistent issue remains: the accurate translation of syntactic structures, particularly when translating between languages with vastly different grammatical frameworks.

Veselovská (2017) provides a comprehensive exploration of English syntax, focusing on the core structural elements that govern sentence formation. In her work, she emphasizes the importance of understanding both simple and complex syntactic structures, offering insights into how these structures affect sentence processing and meaning. The study also covers various syntactic phenomena, such as word order, clause structure, and syntactic relationships, making it a valuable resource for those interested in the intricacies of English grammar and sentence construction.

The variability in English sentence structures-ranging from simple sentences to highly complex ones with embedded clausescomplicates the task of ensuring syntactic accuracy in MT. Additionally, English sentences often involve dependencies between words and phrases that are not easily translatable into languages with more rigid syntactic structures, such as those that follow a strict subjectverb-object (SVO) order or rely heavily on case-marking (Sutskever, Ilya, Oriol Vinyals, and Quoc V. Le:2014). For automatic translation systems to maintain the meaning, context, and natural flow of the original text, it is essential to preserve the syntactic integrity of the source language. Failing to do so can result in mistranslations, awkward phrasing, or even complete misinterpretation of the text. Thus, achieving precise syntactic analysis is critical for generating high-quality translations that are both grammatically accurate and contextually appropriate.

This research investigates the role of English syntactic analysis in automatic translation, with a specific focus on how sentence structure accuracy impacts translation quality. The study examines various syntactic elements such as word order, verb phrase structure, and clause dependencies and their influence on the performance of translation systems (Sutskever Ilya:2014). By exploring how syntactic parsing algorithms interact with the rules of English grammar, the study aims to identify the challenges posed by English syntax in translation tasks and assess the effectiveness of current methodologies in preserving syntactic correctness. Additionally, the paper emphasizes how recent advancements in natural language processing, especially deep learning and transformer-based models, have the potential to enhance syntactic analysis in machine translation systems.

This research focuses on the relationship between translation and the syntax of complex sentences. It addresses several key topics, including: (1) the distinction between clauses and sentences, (2) the interaction between syntax and translation, (3) the concept of a "clause," (4) the notion of a "phrase," (5) transformational grammar and its application to complex sentences, and (6) adverbial subordinators. The study concludes by summarizing the findings (Koehn, Philipp, and Rebecca Knowles:2017).

As both linguists and translators are well aware, translation is impossible without first analyzing the syntax of the source text and then restructuring it to fit the syntactic norms of the target language. Regarding word order, in English, subordinating conjunctions are typically followed by a noun phrase. In terms of sentence structure, the main (superordinate) clause typically precedes the subordinate clause. However, transformations may allow for the subordinate clause to be placed at the beginning of the sentence. Additionally, in some English non-finite sentences, subordinators can be omitted. Some finite adverbial clauses in English can also be converted into non-finite or verbless clauses, either by omitting certain parts of the subordinate clause or by changing the finite verb into a non-finite form. Lastly, several English subordinators are polysemous, meaning they can introduce different types of clauses, which can present challenges for translators when determining the appropriate translation (Rios, Ana, and Antonio:2010).

In the introduction, the paper highlights that many students struggle with understanding long and complex sentences in English. These sentences are often perceived as particularly challenging, yet they are a common feature of the language, influenced by the inherent characteristics of English itself. In the methodology section, the paper begins by examining the structural aspects of English sentences, focusing on two key dimensions. It then introduces an analytical approach to English sentence structures, exploring various sentence types and using a case study to demonstrate how this method can be applied in practice (Bahdanau, Cho, dan Bengio, 2014). The discussion emphasizes the significance of analyzing sentence structures in various aspects of English language learning, including listening, reading comprehension, cloze tests, translation, oral expression, and writing. The paper concludes by asserting that a deep understanding of sentence structure is crucial for mastering English. It highlights those students, particularly those majoring in English, should actively engage in practices like grammar study, text analysis, and regular reading aloud to enhance their overall language proficiency.

METHOD

This research utilizes a qualitative design to explore the syntactic accuracy of machine-generated translations from English into various target languages, focusing specifically on the structure of sentences. Considering the context-sensitive and intricate nature of syntactic mistakes in automated translations, a qualitative approach is well-suited for delving into the various forms these errors take and understanding the underlying causes. The study aims to identify and examine the types of syntactic variations and inaccuracies present in the machine translations and assess how these affect the overall quality of the translated texts.

Data was gathered from three widely used machine translation systems: Google Translate, DeepL, and Microsoft Translator. A total of 500 English sentences, spanning diverse domains such as academic articles, news, technical documents, and everyday conversations, were selected to ensure a broad range of sentence structures. These sentences were translated into Indonesian and Spanish, chosen for their distinct syntactic features compared to English, such as word order, grammatical agreement, and syntactic rules, thus providing insights into how these systems handle syntactic transformations.

Qualitative content analysis was used to identify and categorize the types of syntactic errors in the translations. Errors were classified into categories including word order problems, agreement issues, omissions, inappropriate additions, and clause structure errors. Each error was examined in context to understand the specific challenges faced by the machine translation systems, such as difficulties with longer sentences or complex clause structures. Additionally, the severity of the errors was evaluated, ranging from minor shifts in word order that preserved meaning to more significant errors that rendered the translation grammatically incorrect or nonsensical. A comparative analysis of the three translation systems helped identify strengths and weaknesses in their handling of syntactic issues.

The findings revealed several key patterns. Complex sentences with multiple clauses or embedded structures posed significant challenges, with Google Translate and Microsoft Translator making more frequent errors compared to DeepL, which performed better with such sentences, though it still struggled with certain subordination structures. The study also found that the target languages introduced specific challenges: while Spanish translations were generally more accurate, both Google Translate and DeepL faced greater difficulty with Indonesian, due to its flexible word order and distinct sentence structure. Agreement errors were common across all systems, especially in sentences with compound subjects or complex verb tenses. Word order issues were more frequent in Google Translate's translations, particularly with longer sentences, while DeepL demonstrated a better ability to manage word order.

Despite producing generally understandable translations, none of the systems replicated the fluency of human translations. Syntactic errors that went uncorrected led to awkward or unclear phrasing, although DeepL's translations were more likely to maintain a natural flow despite occasional errors. To ensure the credibility of the findings, multiple experts in linguistics and machine translation independently reviewed a subset of the translations for consistency. Expert feedback was also solicited to validate the analysis, and a detailed account of the errors was provided to ensure transparency and replicability.

RESULTS AND DISCUSSION

The Role of Syntactic Analysis in Machine Translation

Syntactic analysis involves examining the structure of sentences to understand how their components, such as words, phrases, and clauses, are organized and interrelated. This process is crucial in machine translation (MT) as it directly impacts the precision and naturalness of translations. By analyzing the syntactic structure, the translation system can determine how different elements of the source language relate to each other, which influences how the target language text is generated (Koehn, P:2017).

There are two primary approaches to syntactic analysis in MT: syntax-based MT and neural machine translation (NMT). Syntax-based MT follows established linguistic principles to model sentence structures explicitly, while NMT, which has become increasingly popular, leverages deep learning techniques to learn syntactic patterns from vast amounts of data without relying on predefined rules (Ranathunga et al., 2021). Although NMT has greatly enhanced the fluency of translations, challenges remain when dealing with intricate syntactic constructions, such as long-range dependencies, coordination, and subordination (Dabre et al., 2020). As a result, effective syntactic parsing continues to play a vital role in improving the performance of translation systems, even with the rise of neural network-based approaches.

Machine translation plays a crucial role in enabling communication across different languages by converting text from one language to another. However, ensuring that translations are both accurate and contextually appropriate remains a significant challenge (Pal & Saha, 2015) This review examines the obstacles involved in achieving translation accuracy, particularly in terms of interpreting context, resolving ambiguity in polysemous words, and dealing with idiomatic expressions, cultural differences, and specialized terminology. The article underscores the necessity of preserving grammatical correctness and syntactic integrity, while also capturing the cultural nuances embedded in the original text. Calzolari et al. (2012) discuss the advancements and challenges in machine translation during the Eighth International Conference on Language Resources and Evaluation. The proceedings cover various topics, including the complexities of translating intricate sentence structures, which involve navigating grammatical transformations and syntactic nuances. The issues highlighted in this study are crucial for improving translation systems that deal with complex linguistic forms, underscoring the need for more robust grammatical handling in machine translation.Ultimately, the review highlights the importance of improving machine translation systems to overcome language barriers, promote cross-cultural understanding, and enhance global cooperation.

Rule-based machine translation (RBMT) operates by relying on a set of preestablished linguistic rules and a comprehensive dictionary to convert text from one language to another. The rules for RBMT are crafted manually by linguists and experts in the language, and they cover a wide range of elements, including grammar, syntax, lexical correspondences, and transformation procedures. The system first analyzes the source text, applying these predefined rules, and subsequently generates a translation in the target language (Rivera-

Trigueros:2022). One key benefit of Rule-Based Machine Translation (RBMT) is its ability to offer precise control over grammar and syntax, making it especially valuable in specialized fields such as legal or technical translation, where consistency and adherence to specific rules are crucial. However, RBMT has notable drawbacks. It tends to face challenges with idiomatic expressions, slang, and variations in language use, and requires continuous updates and maintenance of extensive lexical resources. Since the translation process is rule-driven, RBMT lacks the flexibility needed to handle the evolving nature of language.

On the other hand, statistical machine translation (SMT) takes a very different approach by using statistical models to predict the best possible translation. In SMT, the system is trained using large sets of parallel corpora, which are collections of sentences in the source and target languages that have been aligned or matched. During training, SMT systems analyze these corpora and learn patterns of translation based on statistical relationships between words, phrases, and sentence structures in both languages. Once trained, the system generates translations by estimating which translation would be most probable, based on these learned patterns. One of the significant strengths of SMT is its ability to handle diverse and complex linguistic structures across different languages. It can adapt to the statistical regularities of each language pair, even if they differ significantly. However, SMT is not without its challenges. It often requires large, carefully curated linguistic resources, and while it does well with frequent and predictable sentence patterns, it struggles when faced with long-range dependencies or highly contextual translations, such as metaphorical or figurative language.

Neural machine translation (NMT), a more recent development in the field, marks a substantial shift in how machine translation is approached. Rather than relying on predefined rules or statistical mappings, NMT uses deep learning techniques, particularly artificial neural networks, to perform translation. The most successful NMT models, such as those based on transformer architectures, include two main components: the encoder and the decoder. These components work in tandem to process and generate translations, with the encoder first converting the input (source text) into a series of abstract representations, while the decoder then generates the translated output in the target language. An important feature of modern NMT models is the attention mechanism, which enables the system to focus on different parts of the source sentence at various stages of the translation process, allowing it to capture complex contextual dependencies across longer spans of text (D. Shterionov, R. Superbo, P. Nagle, L. Casanellas, T. O'Dowd:2018).

One of the most significant advantages of NMT is its ability to generate translations that are both fluent and contextually accurate. This is because NMT learns to understand and process language in a more holistic way, considering entire sentences or even paragraphs, rather than just individual words or phrases. As a result, NMT is particularly good at handling idiomatic expressions, slang, and other context-dependent language features that rule-based systems and SMT often struggle with. Moreover, NMT has shown impressive results across many different language pairs and is highly adaptable, capable of learning from vast amounts of data without explicit programming for each language combination. However, while NMT outperforms other methods in terms of fluency and contextual accuracy, it still faces challenges, particularly when translating highly specialized or domain-specific content where the training data might be insufficient or imbalanced.

In summary, the field of machine translation has evolved through several stages, from rule-based systems to statistical approaches and finally to the adoption of deep learning models. While RBMT offers accuracy in terms of structure and control, it lacks flexibility and scalability. SMT, on the other hand, provides more flexibility but struggles with long-range dependencies. NMT represents the cutting edge, handling context and idiomatic expressions with impressive accuracy, but it still requires large datasets and faces challenges in highly specialized domains. Each of these systems has contributed to the overall advancement of machine translation, but NMT has emerged as the most promising approach for achieving high-quality, fluent, and contextually appropriate translations across a wide range of languages.

Challenges and Approaches to English Syntactic Structures in Machine Translation

English syntax presents notable challenges for machine translation (MT) due to its flexible word order, extensive use of auxiliary verbs, and intricate subordination structures.

These elements complicate the task of producing accurate translations, especially when converting English text into languages with distinct grammatical rules. Some of the syntactic features of English that pose challenges in translation include:

a) Word Order: While English generally follows a Subject-Verb-Object (SVO) structure, certain constructions such as questions, passive forms, or topicalization can disrupt this pattern. Translation systems must account for these variations to ensure proper sentence construction in the target language.
b) Auxiliary Verb: The frequent use of auxiliary verbs like "do," "have," and "be" in English plays a key role in indicating tense, aspect, modality, and voice. Translating these constructions accurately is critical, especially when the target language has a different approach to tense or aspect.
c) Long-Distance Dependencies: English sentences often feature long-distance dependencies, such as relative clauses or embedded wh-questions, where the relationship between sentence components may not be immediately clear. These complex structures present challenges for translation systems trying to preserve the sentence meaning in the target language.
d) Coordination and Subordination: English distinguishes between coordinated structures (e.g., "and," "or") and subordinated structures (e.g., "because," "although"). Misunderstanding these relationships can

significantly alter the meaning of a sentence and result in incorrect translations.

Machine translation (MT) systems have evolved significantly over the years, with various approaches developed to tackle the complex syntactic challenges inherent in translating languages, particularly English. These approaches, each with their distinct advantages and limitations, aim to address issues related to sentence structure, word order, and grammar. The earliest machine translation systems relied on rule-based methods, where linguists manually crafted extensive sets of grammar rules for sentence parsing and translation. RBMT systems excelled in controlled linguistic environments where the source language adhered to well-defined structures. However, they struggled with the flexible and often ambiguous nature of natural language, particularly with languages like English, where word order and grammar can vary widely depending on context. This rigidity made RBMT systems less effective for real-world translation tasks, where flexibility is crucial for handling diverse sentence constructions.

With the introduction of statistical machine translation, MT systems began to shift from rule-based methods to data-driven approaches. SMT models rely on large bilingual corpora, using statistical algorithms to analyze sentence pairs and estimate translation probabilities. These models learn from data and can identify syntactic patterns across different languages. While SMT improved the ability to handle complex syntactic variations, its reliance on word-level alignments still posed challenges. For instance, SMT struggled with translating long-range dependencies-where elements of a sentence are far apart-or with handling structural mismatches between languages, such as differences in word order or grammatical features like gender or case marking.

The emergence of neural machine translation, particularly with deep learning models such as transformers, marked a significant leap forward in translation quality. NMT systems are capable of learning from vast amounts of parallel text data, and they use attention mechanisms to better capture syntactic relationships, allowing them to handle longer sentences and more intricate syntactic structures. One of the major strengths of NMT is its ability to generate fluent, contextually appropriate translations by implicitly learning linguistic patterns. However, NMT models are not explicitly programmed with syntactic rules, which can sometimes lead to errors in sentence structure, particularly when dealing with complex constructions or highly ambiguous phrases.

To leverage the advantages of both rule-based and neural approaches, contemporary machine translation systems often employ hybrid methods. These models seek to combine the accuracy and structure of rule-based techniques with the flexibility and fluency offered by neural networks. This combination aims to enhance the handling of syntactic dependencies while preserving the natural flow of language in the translation output (Silvestre-Cerdà, Andrés-Ferrer, & Civera, 2011). Hybrid systems are seen as promising for overcoming the weaknesses inherent in both individual approaches, offering greater flexibility and scalability.

Jap (2020) examines the role of syntactic frequency in sentence processing, focusing on how variations in syntactic structures can impact translation systems. In his research, Jap underscores the importance of advanced syntactic parsing techniques to address the complexities involved in translating sentences with non-canonical structures, such as those with embedded clauses or parenthetical elements. These intricate sentence constructions often challenge both rule-based and neural machine translation systems, highlighting the need for more sophisticated methods to achieve accurate and fluent translations.

A fundamental challenge in machine translation is the need to account for the syntactic differences between English and other languages. Variations in word order, subject-object agreement, and case marking, for instance, require MT systems to adapt the English sentence structure to the grammar rules of the target language. Effective handling of these cross-linguistic differences is crucial for generating grammatically accurate translati ons.

English, like many other languages, is prone to syntactic ambiguity. Phrases or sentences can often be interpreted in multiple ways depending on their context, and determining the correct interpretation can be difficult for MT systems. This is particularly evident in noun phrases, multi-clause sentences, and complex syntactic structures that can be interpreted in various ways. Both statistical and neural models encounter considerable difficulties in accurately resolving these ambiguities (Zhai et al., 2013).

As machine translation continues to evolve, there is a clear trend toward using more sophisticated deep learning techniques to handle syntactic challenges. Models based on transformers, along with advancements in self-supervised learning, offer promising solutions for better modeling complex syntactic structures and contextual nuances in translation (Neubig, G., & Dyer:2016). These techniques could improve how translation systems handle long-range dependencies, embedded clauses, and word order discrepancies, ultimately leading to more fluent and accurate translations.

Moreover, multilingual models-those capable of processing multiple languages simultaneously-hold the potential to alleviate some of the challenges posed by syntactic variation. By enabling MT systems to process syntactic structures from several languages in parallel, multilingual models could help overcome the limitations of language-specific syntactic rules, offering more robust translation solutions across a wide range of language pairs (Johnson, M:2017). In conclusion, while significant progress has been made in syntactic analysis for machine translation, continued research and development are necessary to address the remaining challenges. The integration of advanced deep learning techniques and the development of hybrid models will likely play a key role in the next generation of MT systems, helping to break down language barriers and improve communication across linguistic divides.

CONCLUSION

Tthis study highlights the importance of syntactic analysis in improving machine translation (MT) accuracy, particularly when translating English. The complexity of English sentence structures, with its varied word order, auxiliary verbs, and clause dependencies, presents challenges for MT systems, especially when translating to languages with different grammatical rules. Despite advances, such as neural machine translation (NMT), issues like long-range dependencies and syntactic ambiguities remain significant obstacles. Early MT systems, like rule-based machine translation (RBMT), struggled with the dynamic nature of natural language, while statistical machine translation (SMT) improved through data-driven approaches but still struggled with complex structures. NMT, using deep learning and transformers, marked significant progress but faces limitations in handling ambiguous or specialized structures. Hybrid models, combining rule-based and neural approaches, offer promising solutions by blending precision and adaptability. Additionally, multilingual models that process multiple languages at once show potential for improving translation across diverse language pairs. While progress has been made, further innovation in deep learning and hybrid models is essential to address challenges and enhance the system's ability to handle complex syntax, ensuring better translations and fostering global communication.

REFERENCES

Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473.* https://doi.org/10.48550/arXiv.1409.0473

Calzolari, N., Choukri, K., Declerck, T., Doğan, M. U., Maegaard, B., Mariani, J., Moreno, A., Odijk, J., & Piperidis, S. (2012). *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*. European Language Resources Association (ELRA), Istanbul, Turkey.

Dabre, R., Chu, C., & Kunchukuttan, A. (2019). A Survey of Multilingual Neural Machine Translation. *ACM Computing Surveys (CSUR), 53*, 1 - 38. https://doi.org/10.1145/3406095

Jap, B. (2020). *Syntactic frequency and sentence processing in Standard Indonesian* (Doctoral dissertation). University of Groningen. https://doi.org/10.33612/diss.143947876

Johnson, M., Schuster, M., Le, Q. V., Krikun, M., Wu, Y., Chen, Z., Thorat, N., Viégas, F., Wattenberg, M., Corrado, G., Hughes, M., & Dean, J. (2017). Google's multilingual neural machine translation system: Enabling zero-shot translation. *Transactions of the Association for Computational Linguistics, 5,* 339–351. https://doi.org/10.1162/tacl_a_00065

Koehn, P., & Knowles, R. (2017). Six challenges for neural machine translation. *In Proceedings of the First Workshop on Neural Machine Translation* (pp. 28–39). Vancouver: Association for Computational Linguistics. https://doi.org/10.18653/v1/W17-3204

Neubig, G., & Dyer, C. (2016). Neural machine translation and syntax: A thorough evaluation. *Proceedings of the NAACL-HLT* (pp. 13–21).

Pal, A. R., & Saha, D. (2015). Word sense disambiguation: A survey. International Journal of Control Theory and Computer Modeling, 5(3). https://doi.org/10.5121/ijctcm.2015.5301

Ranathunga, S., Lee, E. A., Skenduli, M. P., Shekhar, R., Alam, M., & Kaur, R. (2021). Neural machine translation for low-resource languages: A survey.

ACM Computing Surveys, 55(1), 1–37. https://doi.org/10.48550/arXiv.2106.15115

Rios, A., & de Carvalho, A. F. P. S. (2010). Parsing techniques in machine translation. *Machine Translation Journal, 23*(2), 128–142.

Rivera-Trigueros, I. Machine translation systems and quality assessment: a systematic review. *Lang Resources & Evaluation* **56**, 593–619 (2022). https://doi.org/10.1007/s10579-021-09537-5

Shterionov, D., Superbo, R., Nagle, P., Casanellas, L., O'Dowd, T., & Way, A. (2018). Human versus automatic quality evaluation of NMT and PBSMT. *Machine Translation*, *32*(3), 217–235. http://www.jstor.org/stable/44987870

Silvestre-Cerdà, J.A., Andrés-Ferrer, J., Civera, J. (2011). Explicit Length Modelling for Statistical Machine Translation. In: Vitrià, J., Sanches, J.M., Hernández, M. (eds) Pattern Recognition and Image Analysis. IbPRIA 2011. Lecture Notes in Computer Science, vol 6669. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-21257-4_34

Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. *In Proceedings of the 27th International Conference on Neural Information Processing Systems* (pp. 3103–3112). https://doi.org/10.48550/arXiv.1409.3215

Veselovská, L. (2017). *English syntax (1)*. Palacký University Olomouc.

Zhai, F., Zhang, J., Zhou, Y., & Zong, C. (2013). Handling ambiguities of bilingual predicate-argument structures for statistical machine translation. In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics 1127–1136. Association for Computational Linguistics.