Text Complexity Classification and Simplification

Ryan Wiesenberg

Abstract—**As businesses and organizations continue to grow globally interconnected, many groups require their internal documentation and presentations to be written in a manner easily understandable by non-native speakers. Due to this requirement, most team members will develop an ability to structure generated text in a simple manner. However, gaining this skill is not generally considered as part of the person's core work and requires a significant time to develop and master. Companies may seek to artificially augment their employees through automated sentence difficulty detection and simplification to alleviate this burden. To provide researchers with a comprehensive literature collection, this paper performs a literature review across existing research in sentence difficulty classification and simplification. It is the goal of this analysis to assist the research community in consolidating and discovering trends in existing research. Additionally, this paper presents potential paths for future developments based on the findings therein.**

I. INTRODUCTION

Text simplification (alternatively sentence or language simplification) is the process of reducing the complexity of a sentence or body of text, making it easier to understand, while maintaining the core meaning of the original text. This process has uses for machines, making it easier for other natural language processing (NLP) tasks. More commonly this task is performed to ease human ingestion for the benefit of non-native speakers, children, and persons with disabilities.

A. Historical Background and Approach

Historically the sentence simplification process involved simplification through some form of lexical or syntactical simplification. This process then generates a new text containing the same meaning. This process is not to be confused with the act of explanation generation, wherein uncommon or complex words are accompanied by a more simplistic term or phrase to provide additional resources for the more-common reader [1]. Artificial intelligence simplification agents will use some combination of lexical or syntactical simplification to determine both complexity and generate a less complex sentence, in a similar manner to that conducted by a human.

Lexical simplification is the process of removing or replacing complex words a simpler counterpart. This process creates a sentence or text with generally less or more common words while still maintaining the core meaning. When this task is performed by a human agent, it consists of two phases. First the agent will identify the complex words or phrases within a text. Then the agent will select the most appropriate replacement candidate from a generated list of words or phrases with an approximately similar meaning. Blum and Levenston. [2] further break down the generation of substitution candidates into various tasks such as synonym or superordinate substitution, paraphrasing, and meaning approximation.

Syntactic simplification is the process of breaking a complex text into its component parts and generating a subset of the text with a reduced grammatical complexity. Unlike lexical simplification which seeks to remove complicated words or phrases, syntactic simplification seeks to make the sentence easier to ingest by creating several shorter sentences of the same meaning. Additionally, this process removes inherent complexity in the sentence's grammar (passive tenses, sentence ordering, etc.). Syntactic simplification can make it easier for groups with disabilities or second-language learners to understand more complex text at a faster rate [3], [4].

B. Problem Description and Motivation

Because of the relative success of this process when performed by human agents, intelligence agents will similarly attempt lexical or syntactical simplification or some hybrid of the two. However, there is no built-in metric to language to aid this process of relative complexity of approximate meaning, making simplification a non-trivial task for human agents. Additionally, text simplification skills are rising in importance for people members of all professions due to the increasing globalization of businesses and organizations. Team members are requested to generate documentation in a manner that is easily understandable by non-native speakers or others less versed in the topic. However, this skill is generally not part of the person's core work or job description, and text simplification is a difficult task without significant training. Allowing this process to be automated or artificially augmented provides two potential benefits:

- Increasing efficiency by minimizing the cognitive burden of this auxiliary task
- Increasing effectiveness by suggesting or replacing complex text the author may not have self-detected due to lack of complete training

To allow for ease of model creation, the development of a comprehensive collection of text complexity classification and simplification research for a researcher to reference is thus seen as the first step towards future progress.

C. Research Approach

This paper conducts a literature review of existing methods of sentence complexity and simplification. This review will allow researchers to easily reference existing research on the topic and guide future work. Additionally, this paper will showcase regions of research that have not been fully explored and likely have gaps in their full understanding or potential use cases. As such, this paper seeks to achieve the following:

- Provide an overview of current resources used as references for varying levels of text complexity, methods used to classify sentence complexity, and architectures created for simplification tasks.

- Discover trends and gaps in previous research and further identify potential areas for future research

D. Related Surveys

Prior to this work, a few researchers have summarized the state of the field at the time of their publication. Feng (2008) [5] provides an analysis of several early works in text simplification for both machine ingestion [3], [6], [7] and human ingestion [8]–[15]. Feng gave an overview of several common simplification algorithms and presented a detailed analysis of each simplification method including both strengths and weaknesses. Shardlow [1] provided an summary of the field of text simplification in 2014. The paper collected various methods of lexical and syntactical simplification work from previous research work and briefly summarized them. Additionally, Shardlow provided a general trend and the challenges of the field, briefly touching on some emerging classification methods and simplification systems.

Since the publication of the aforementioned papers, there has a been a proliferation in corpora and more research into more automated simplification systems and additional identification of existing system shortcomings. This survey paper intends to collect and summarize corpus datasets, methods of complexity classification, and simplification architectures. This paper will prove an updated viewpoint of the field of text simplification and analyze popular trends and identify potential future research areas.

II. RESEARCH METHOD

To achieve the research goals of this paper, a systematic literature review was conducted following the guidelines provided by Kitchenham [16] and the structure of this paper was influenced by several other literature reviews [17], [18]. Much like the references provided, this paper structures the review into three phases: planning, conducting, and reporting. The current section will be used to outline the research protocol including the research questions, tasks for qualification, and literature search process. Each of the following sections, III, IV, and V, will address one of the research questions posed in the next segment. These sections will discover potential answers to the research question and report of the state of the field based on literature discovered by the search and inclusion process. Finally, this paper identifies potential future research directions and presents a conclusion of the findings.

A. Research Questions

The following research questions have been derived, based on the first research objective described in Section I.C. These questions will become the basis of the literature review and will be addressed in Sections III, IV, and V respectively.

- **RQ1. Text Corpora** What current text corpora are being used as a basis for artificial intelligence units in text complexity classification and simplification?

- **RQ2. Complexity Classification** What modules or architectures are currently used for determining the complexity level of a text?

- **RQ3. Simplification Architectures** What methods or architectures are currently used by researchers to create text simplification models?

B. Research Tasks

To answer the three research questions, several tasks have been conducted: one task to set up the literature review, a second task to classify each literature based on its methods and resources, and a third task to identify gaps and propose future research. These tasks were conducted sequentially and the cumulation of the work will be summarized at the conclusion of this paper.

The first task, setting up the literature review, includes the definition of the search and review process, the selection of search engines, definition of a keyword set, and a final selection of papers generated from the keyword list compared against a set of inclusion and exclusion criteria. This step is explained further in the following section *Literature* Search *Process*.

In the second research task, each paper collected in the previous setup task was classified according to which methods of each **RQ1**, **RQ2**, **RQ3** was used in the research. The aim of this task is to provide an overview of the field of text complexity and simplification. Prevailing trends and methodologies for each research question are presented along with potential strengths and weaknesses in Sections III.D, 0, and V respectively. Additionally, this task will create a full research space for the next and final research task.

In the final research task, each paper gathered and classified in the previous task is cross analyzed against the collected trends to discover areas of future research in Section VI. Additionally, a closing summary of the paper and conclusions about the state of the field will be presented in Section VII.

C. Literature Search Process

The search strategy for this review was directed at finding published research papers from conferences or journals. Google Scholar, IEEE Xplore, and ACM Digital Library were all used to find and collect the publications. TABLE I. Search Terms lists all the search terms used in the search for articles related to each research question and a second table applied to each research question to keep the results confined to text simplification research.

TABLE I.　　SEARCH TERMS

Research Question	Search Terms
Complexity Classification	measurement OR classification OR estimation
Text Corpora	corpora OR corpus OR dataset OR training data OR source data
Simplification Architectures	machine learning OR deep learning OR neural OR artificial OR artificial intelligence OR AI OR adaptation OR automatic OR architecture

AND

	Search Terms
Text Simplification	text simplification OR sentence simplification OR language simplification OR lexical simplification OR syntactic simplification OR text complexity OR sentence complexity

Due to the broad nature of the research questions the inclusion criteria was equally broad for each literature item:

1. The item must present or use at least one of: complexity classified text corpora, automated complexity classifications, or simplification architecture.
2. The item must present some form of critique or cost-benefit analysis of the identified method to be considered related to that method.
3. The method presented must include some form of automation if dealing with complexity classification or simplification architectures

Any that did not contain the two items above is excluded from the systematic literature review that follows. Additionally, articles or methods were excluded if they:

1. modernized or updated but did not add additional layers or methods to an existing classifier, corpus, or learning architecture
2. were hidden behind a paywall inaccessible by university or personal resources
3. did not have a version of the text available in English

D. Threats to Validity

The primary thread to the validity of this systematic literature review is incompleteness of literature available. This issue stems from the potentially limited nature of the search engines and associated search terms. To minimize this threat, this paper uses an iterative search approach, adding search terms as they are discovered through the reading of literature discovered by the initial search criteria. Additionally, multiple search engines were used to mitigate this risk.

III. TEXT CORPORA

To define text complexity and simplification tasks and success criteria, it is important to generate sufficient text corpora by which to judge the models. In some of the early works on text simplification, these models were created by hand to suit aphasic needs [11]–[15] by subject-matter experts and linguistic professionals. The corpora in these studies was limited to a small selection, just sufficient to validate the hand-applied rule-based simplification model. For automated text complexity classification and simplification models a larger corpora network is needed to ensure applicability across domains and for sufficient training data for more complex models such as neural networks.

A. News Corpora

News sources are commonly used for text complexity estimation tasks because multiple versions of the same article can be found on the same topic. This allows for systems to learn how to determine similarity across different writing styles and complexity levels. Additionally, due to the diverse topic nature of news articles, models trained on news documentation is widely applicable to a variety of domains. In 2016, Xu et al. [19] produced the Newsela corpus, a set of news texts, translated into a variety of simplification levels and now commonly used by other text complexity research [19], [20], [21]–[25].

B. Encyclopedia Corpora

Another common source of corpus for simplification tasks is from encyclopedia-like sources. Its frequent use in complexity calculation and simplification tasks stems from the original source, Britannica [26]–[28] and Wikipedia [19], [23]–[25], [27]-[29], [30]–[43], having a similar resource with the same content but written for an audience of a lower reading level to increase accessibility.

Wikipedia in particular has received a lot of attention in the NLP and text simplification communities due to the large efforts to align and create parallel corpuses. The two most used corpora are through WikiSmall [29], Turk Corpus [19], and Coster and Kauchak Corpus [38], all of which contain both 1-to-1 and 1-to-many simplification alignments. Hwang [44], and Kaji and Komachi [45] are two lesser-used corpora, only having 1-to-1 simplification alignments, However, it is worth noting that these simplifications are only rough approximations of an equivalent simplification. As such, these alignments are not ideal because they do not correspond to a minimum simplification, only a less complex version of a sentence with a similar meaning.

C. Reading-Level Corpora

Perhaps the most obvious corpora for text simplification is based on approximate human reading-level. This classification of corpora is created as a result of the need for appropriate per educational level. Each text is annotated either with a grade level in the case of Common Core [34], [46]–[49], or by indented audience as with Lexile score [46], [47], [49]. The drawback to this dataset is that it cannot be readily used for simplification tasks as there is no obvious parallel corpora like is available the previous two sections, but it does provide a good reference for complexity classification tasks.

D. Scientific Corpora

The scientific field has also been actively investigating the use of text simplification. Due to the large gap between complexity of doctoral text and the typical reader's understanding of the field, this research primarily focuses on the generation of lay-person friendly documentation [50]–[52]. As a result, several researchers have created parallel simplified text corpora for a variety of scientific fields, especially medical texts. However, due to the niche nature of the text and wording within the text, each corpus is of relatively limited scope. The automated classification and simplification similarly methods associated possible with the corpora with them are similarly disadvantaged.

IV. COMPLEXITY CLASSIFICATION

The earliest form of text complexity estimation involves the use of a Likert scale. Based on its perceived complexity to the reader, the participants are tasked rating the text on a relative, finite scale typically from 1 to 5 when compared of other texts. The results of this test would then be aggregated, and the average score would then be used to compare the text against other texts rated by the same or similar group. In practice this measurement has been found to correspond well to approximate reading level [53]. However, this task is known to be time and

resource intensive in both sourcing sufficient participation and developing a comprehensive survey model. As a result, researchers have developed several automated methods of estimating reading level through simple complexity metrics. These metrics are sourced or roughly correspond to different lexical and syntactic elements such sentence length, number of syllables, and tense usage.

A. Relative Complexity Level

One of the earliest forms of automated reading level complexity estimation was developed by Flesch [54] as the Flesch reading-ease test and then further refined with assistance by Kincaid [55] into the Flesh-Kincaid reading grade level for the US navy. These methods are simple and only consider the length of the sentences and syllables within those words. As such, it is easy to score well on the test by using sentences with minimal words and short syllables. When this is paired with automated simplification tools, the resultant simplified sentences can score high while not maintaining grammar or the original meaning [36]. While it is common for more classical simplification papers to use the Flesch-Kincaid score as the primary measurement [33], [34], [39], [41], [56], it is more modernly used alongside another metric that does not have the same disadvantages [23], [24], [29], [36], [47], [48].

One method of augmenting the Flesch-Kincaid score is through the incorporation of a core vocabulary list. This core vocabulary list is derived from the analysis of literature of a given language and identifying how often words are used in writings across the language. The resulting metric creates a stronger correlation between language competency of the writer and of the text where passages with less common words receive a higher complexity score, typically referred to as a Lexile classification due to its similarity to the educational tool of the same name. However, much like the Flesch-Kincaid score, this metric fails to address the issue of lost meaning or poor sentence structure and grammar. This type of classification is still used in some sentence complexity research [9], [46], [47], [57], [58] but is more commonly used to classify the complexity of reading-level corpora as described in Section III.B.

B. Corpora-Tuned Complexity

The most common method of incorporating the cohesion of a sentence in text-complexity calculation is to include an n-gram analysis. Developing an n-gram precision model involves the creation of continuous sequences of items, in this instance words, that frequently occur together, as in a sentence, across a given training data. These n-grams typically correlate well to human constructed sentences in both meaning and grammar because of their essentially human-generated training. One of the most common n-gram models is BLEU, created by Papineni et al. in 2012 [59]. Many studies have been conducted with the use of BLEU as a primary simplification metric [19], [22], [22], [24], [25], [27], [29], [31], [33], [35]–[38], [40], [42], [43], [58], [60]–[63]. However, when simple n-gram classifiers like BLEU are used sentence simplification tasks, give high value to words that are commonly adjacent, even if the actual meaning of the sentence is not maintained [35].

SARI, developed alongside the Turk Wikipedia Corpora by Xu et al. [19], seeks to compensate for BLEU's shortcomings. SARI ingests known simplicity conversions for the given corpora during its calculation to force a simplification model to maintain meaning as well as decrease complexity. As shown in its introductory paper, SARI shows similar performance to BLEU in cohesion with the added benefit of additionally analyzing valid simplifications. However, many studies have still found some discrepancies between high-scoring sentences and human readability. Despite the potential shortcomings, This model of grouping n-gram generation with simplicity parallels is the main metric of analyzing sentence complexity and simplification success along with BLEU [19], [22]–[25], [31], [35], [37], [40], [43], [58].

In recent years, some research has been performed in determining text complexity through a neural network. These utilize Recurrent Neural Networks and their ability to cumulatively build up a model of meaning over a body of text and then distill the information down to a single metric [64], [65]. Additionally, Lai. [65] proposed the use of both convolution and recursion in the neural network to create an RNN/CNN hybrid that outperforms either system independently, but more research is needed to discover the shortcomings of either approach.

V. Simplification Architectures

Text simplification not only improves the learning potential of lay people in a skilled field, but also allows non-native speakers or disabled persons to more easily communicate with others. Historically, simplification of text was a largely manual process, requiring a highly trained or experienced individual. To help ease the skilled, manual effort of simplification, researchers developed several varieties of automated architectures for text simplification.

A. Rule-based Approaches

Rule-based simplification, also referred to as tree-based simplification, is the earliest automated simplification process. This method was popularized by the PSET model created by Carrol [15] and improved by Canning et al. [12], [13], [13], [66] in subsequent years. This method continues to be popular [3], [7], [7], [9], [10], [39], [50]–[53], [67]–[71] because of its high performance, adaptability to any domain, and transparency of creation when compared to Statistical Machine Translation and Neural Network simplification as discussed in the following two sections. This method's main drawback is its highly intensive creation process and tuning requiring extensive rules for any potential lexical or syntactical simplification requirement spanning any part of speech, tense, sentence ordering, and more. Additionally, this model is limited in application to the text language and subject matter that it was intentionally designed for.

A more modern approach to this task involves a list of known word substitutions in conjunction with an automated complexity algorithm such as Flesch-Kincaid, BLEU, or SARI. This approach minimize the complexity metric or metrics while cycling through potential substitutions. These word substitutions most commonly come from the paraphrase

database (PPDB) [19], [23], [31], a global pairing of source text with others of similar meaning, or WordNet [39], a collection of semantic relationships between words. These databases can be used when attempting to simplify text that does not readily have a parallel corpus or a corpus of known complexity. Much like the purely rule-based approach above, this method requires a non-insignificant amount of tuning to avoid incorrect substitutions. In contrast, this architecture is easier to construct than a direct rule-based approach because PPDB and WordNet cover almost any topic and a wide array of languages.

B. Statistical Machine Translation (SMT)

Statistical machine translation is the process of determining the relationship of words according to the frequency of the presence of the term in parallel corpus. Typically, this approach is applied to bi-lingual texts, but has been used in simplification tasks by treating it as a mono-lingual translation from a complex language to a simpler language [6], [27], [41], [63], [67], [72]. Unlike rule-based translation, there is no need to tailor the simplification model toward a specific language or topic, requiring less manual overhead in model creation. The tradeoff for the ease of simplification model creation is the requirement of significant parallel corpus for developing statistical significance. Additionally, there is a large memory overhead to track and remember words or phrases with a high correlation, especially across diverse corpora.

C. Deep Learning

The most recent development in the area of text simplification is the use of neural networks as a simplification architecture. The most common deep learning simplification model involves tuning a neural machine translation (NMT) model like seq2seq, which encodes text intent into a summary vector and then decodes it into a simpler form, on parallel corpus [22]–[25], [31], [43]. This method operates much the same way as the SMT architecture, treating simplification as a bi-lingual translation task from a complex language to a simple one. Both methods require significant parallel corpus and similarly improve over the rule-based model in required human tuning. Additionally, NMT models tend to have a smaller footprint over SMT models because they do not have to remember every statistically significant correlation, instead remembering how to encode or decode simplistic vectors. However, these models tend to produce less human-like results than rule-based or SMT approaches due to the encoding process which can potentially lose some meaning.

The newest form of a deep learning simplification model was published in 2019 by Kawashima and Takagi [37]. Their paper utilized a generative adversarial network (GAN) to created simplified sentences from the Simple Wikipedia corpora. The advantages of a GAN over typical NMT models is the ability to use non-parallel corpus as well as continuously improve the human-like structure of the resultant sentences by tuning a discriminator alongside the simplification model. However, this is the first research of its type and does not provide much study into the potential shortcomings of the approach.

VI. RECOMMENDATIONS FOR FUTURE RESEARCH

Based on the results of the literature review conducted in the previous three sections, one can see the proliferation of research in sentence complexity classification and simplification. The major pain point across text complexity classification and simplification continues to be the requirement of significant manual work.

The creation of automatically aligned parallel corpus is an area of future research. It is possible to align two sets of corpus that are known to have parallels without direct human intervention as Coster and Kauchack have shown with the Simple Wikipedia corpus [38]. Automatic alignment of two corpora could also similarly be applied to the Britannica dataset or news corpus to create a wider array of available simplification datasets.

There is still a need for a complexity metric that perfectly reflects the difficulty of a sentence while maintaining meaning and structure. Bosco et al. [64] and Lai et al. [65] have proposed methods of classification using a neural network and the results are promising, but more research is needed to compare these methods to human classifications. Because this application is new and relatively unexplored, more research could be conducted to develop additional network models and tune the architecture to get a more accurate estimation of complexity.

Another potential area of research is in the application of GANs to different corpora, similar to the research performed by Kawashima and Takagi [37]. As was mentioned in the section on Wikipedia (Section III.B), the parallelism of Wikipedia to Simple Wikipedia is not an optimal simplification, just an estimation. Additionally, it is not important to have a perfectly parallel corpora for a GAN, just one with known relative complexity values such as Common Core or Lexical corpora. As subsequent layer, the GAN ecosystem could implement a neural classifier to determine the lexical level of the minimized text instead of the BLEU and SARI method performed in the paper.

VII. CONCLUSIONS

This paper presented a systematic literature review on the developments on automated text complexity classification and simplification models and architectures. It is the intent of this review to identify trends in the current state of the field and identify ongoing issues of the practice. It is the hope of the author that this paper can be used as a reference to discover leading literature and methodologies in the field including the implementations of various corpora, complexity calculations, and text simplification models.

Despite many advances over the previous 25 years of sentence complexity classification and simplification, the primary difficulty in the field is still the need for hand-tuning by researchers and subject matter experts. This paper outlines some recommended future research involving heavier utilization of neural networks for corpora generation, text complexity calculation, and text simplification architectures as described in section VI. Given more time and attention, these methods could further minimize the need for human effort and improve automated simplification tasks.

REFERENCES

[1] M. Shardlow, "A Survey of Automated Text Simplification," *Int. J. Adv. Comput. Sci. Appl.*, vol. 4, no. 1, 2014, doi: 10.14569/SpecialIssue.2014.040109.

[2] S. Blum and E. A. Levenston, "Universals of Lexical Simplification," *Lang. Learn.*, vol. 28, no. 2, pp. 399–415, 1978, doi: 10.1111/j.1467-1770.1978.tb00143.x.

[3] A. Siddharthan, "Syntactic Simplification and Text Cohesion," *Res. Lang. Comput.*, vol. 4, no. 1, pp. 77–109, Jun. 2006, doi: 10.1007/s11168-006-9011-1.

[4] F. Kiliçkaya, *Authentic Materials and Cultural Content in EFL Classrooms*, vol. 10, no. 7. 2004.

[5] L. Feng, "Text Simplification: A Survey," p. 35.

[6] B. Beigman Klebanov, K. Knight, and D. Marcu, "Text Simplification for Information-Seeking Applications," in *On the Move to Meaningful Internet Systems 2004: CoopIS, DOA, and ODBASE*, Berlin, Heidelberg, 2004, pp. 735–747, doi: 10.1007/978-3-540-30468-5_47.

[7] R. Chandrasekar, C. Doran, and B. Srinivas, "Motivations and methods for text simplification," in *Proceedings of the 16th conference on Computational linguistics - Volume 2*, Copenhagen, Denmark, Aug. 1996, pp. 1041–1044, doi: 10.3115/993268.993361.

[8] J. Burstein, J. Shore, J. Sabatini, Y.-W. Lee, and M. Ventura, "The Automated Text Adaptation Tool," in *Proceedings of Human Language Technologies: The Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT)*, Rochester, New York, USA, Apr. 2007, pp. 3–4, Accessed: Jun. 16, 2020. [Online]. Available: https://www.aclweb.org/anthology/N07-4002.

[9] S. Williams and E. Reiter, "Generating readable texts for readers with low basic skills," presented at the Proceeding of the 10th European Workshop on Natural Language Generation, Aberdeen, UK, 2005, Accessed: Jun. 16, 2020. [Online]. Available: http://www.ling.helsinki.fi/~gwilcock/ENLG-05/.

[10] K. Inui, A. Fujita, T. Takahashi, R. Iida, and T. Iwakura, "Text simplification for reading assistance: a project note," in *Proceedings of the second international workshop on Paraphrasing - Volume 16*, Sapporo, Japan, Jul. 2003, pp. 9–16, doi: 10.3115/1118984.1118986.

[11] S. L. Devlin, "Simplifying natural language for aphasic readers.," Ph.D., University of Sunderland, 1999.

[12] S. Devlin and G. Unthank, "Helping aphasic people process online information," in *Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility*, Portland, Oregon, USA, Oct. 2006, pp. 225–226, doi: 10.1145/1168987.1169027.

[13] Y. M. Canning, "Syntactic simplification of text," Ph.D., University of Sunderland, 2002.

[14] Y. Canning and J. Tait, "Syntactic Simplification of Newspaper Text for Aphasic Readers," *ACM SIGIR Forum*, vol. 33, no. 1, pp. 6–11, Sep. 1999, doi: 10.1145/331403.331410.

[15] J. Carroll, G. Minnen, Y. Canning, S. Devlin, and J. Tait, "Practical Simplification of English Newspaper Text to Assist Aphasic Readers," p. 5, Jul. 1998.

[16] B. Kitchenham, "Procedures for Performing Systematic Reviews," p. 33.

[17] M. Harman, S. A. Mansouri, and Y. Zhang, "Search-based software engineering: Trends, techniques and applications," *ACM Comput. Surv.*, vol. 45, no. 1, p. 11:1–11:61, Dec. 2012, doi: 10.1145/2379776.2379787.

[18] A. Aleti, B. Buhnova, L. Grunske, A. Koziolek, and I. Meedeniya, "Software Architecture Optimization Methods: A Systematic Literature Review," *IEEE Trans. Softw. Eng.*, vol. 39, no. 5, pp. 658–683, May 2013, doi: 10.1109/TSE.2012.64.

[19] W. Xu, C. Napoles, E. Pavlick, Q. Chen, and C. Callison-Burch, "Optimizing Statistical Machine Translation for Text Simplification," *Trans. Assoc. Comput. Linguist.*, vol. 4, pp. 401–415, Dec. 2016, doi: 10.1162/tacl_a_00107.

[20] R. Stodden and L. Kallmeyer, "A multi-lingual and cross-domain analysis of features for text simplification," May 2020, pp. 77–84, Accessed: Jun. 17, 2020. [Online]. Available: https://www.aclweb.org/anthology/2020.readi-1.12.

[21] S. Stajner, S. P. Ponzetto, and H. Stuckenschmidt, "Automatic Assessment of Absolute Sentence Complexity," in *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, Melbourne, Australia, Aug. 2017, pp. 4096–4102, doi: 10.24963/ijcai.2017/572.

[22] R. Kriz *et al.*, "Complexity-Weighted Loss and Diverse Reranking for Sentence Simplification," *ArXiv190402767 Cs*, Apr. 2019, Accessed: Jun. 18, 2020. [Online]. Available: http://arxiv.org/abs/1904.02767.

[23] S. Zhao, R. Meng, D. He, S. Andi, and P. Bambang, "Integrating Transformer and Paraphrase Rules for Sentence Simplification," *ArXiv181011193 Cs*, Oct. 2018, Accessed: Jun. 18, 2020. [Online]. Available: http://arxiv.org/abs/1810.11193.

[24] X. Zhang and M. Lapata, "Sentence Simplification with Deep Reinforcement Learning," in *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, Copenhagen, Denmark, Sep. 2017, pp. 584–594, doi: 10.18653/v1/D17-1062.

[25] T. Vu, B. Hu, T. Munkhdalai, and H. Yu, "Sentence Simplification with Memory-Augmented Neural Networks," *ArXiv180407445 Cs*, Apr. 2018, Accessed: Jun. 18, 2020. [Online]. Available: http://arxiv.org/abs/1804.07445.

[26] R. Barzilay and N. Elhadad, "Sentence alignment for monolingual comparable corpora," in *Proceedings of the 2003 conference on Empirical methods in natural language processing*, USA, Jul. 2003, pp. 25–32, doi: 10.3115/1119355.1119359.

[27] S. Štajner, H. Béchara, and H. Saggion, "A Deeper Exploration of the Standard PB-SMT Approach to Text Simplification and its Evaluation," in *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, Beijing, China, Jul. 2015, pp. 823–828, doi: 10.3115/v1/P15-2135.

[28] J. J. Li and A. Nenkova, "Fast and Accurate Prediction of Sentence Specificity," presented at the Twenty-Ninth AAAI Conference on Artificial Intelligence, Feb. 2015, Accessed: Jun. 18, 2020. [Online]. Available: https://www.aaai.org/ocs/index.php/AAAI/AAAI15/paper/view/9941.

[29] Z. Zhu, D. Bernhard, and I. Gurevych, "A monolingual tree-based translation model for sentence simplification," in *Proceedings of the 23rd International Conference on Computational Linguistics*, Beijing, China, Aug. 2010, pp. 1353–1361, Accessed: Jun. 17, 2020. [Online].

[30] S. Vajjala and D. Meurers, "Assessing the relative reading level of sentence pairs for text simplification," in *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*, Gothenburg, Sweden, Apr. 2014, pp. 288–297, doi: 10.3115/v1/E14-1031.

[31] S. Nisioi, S. Štajner, S. P. Ponzetto, and L. P. Dinu, "Exploring Neural Text Simplification Models," in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, Vancouver, Canada, Jul. 2017, pp. 85–91, doi: 10.18653/v1/P17-2014.

[32] W. Coster and D. Kauchak, "Learning to simplify sentences using Wikipedia," in *Proceedings of the Workshop on Monolingual Text-To-Text Generation*, Portland, Oregon, Jun. 2011, pp. 1–9, Accessed: Jun. 17, 2020. [Online].

[33] K. Woodsend and M. Lapata, "Learning to simplify sentences with quasi-synchronous grammar and integer programming," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Edinburgh, United Kingdom, Jul. 2011, pp. 409–420, Accessed: Jun. 17, 2020. [Online].

[34] S. Vajjala and D. Meurers, "Readability-based Sentence Ranking for Evaluating Text Simplification," *ArXiv160306009 Cs*, Mar. 2016, Accessed: Jun. 18, 2020. [Online]. Available: http://arxiv.org/abs/1603.06009.

[35] E. Sulem, O. Abend, and A. Rappoport, "Semantic Structural Evaluation for Text Simplification," in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, New Orleans, Louisiana, Jun. 2018, pp. 685–696, doi: 10.18653/v1/N18-1063.

[36] S. Wubben, A. van den Bosch, and E. Krahmer, "Sentence Simplification by Monolingual Machine Translation," in *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Jeju Island, Korea, Jul. 2012, pp. 1015–1024, Accessed: Jun. 15, 2020. [Online]. Available: https://www.aclweb.org/anthology/P12-1107.

[37] T. Kawashima and T. Takagi, "Sentence Simplification from Non-Parallel Corpus with Adversarial Learning," in *IEEE/WIC/ACM International Conference on Web Intelligence*, Thessaloniki, Greece, Oct. 2019, pp. 43–50, doi: 10.1145/3350546.3352499.

[38] W. Coster and D. Kauchak, "Simple English Wikipedia: a new text simplification task," in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers - Volume 2*, Portland, Oregon, Jun. 2011, pp. 665–669, Accessed: Jun. 17, 2020. [Online].

[39] J. De Belder and M.-F. Moens, "Text simplification for children," in *Prroceedings of the SIGIR workshop on accessible search systems*, Jan. 2010, pp. 19–26, Accessed: Jun. 18, 2020. [Online]. Available: https://lirias.kuleuven.be/1655494.

[40] C. Scarton, G. Paetzold, and L. Specia, "Text Simplification from Professionally Produced Corpora," presented at the LREC 2018, Miyazaki, Japan, May 2018, Accessed: Jun. 18, 2020. [Online]. Available: https://www.aclweb.org/anthology/L18-1553.

[41] N. Bach, Q. Gao, S. Vogel, and A. Waibel, "TriS: A Statistical Sentence Simplifier with Log-linear Models and Margin-based Discriminative Training," in *Proceedings of 5th International Joint Conference on Natural Language Processing*, Chiang Mai, Thailand, Nov. 2011, pp. 474–482, Accessed: Jun. 18, 2020. [Online]. Available: https://www.aclweb.org/anthology/I11-1053.

[42] S. Narayan and C. Gardent, "Unsupervised Sentence Simplification Using Deep Semantics," *ArXiv150708452 Cs*, Sep. 2016, Accessed: Jun. 18, 2020. [Online]. Available: http://arxiv.org/abs/1507.08452.

[43] A. Wang, R. Kurohara, and D. Liang, "Wikipedia Sentence Simplification," p. 11.

[44] W. Hwang, H. Hajishirzi, M. Ostendorf, and W. Wu, "Aligning Sentences from Standard Wikipedia to Simple Wikipedia," in *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Denver, Colorado, May 2015, pp. 211–217, doi: 10.3115/v1/N15-1022.

[45] T. Kajiwara and M. Komachi, "Building a Monolingual Parallel Corpus for Text Simplification Using Sentence Similarity Based on Alignment between Word Embeddings," in *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, Osaka, Japan, Dec. 2016, pp. 1147–1158, Accessed: Jun. 19, 2020. [Online]. Available: https://www.aclweb.org/anthology/C16-1109.

[46] M. Glaus, "Text Complexity and Young Adult Literature," *J. Adolesc. Adult Lit.*, vol. 57, no. 5, pp. 407–416, 2014, doi: 10.1002/jaal.255.

[47] K. M. Sheehan, M. Flor, and D. Napolitano, "A Two-Stage Approach for Generating Unbiased Estimates of Text

Complexity," in *Proceedings of the Workshop on Natural Language Processing for Improving Textual Accessibility*, Atlanta, Georgia, Jun. 2013, pp. 49–58, Accessed: Jun. 18, 2020. [Online]. Available: https://www.aclweb.org/anthology/W13-1506.

[48] K. M. Sheehan, "Aligning TextEvaluator® Scores With the Accelerated Text Complexity Guidelines Specified in the Common Core State Standards," *ETS Res. Rep. Ser.*, vol. 2015, no. 2, pp. 1–20, 2015, doi: 10.1002/ets2.12068.

[49] K. M. Sheehan, I. Kostin, Y. Futagi, and M. Flor, "Generating Automated Text Complexity Classifications That Are Aligned with Targeted Text Complexity Standards," *ETS Res. Rep. Ser.*, vol. 2010, no. 2, pp. i–44, 2010, doi: 10.1002/j.2333-8504.2010.tb02235.x.

[50] S. Jonnalagadda and G. Gonzalez, "BioSimplify: an open source sentence simplification engine to improve recall in automatic biomedical information extraction," *AMIA. Annu. Symp. Proc.*, vol. 2010, pp. 351–355, 2010.

[51] M. Kvist and S. Velupillai, "Professional Language in Swedish Radiology Reports – Characterization for Patient-Adapted Text Simplification," 2013, pp. 55–59, Accessed: Jun. 17, 2020. [Online]. Available: http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-95622.

[52] I. Temnikova, "TEXT COMPLEXITY AND TEXT SIMPLIFICATION," Apr. 2012, Accessed: Jun. 17, 2020. [Online]. Available: https://wlv.openrepository.com/handle/2436/297482.

[53] W. S. Lasecki, L. Rello, and J. P. Bigham, "Measuring text simplification with the crowd," in *Proceedings of the 12th Web for All Conference*, Florence, Italy, May 2015, pp. 1–9, doi: 10.1145/2745555.2746658.

[54] R. Flesch, "A new readability yardstick," *J. Appl. Psychol.*, vol. 32, no. 3, pp. 221–233, 1948, doi: 10.1037/h0057532.

[55] J. P. Kincaid, Jr. Fishburne, R. Robert P., C. Richard L., and Brad S., "Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel:," Defense Technical Information Center, Fort Belvoir, VA, Feb. 1975. doi: 10.21236/ADA006655.

[56] N. Karpov, J. Baranova, and F. Vitugin, "Single-Sentence Readability Prediction in Russian," in *Analysis of Images, Social Networks and Texts*, Cham, 2014, pp. 91–100, doi: 10.1007/978-3-319-12580-0_9.

[57] M. Z. Kurdi, "Lexical and Syntactic Features Selection for an Adaptive Reading Recommendation System Based on Text Complexity," in *Proceedings of the 2017 International Conference on Information System and Data Mining*, Charleston, SC, USA, Apr. 2017, pp. 66–69, doi: 10.1145/3077584.3077595.

[58] T. Maruyama and K. Yamamoto, "Sentence simplification with core vocabulary," in *2017 International Conference on Asian Language Processing (IALP)*, Dec. 2017, pp. 363–366, doi: 10.1109/IALP.2017.8300618.

[59] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, "Bleu: a Method for Automatic Evaluation of Machine Translation," in *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, Philadelphia, Pennsylvania, USA, Jul. 2002, pp. 311–318, doi: 10.3115/1073083.1073135.

[60] A. Katsuta and K. Yamamoto, "Crowdsourced Corpus of Sentence Simplification with Core Vocabulary," 2018.

[61] D. Ștefănescu, R. Ion, and S. Hunsicker, "Hybrid Parallel Sentence Mining from Comparable Corpora," p. 8.

[62] D. Feblowitz and D. Kauchak, "Sentence Simplification as Tree Transduction," in *Proceedings of the Second Workshop on Predicting and Improving Text Readability for Target Reader Populations*, Sofia, Bulgaria, Aug. 2013, pp. 1–10, Accessed: Jun. 18, 2020. [Online]. Available: https://www.aclweb.org/anthology/W13-2901.

[63] L. Specia, "Translating from Complex to Simplified Sentences," in *Computational Processing of the Portuguese Language*, Berlin, Heidelberg, 2010, pp. 30–39, doi: 10.1007/978-3-642-12320-7_5.

[64] G. L. Bosco, G. Pilato, and D. Schicchi, "A Sentence based System for Measuring Syntax Complexity using a Recurrent Deep Neural Network," p. 8, Nov. 2018.

[65] S. Lai, L. Xu, K. Liu, and J. Zhao, "Recurrent Convolutional Neural Networks for Text Classification," presented at the Twenty-Ninth AAAI Conference on Artificial Intelligence, Feb. 2015, Accessed: Jun. 18, 2020. [Online]. Available: https://www.aaai.org/ocs/index.php/AAAI/AAAI15/paper/view/9745.

[66] S. L. Devlin, "Simplifying natural language for aphasic readers.," Ph.D., University of Sunderland, 1999.

[67] S. Bott, H. Saggion, and D. Figueroa, "A hybrid system for Spanish text simplification," in *Proceedings of the Third Workshop on Speech and Language Processing for Assistive Technologies*, Montreal, Canada, Jun. 2012, pp. 75–84, Accessed: Jun. 17, 2020. [Online].

[68] D. Vickrey and D. Koller, "Applying sentence simplification to the CoNLL-2008 shared task," in *Proceedings of the Twelfth Conference on Computational Natural Language Learning*, Manchester, United Kingdom, Aug. 2008, pp. 268–272, Accessed: Jun. 17, 2020. [Online].

[69] E. Abrahamsson, T. Forni, M. Skeppstedt, and M. Kvist, "Medical text simplification using synonym replacement: Adapting assessment of word difficulty to a compounding language," in *Proceedings of the 3rd Workshop on Predicting and Improving Text Readability for Target Reader Populations (PITR)*, Gothenburg, Sweden, Apr. 2014, pp. 57–65, doi: 10.3115/v1/W14-1207.

[70] A. Candido, E. Maziero, L. Specia, C. Gasperin, T. Pardo, and S. Aluisio, "Supporting the Adaptation of Texts for Poor Literacy Readers: a Text Simplification Editor for Brazilian Portuguese," in *Proceedings of the Fourth Workshop on Innovative Use of NLP for Building Educational Applications*, Boulder, Colorado, Jun. 2009, pp. 34–42, Accessed: Jun. 17, 2020. [Online]. Available: https://www.aclweb.org/anthology/W09-2105.

[71] G. Leroy, J. E. Endicott, D. Kauchak, O. Mouradi, and M. Just, "User Evaluation of the Effects of a Text Simplification Algorithm Using Term Familiarity on

Perception, Understanding, Learning, and Information Retention," *J. Med. Internet Res.*, vol. 15, no. 7, p. e144, 2013, doi: 10.2196/jmir.2569.

[72] C. Quirk, C. Brockett, and W. Dolan, "Monolingual Machine Translation for Paraphrase Generation," in *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, Barcelona, Spain, Jul. 2004, pp. 142–149, Accessed: Jun. 17, 2020. [Online]. Available: https://www.aclweb.org/anthology/W04-3219.