

## Article

# Korean Semantic Role Labeling with Bidirectional Encoder Representations from Transformers and Simple Semantic Information

Jangseong Bae <sup>1</sup>  and Changki Lee <sup>2,\*</sup>
<sup>1</sup> Language AI Lab, LG CNS, Seoul 07795, Korea; bjs@lgcns.com

<sup>2</sup> Department of Computer Science and Engineering, Kangwon National University, Chuncheon 24341, Korea

\* Correspondence: leeck@kangwon.ac.kr

**Abstract:** State-of-the-art semantic role labeling (SRL) performance has been achieved using neural network models by incorporating syntactic feature information such as dependency trees. In recent years, breakthroughs achieved using end-to-end neural network models have resulted in a state-of-the-art SRL performance even without syntactic features. With the advent of a language model called bidirectional encoder representations from transformers (BERT), another breakthrough was witnessed. Even though the semantic information of each word constituting a sentence is important in determining the meaning of a word, previous studies regarding the end-to-end neural network method did not utilize semantic information. In this study, we propose a BERT-based SRL model that uses simple semantic information without syntactic feature information. To obtain the latter, we used PropBank, which described the relational information between predicates and arguments. In addition, text-originated feature information obtained from the training text data was utilized. Our proposed model achieved state-of-the-art results on both Korean PropBank and CoNLL-2009 English benchmarks.

**Keywords:** Korean semantic role labeling; BERT; semantic information; text-originated feature information



**Citation:** Bae, J.; Lee, C. Korean Semantic Role Labeling with Bidirectional Encoder Representations from Transformers and Simple Semantic Information. *Appl. Sci.* **2022**, *12*, 5995. <https://doi.org/10.3390/app12125995>

Academic Editors: Jae-Hoon Kim and Kichun Lee

Received: 20 May 2022

Accepted: 10 June 2022

Published: 13 June 2022

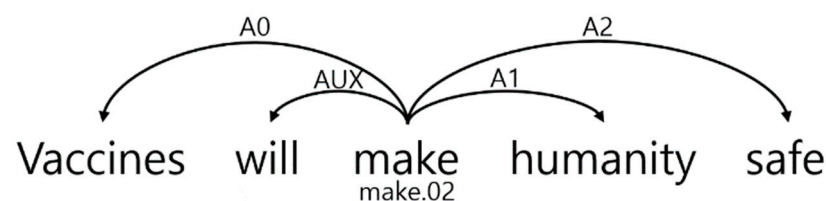
**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Semantic role labeling (SRL) is one of the fundamental tasks in natural language understanding (NLU). It aims to understand the predicate–argument structure of a sentence, to identify “who did what to whom,” “when,” “where,” etc. Given a sentence (in Figure 1), SRL has several subtasks in a sequential framework, including: predicate identification (make), predicate classification (make.02), argument identification (e.g., Vaccines), and argument classification. (Vaccines is A0 for the predicate make). This task finds its use in several downstream tasks, such as dialog systems [1], answering questions [2,3], and extracting information [4,5].



**Figure 1.** Example of a dependency-based SRL.

In recent years, breakthroughs achieved using end-to-end neural network models [6–9] have resulted in a state-of-the-art SRL performance even without syntactic features. These performances appear to falsify the long-held belief that syntactic features are fundamental

prerequisites for performing SRL [10,11]. However, state-of-the-art neural network models [12–15] can produce results with even better accuracy by leveraging syntactic feature information. Roth et al. [12] discussed the requirement of syntactic capabilities to achieve competitive performance in dependency-based SRL. Strubell et al. [13] showed that linguistic information, including syntactic feature information, can help improve SRL performance only when used adequately and efficiently on the SRL system during the training phase. Even in a previous study [7], it was shown that the use of gold syntax feature information could dramatically improve the SRL performance with no explicit linguistic features.

Recently, the natural language processing (NLP) community has shown interest in neural models that heavily use pre-training based on language modeling [16–18]. Pre-training based on language models has shown impressive benefits in a variety of NLP tasks, from sentence classification to sequence labeling. Among these, the most recently developed bidirectional encoder representations from transformers (BERT) [18] was trained using the encoder part of the transformer model [19]; furthermore, it outperformed the other two language models. Most recently, in addition to the previously mentioned language models, there are language models that have trained more extensive parameters, such as XLNet [20] and GPT-3 [21]. However, many studies are still being conducted based on the BERT model because these models require very large computing power, even in inference time. Apart from the interest in the NLP community, there have been studies using the BERT model for semantic role labeling [22]. Shi et al. [22] showed that the pre-trained BERT model performed sufficiently well in SRL tasks without any linguistic information.

The semantic information of every word of the sentence is highly important in determining the meaning-role of the word, but to utilize any information from those words requires WordNet or a comparable thesaurus, thus limiting its practical usage. Therefore, in previous studies of semantic role labeling using machine learning, semantic information was excluded, and the study was primarily conducted using syntax analysis qualities or end-to-end neural network methods.

Some common queries include: (1) Does BERT show good performance even when the model is applied to Korean, which is a low-resource language? (2) Can semantic information be obtained in a straightforward way? Can SRL be applied to the BERT-based semantic role-labeling model? (3) Will adding simple semantic information and text-originated feature information improve SRL performance? The answer to all such questions is positive. We propose a BERT-based Korean semantic role labeling system that utilizes two types of feature information—one that requires the extraction of the arguments that may be contained in a predicate from a corpus such as PropBank [23], where the predicate–argument structure is described and used as an input to the BERT model, and the other is the way to use the properties extracted from a raw sentence, such as the distance between the predicate and argument in the text. We showed that our proposed model yields state-of-the-art results in Korean semantic role labeling. The model also shows improvement in performance on the CoNLL-2009 English benchmark. The remainder of this paper describes our models and experimental results for semantic role labeling.

## 2. Korean Semantic Role Labeling

### 2.1. Task Definition

As mentioned in the introduction, semantic role labeling can be decomposed into four subtasks: predicate identification, predicate classification, argument identification, and argument classification. There are two representations for argument annotation: span-based and dependency-based. Semantic corpora such as PropBank usually represent arguments as syntactic constituents (spans), whereas the CoNLL-2009 shared task proposes dependency-based SRL, where the goal is to identify the syntactic heads of arguments rather than the entire span. In recent years, dependency-based approaches have become more popular than span-based approaches. Therefore, following this trend, we deal with the Korean SRL as a dependency-based annotation problem. In the Korean SRL benchmark, predicates are provided during training and testing, analogous to the English SRL benchmark. Most of

the recent studies have focused on the identification and classification of the arguments. In this study, we only discuss argument identification and classification.

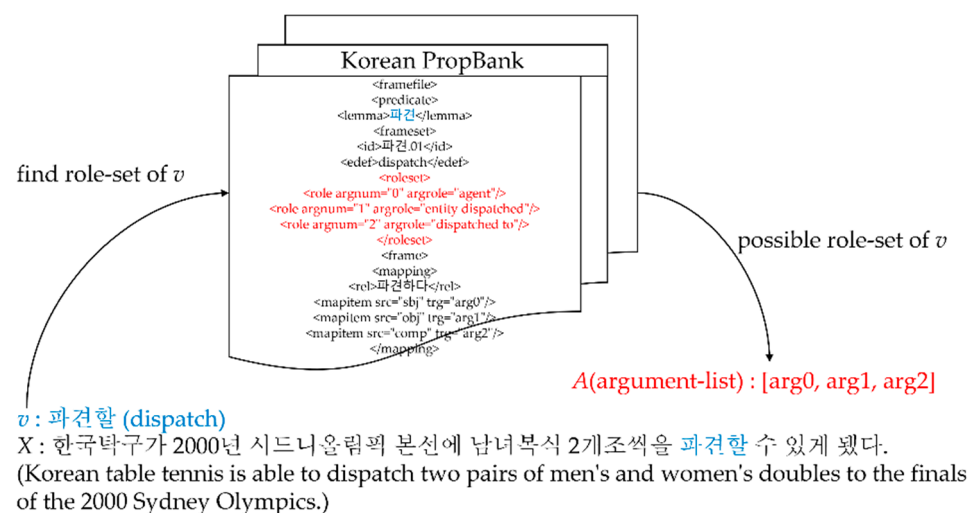
## 2.2. Argument Identification and Classification

The argument identification and classification task is to detect the argument syntactic heads and assign them with the correct semantic role labels. We formulated this task as a sequence-labeling problem. The input sentence is fed into the byte pair encoding (BPE) tokenizer, which splits some words into sub-tokens. To begin, for Korean input sentences, morpheme analysis was performed, and subsequently, the BPE tokenizer was applied. Unlike in English, in Korean, morpheme analysis is crucial because the Korean language uses a specific word to convey its meaning and grammatical utility. When given a sentence–predicate pair  $(X, v)$ , our task is defined as predicting a sequence of tags  $y$ , where the label set draws from the corresponding BIO tagging scheme and the arguments of the predicate (e.g., B-ARG1).

## 2.3. Simple Features

### 2.3.1. Semantic Feature

A convenient way to obtain semantic information is by utilizing the frameset files of PropBank. PropBank is a corpus annotated with verbal propositions and arguments. It also provides the frameset files, which define the argument labels used by the annotations on a per-verb basis. Each frameset file has one or more predicates, such as “make.01” or “make.02,” each of which is divided into coarse-grained word senses called role-sets. For each role-set, the frameset file provides descriptions of the argument roles, along with examples. We note the role-sets of PropBank. It considers a list of arguments that each predicate can take in the role-sets and uses it as input to the BERT model. Therefore, contrary to our previous definition, the input of our task changes from sentence–predicate pairs to sentence–predicate–argument-list triples  $(X, v, A)$ . Figure 2 shows an example of how to find the  $A$ (argument-list) described above in PropBank when the sentence  $X$  and the verb  $v$  are given (Korean PropBank is used as an example).



**Figure 2.** Example of finding an argument-list.

### 2.3.2. Text Feature

Among the features used for shallow learning in the past, positional information was available in the raw text. Even language models such as BERT, which demonstrate good results in the recent NLP tasks, use raw-text location information. Similarly, we used simple location information as features of the neural network model in the fine-tuning stage. Our locative features are as follows: (1) a word range feature that marks the beginning and the middle of a word, (2) a feature that indicates the textual distance between the word

to which each token constituting the sentence belongs and the word where the target predicate is located. Figure 3 shows an example of a text feature. The entered Korean sentence reads, “한국탁구가 2000 년 시드니올림픽 본선에 남녀복식 2 개조식을 파견할 수 있게 됐다. (Korean table tennis is able to dispatch two pairs of men’s and women’s doubles to the finals of the 2000 Sydney Olympics.)” In Figure 3, X is the input sentence, F1 is a word range feature, and F2 is the textual distance between each token and the predicate token. Each token is the result of applying BPE to the result of the morphology analysis, which will be explained later.

Korean table tennis				2000 Sydney Olympics						two pairs of men's and women's doubles					is able to dispatch							
X	한국	탁	구가_	2000	년_	시	드	니	올림픽_	...	2	개	조	식	을_	파	견	할_	수_	있게	됐다	.
F1	B	I	E	B	E	B	I	I	E	...	B	I	I	I	E	B	I	E	S	S	B	E
F2	-6	-6	-6	-5	-5	-4	-4	-4	-4	...	-1	-1	-1	-1	-1	0	0	0	1	2	3	3

Figure 3. Example of a text feature.

#### 2.4. Pre-Training Korean BERT

BERT is a language model consisting of a bidirectional transformer encoder stacked with several transformer blocks. The BERT consists of N transformer encoder blocks. The base model consists of 12 blocks, and the large model consists of 24 blocks, implying that the entire input sentence is constructed N times.

The input used to train the Korean BERT model is shown in Figure 4. Position embedding was used to indicate the location information in an input sentence. The BERT model received two sentences simultaneously and used the SEP token to separate the input sentences. Segment embedding was performed by feeding the value 0 until the first [SEP] token and 1 until the second [SEP] token appeared. Token embedding used BPE applied to the analysis results of Korean morphemes, and the last token of each morpheme was labeled with “\_” to separate the morpheme units. The first token of every input sentence started with the [CLS] token, the token vector used for classification problems. Finally, the input token vector created after the element-wise summation of the three embeddings was used as the input for the Korean BERT model.

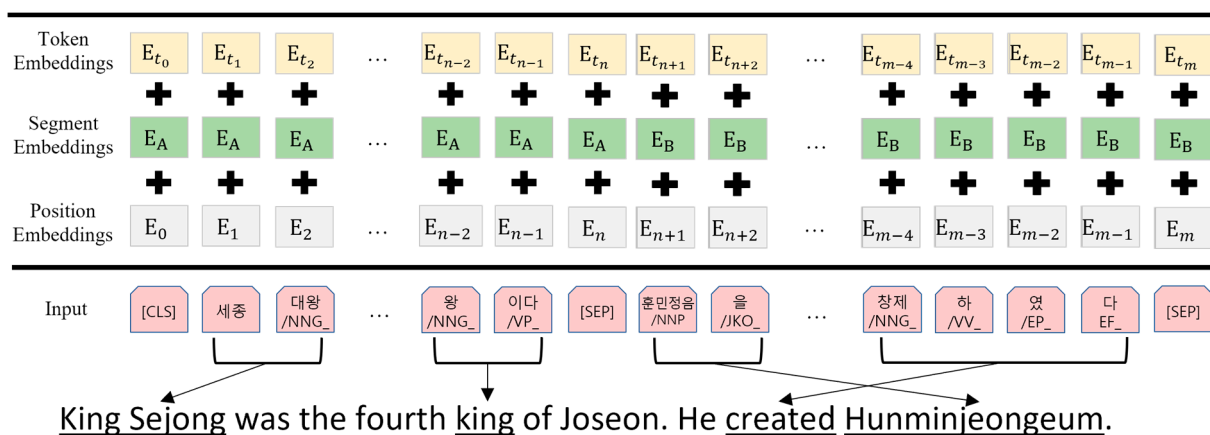


Figure 4. Example of input for Korean BERT model.

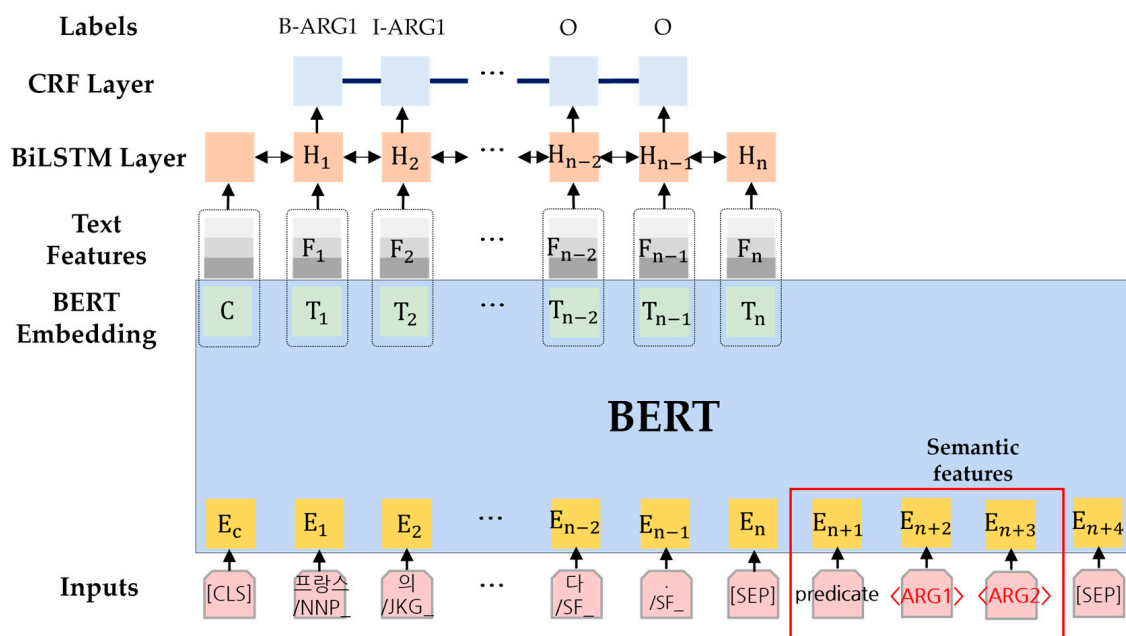
#### 2.5. Fine-Tuning Korean SRL Model

It is known that the BERT model can classify the label of the token by adding only the MLM classifier without any special architecture. However, many NLP studies mainly use the method of using BERT’s output as an input to other deep learning models. The study in [24] discussed the performance improvement when combining other deep learning models at the top of the BERT model. Therefore, this paper uses a Bidirectional Long Short-Term Memory Recurrent Neural Network (LSTM RNN) model that can propagate input

feature information in both directions. We also use a conditional random field (CRF) layer to leverage bidirectional information on output labels. CRF is known to help overcome constraints such as the inability of “I” tags to appear first in the BIO tagging method [25]. Following the precedent, we also use a BERT model combined with BiLSTM RNN and CRF.

The BERT-based Korean SRL model architecture is illustrated in Figure 5 at the point in the inference process where it outputs a tag for each token. To encode the sentence in a predicate-aware manner, we design the input as ([CLS] sentence [SEP] predicate, argument-list [SEP]), allowing the representation of the predicate and semantic information to interact with the entire sentence. The training data consist of the following: input sentence  $X = \{x_1, x_2, \dots, x_n\}$ , features  $F = \{f_1, f_2, \dots, f_n\}$ , and semantic label  $Y = \{y_1, y_2, \dots, y_n\}$ . The [unused] token is used to feed the argument-list as input to the BERT model. For example, the token [unused1] is replaced by the token <ARG1>. The input sequence  $x_i$  becomes  $E_i$  and is fed into the BERT encoder with the amalgamation of the three embeddings. The contextual representation of the sentence from the BERT model is expressed as follows:

$$\{T_1, \dots, T_n\} = \text{BERT}(E_1, \dots, E_n)$$



**Figure 5.** Architecture of our semantic role labeling model.

The  $T_i$  created through the BERT encoder is concatenated with the embedding  $f_i$  containing the feature information to become the input of the BiLSTM RNN, which is expressed as follows:

$$\{h_1, \dots, h_n\} = \text{BiLSTM RNN}(T_1; f_1, \dots, T_n; f_n)$$

For the final prediction, each token  $h_i$  is fed into the CRF layer and classified over the label set.

### 3. Experiments

#### 3.1. Experimental Setup

We conducted experiments on two SRL tasks: the Korean PropBank (KPB) [26] and CoNLL-2009 English. For the Korean dataset, we constructed the same training and evaluation data as in previous studies. A total of 19,302 sentences were used for training, and 3733 sentences were used for evaluation. In English, we followed standard splits for training, development, and test sets. In order to compose a simple semantic feature for



Korean SRL, 2835 frameset files in KPB were used. There were 1217 sentences in KPB where a simple semantic feature was not available. In this case, the “<None>” token was used. In addition, 10,687 frameset files in PropBank were used to construct simple English semantic features. In this study, we used a pre-trained KorBERT model for Korean and a pre-trained BERT model for English.

The BiLSTM RNN model used for fine-tuning the Korean SRL model comprised four 1200 dimensional hidden layers. The KorBERT representation comprised 768 dimensions, and the feature embedding had 2000 dimensions. The BiLSTM RNN model used for fine-tuning the English SRL model used one hidden layer of 1600 dimensions. The BERT representation was 768 dimensions for the base models, 1024 dimensions for large models, and 800 dimensions for feature embedding. The dropout value was 0.1, which was the same for both models. Adam was used as the learning algorithm, and CRF was used as the output layer of the BiLSTM RNN for dependency modeling. All the experimental environment variables described above were marked with optimized values.

### 3.2. Experimental Results

Table 1 shows the experimental results of the Korean semantic role labeling. The models in [27,28] used word embedding made with word2vec and did not use syntactic information, such as dependency trees and end-to-end neural network models. Reference [29] applied a structural support vector machine (S-SVM) for semantic role labeling and utilized information, such as syntactic information and name entity recognition results. The experiment showed that when the BERT model was applied to Korean semantic role labeling, the F1 score was 83.62%, which was an improvement of approximately 4% in comparison with that of the previous study. This result seems to be an appropriate answer to the question we asked at the beginning of the paper, whether we can apply BERT to Korean, a low-resource language. So as to determine the effect of the two simple features proposed in this paper on Korean semantic role labeling, experiments were conducted by adding each of them to the BERT-BiLSTRNN-CRF model. The results showed that the F1 score improved by 1.33% when the semantic features were used and by 2.15% when the text features were added. Finally, when the two feature information parts were used together, our model showed an F1 score of 86.36%, ultimately yielding state-of-the-art performance on the Korean PropBank benchmark. The important reasons for this improvement can be that these features are applicable to various datasets, do not increase the structural complexity of the model, and are easy to apply to other language models.

**Table 1.** F1 on the Korean PropBank benchmark with gold predicates.

Model	Test F1
Stacked BiLSTM RNN CRF [27]	78.57
Highway BiLSTM RNN CRF [28]	78.84
S-SVM with syntactic information and extra knowledge [29]	79.54
BERT-BiLSTM-RNN-CRF (our model)	
(Bert-base)	83.62
+ Semantic features	84.95
+ Text features	85.77
+ Semantic features + Text features	86.36

We conducted the experiments to confirm that the proposed features and models performed well in both Korean and English benchmarks. Table 2 shows the experimental results of English semantic role labeling on the CoNLL-2009 English benchmark. All models in [14,15] use neural networks by incorporating syntactic feature information, such as dependency trees. Reference [22] is a model in which BERT is applied to semantic role labeling, and the BiLSTM-RNN-MLP model is used in the fine-tuning step. The model in [22] and our model did not use syntactic features, unlike the [14,15] models. Unlike our

model, the BiLSTM-RNN-MLP model does not use feature information. Results showed that the BERT-BiLSTM-RNN-CRF model with the proposed features showed the best F1 score. Through the model in [22] and the results of our model, we could achieve better performance than the existing models without syntax information. In addition, analogous to the Korean experiment, it was proved that the performance improved when each feature information was added to the English experiment. Finally, there was a higher performance improvement when applied to Korean (+2.74%) than when applied to the text features (+1.87%) in the English experiment, owing to the difference in the characteristics of English, which consists of words, and Korean, which consists of several morphemes. Therefore, text features play a significant role in Korean semantic role labeling.

**Table 2.** F1 on the CoNLL-2009 English benchmark with gold predicates.

Model	Dev F1	Test F1
LSTM-GCNs [14]	84.2	-
Bi-Affine [15]	85.6	-
BERT-BiLSTM-RNN-MLP [22]		
(Bert-large)	88.7	89.8
(Bert-base)	89.3	90.3
BERT-BiLSTM-RNN-CRF (our model)		
(Bert-base)	87.05	88.29
+ Semantic features	87.66	88.90
+ Text features	88.84	89.87
+ Semantic features + Text features	89.08	90.16
(Bert-large)	89.02	90.13
+ Semantic features + Text features	89.34	90.39

#### 4. Discussion

In this study, we proposed simple semantic information for the BERT-based SRL model. It has approximately 2.74% and 1.87% higher F1 scores in Korean and English datasets, respectively. However, we found some notable observations in the experimental results. First, in the English dataset experiment using the BERT-large model, the performance improvement due to the addition of features is not enough. We performed a paired *t*-test to determine whether this result was valid, and as a result, it was confirmed that the *p*-value was valid as 0.0021. This result shows that the use of additional feature information is maximized in the language model of the small size than in the language model of the large size. Next, there is a question about the difference in performance traction between the two features. Next is the observation result of the difference in performance traction between the two features. Text features showed good performance improvement regardless of the dataset, whereas semantic information showed lower performance traction than the English dataset. In the case of text features, features can be extracted regardless of the dataset, whereas semantic features are considered to be affected by each built-in language resource when performing information extraction.

#### 5. Conclusions

In this study, the BERT model was applied to semantic role labeling for Korean, a low-resource language, and state-of-the-art performance was obtained. In addition, we proposed two feature information aspects that can be easily obtained from the corpus. Experimental results demonstrate that both the feature information aspects considerably improved the Korean semantic role labeling performance. In addition, by applying the same method to the CoNLL-2009 English benchmark, it was confirmed that there was an improvement in performance in both Korean and English semantic role labeling. Conclusively, we were able to get all the answers to the questions raised at the beginning of the paper as follows: (1) It was confirmed that BERT worked well in Korean, a low-

resource language. (2) Simple feature information was easily obtained and used easily. (3) There was an improvement in performance in Korean and English datasets through simple feature information.

Further study is required to address the various follow-up questions, such as “will the use of syntactic feature information improve performance further?” and “will the performance improve when we apply our proposed features to other NLP tasks?” This will be dealt with in the follow-up study.

**Author Contributions:** Conceptualization, J.B.; Formal analysis, C.L.; Methodology, J.B.; Writing—original draft, J.B.; Writing—review & editing, C.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the Institute for Information & Communications Technology Promotion (IITP) grants funded by the Korean government (MSIT) (No. 2013-0-00131, Development of Knowledge Evolutionary WiseQA Platform Technology for Human Knowledge Augmented Services).

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

## References

- Chen, Y.N.; Wang, W.Y.; Rudnicky, A.I. Unsupervised induction and filling of semantic slots for spoken dialogue systems using frame-semantic parsing. In Proceedings of the 2013 IEEE Workshop on Automatic Speech Recognition and Understanding, Olomouc, Czech Republic, 8–12 December 2013.
- Shen, D.; Lapata, M. Using semantic roles to improve question answering. In Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL), Prague, Czech Republic, 28–30 June 2007.
- Yih, W.; Richardson, M.; Meek, C.; Chang, M.; Suh, J. The value of semantic parse labeling for knowledge base question answering. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL), Berlin, Germany, 7–12 August 2016.
- Fader, A.; Soderland, S.; Etzioni, O. Identifying relations for open information extraction. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, Edinburgh, UK, 27–31 July 2011.
- Bastianelli, E.; Castellucci, G.; Croce, D.; Basili, R. Textual inference and meaning representation in human robot interaction. In Proceedings of the Joint Symposium on Semantic Processing, Textual Inference and Structures in Corpora, Trento, Italy, 20–23 November 2013.
- Zhou, J.; Xu, W. End-to-end learning of semantic role labeling using recurrent neural networks. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, Beijing, China, 27–31 July 2015.
- He, L.; Lee, K.; Lewis, M.; Zettlemoyer, L. Deep semantic role labeling: What works and what’s next. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, Vancouver, BC, Canada, 30 July–4 August 2017.
- Tan, Z.; Wang, M.; Xie, J.; Chen, Y.; Shi, X. Deep semantic role labeling with self-attention. In Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18), New Orleans, LA, USA, 2–7 February 2018.
- He, L.; Lee, K.; Levy, O.; Zettlemoyer, L. Jointly predicting predicates and arguments in neural semantic role labeling. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, Melbourne, Australia, 15–20 July 2018.
- Levin, B. *English Verb Classes and Alternations: A Preliminary Investigation*; University of Chicago Press: Chicago, IL, USA, 1993.
- Punyakank, V.; Roth, D.; Yih, W. The importance of syntactic parsing and inference in semantic role labeling. *Comput. Linguist.* **2008**, *34*, 257–287. [[CrossRef](#)]
- Roth, M.; Lapata, M. Neural semantic role labeling with dependency path embeddings. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, Berlin, Germany, 7–12 August 2016.
- Strubell, E.; Verga, P.; Andor, D.; Weiss, D.; McCallum, A. Linguistically-informed self-attention for semantic role labeling. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, 2–4 November 2018.
- He, S.; Li, Z.; Zhao, H.; Bai, H. Syntax for semantic role labeling, to be, or not to be. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Melbourne, Australia, 15–20 July 2018.
- Shi, P.; Zhang, Y. Joint bi-affine parsing and semantic role labeling. In Proceedings of the 2017 International Conference on Asian Language Processing (IALP), Singapore, 5–7 December 2017.
- Peters, M.; Neumann, M.; Iyyer, M.; Gardner, M.; Clark, C.; Lee, K.; Zettlemoyer, L. Deep contextualized word representations. In Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, New Orleans, LA, USA, 1–6 June 2018.
- Radford, A.; Narasimhan, K.; Salimans, T.; Sutskever, I. Improving language understanding by generative pre-training. *arXiv* **2018**, arXiv:2012.11747.



18. Devlin, J.; Chang, M.; Lee, K.; Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv* **2018**, arXiv:1810.04805.
19. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.; Kaiser, L.; Polosukhin, I. Attention is all you need. *arXiv* **2017**, arXiv:1706.03762.
20. Yang, Z.; Dai, Z.; Yang, Y.; Carbonell, J.; Salakhutdinov, R.; Le, Q. XLNet: Generalized Autoregressive Pretraining for Language Understanding. *arXiv* **2020**, arXiv:1906.08237v2.
21. Brown, T.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; et al. Language Models Are Few-Shot Learners. *arXiv* **2020**, arXiv:2005.14165v4.
22. Shi, P.; Lin, J. Simple BERT Models for Relation Extraction and Semantic Role Labeling. *arXiv* **2019**, arXiv:1904.05255.
23. Kingsbury, P.; Palmer, M. From treebank to propbank. In Proceedings of the Third International Conference on Language Resources and Evaluation (LREC'02). European Language Resources Association (ELRA), Las Palmas, Spain, 29–31 May 2002.
24. Li, X.; Bing, L.; Zhang, W.; Lam, W. Exploiting BERT for End-to-End Aspect-based Sentiment Analysis. In Proceedings of the 2019 EMNLP Workshop W-NUT: The 5th Workshop on Noisy User-Generated Text, Hong Kong, China, 4 November 2019.
25. Huang, Z.; Xu, W.; Yu, K. Bidirectional LSTM-CRF Models for Sequence Tagging. *arXiv* **2015**, arXiv:1508.01991.
26. Palmer, M.; Ryu, S.; Choi, J.; Yoon, S.; Jeon, Y. Korean Propbank LDC2006T03. Available online: <https://catalog.ldc.upenn.edu/LDC2006T03> (accessed on 19 May 2022).
27. Bae, J.; Lee, C. Korean Semantic Role Labeling Using Stacked Bidirectional LSTM-CRFs. *J. KIISE* **2016**, *44*, 36–43. [[CrossRef](#)]
28. Bae, J.; Lee, C. Korean Semantic Role Labeling with Highway BiLSTM-CRFs. In Proceedings of the HCLT, Daegu, Korea, 13–14 October 2017.
29. Lim, S.; Kim, H. A Study of Korean Semantic Role Labeling using Word Sense. In Proceedings of the HCLT, Jeonju, Korea, 16–17 October 2015.