

Unlocking Korean Verbs: A User-Friendly Exploration into the Verb Lexicon

Seohyun Song^{1*} Eunkyul Leah Jo^{2*} Yige Chen^{3*} Jeon-Pyo Hong^{4*} Kyuwon Kim⁵
Jin Wee⁶ Miyoung Kang⁶ KyungTae Lim^{7†} Jungyeul Park^{2†} Chulwoo Park⁸

¹Seoultech, South Korea ²The University of British Columbia, Canada

³The Chinese University of Hong Kong, Hong Kong ⁴42dot Inc., South Korea

⁵Seoul National University, South Korea ⁶National Institute of Korean Language, South Korea

⁷KAIST, South Korea ⁸Anyang University, South Korea

<https://www.korean.go.kr>

Abstract

The Sejong dictionary dataset offers a valuable resource, providing extensive coverage of morphology, syntax, and semantic representation. This dataset can be utilized to explore linguistic information in greater depth. The labeled linguistic structures within this dataset form the basis for uncovering relationships between words and phrases and their associations with target verbs. This paper introduces a user-friendly web interface designed for the collection and consolidation of verb-related information, with a particular focus on subcategorization frames. Additionally, it outlines our efforts in mapping this information by aligning subcategorization frames with corresponding illustrative sentence examples. Furthermore, we provide a Python library that would simplify syntactic parsing and semantic role labeling. These tools are intended to assist individuals interested in harnessing the Sejong dictionary dataset to develop applications for Korean language processing.

1 Introduction

When examining the verb lexicon for Korean, the Sejong dictionary has been a major language resource. As part of the 21st Century Sejong Project, which aims to advance Korean language information processing, the Sejong dictionary has produced extensive datasets that describe Korean lexicon data in great detail. Although it is designed with a vast amount of morphological, semantic, and syntactic information, there have been minimal attempts to utilize its resources. This paper investigates the Korean verb lexicon, with a particular focus on the subcategorization frames in the Sejong verb dictionary. The objective is to facilitate easier access to and manipulation of the Sejong dictionary, thus promoting further studies and applications in Korean language processing.

*Equally contributed authors. †Corresponding authors: KyungTae Lim and Jungyeul Park.

In annotating corpora with semantic roles, the methodology of Proposition Bank (PropBank) has been widely adopted in natural language processing (Palmer et al., 2005). Similar efforts can be observed in the Sejong dataset, which seeks to provide a meaningful level of linguistic representation and a corpus of annotated data to facilitate further research and analysis. This structured dataset will serve as a foundation for identifying relationships between words, such as organizing and searching for verbs that share the same subcategorization frames.

Within the Sejong dictionary are XML files for each verb (Figure 1). Inside are XML tags that identify each word’s morphological, semantic, and syntactic information with additional metadata. Some annotations make it possible to identify arguments such as semantic roles of verbs in sentences (Figure 2). A set of the underlying semantic roles for each verb and frame groupings are defined for each sense of the verb. Then, in each group are the frames that annotate selection restrictions with example sentences. We believe this detailed verb representation level is vital for many language processing applications. With thorough and manual documentation on subcategorization in the Sejong dictionary, we aim to enhance the dataset’s usability, facilitating the creation of user-friendly systems for all future projects.

In this paper, we introduce a web interface created for individuals with minimal programming experience or those seeking quick access to the verbs in the Sejong dictionary. This interface consolidates verb information and provides easy access to subcategorization frames. It also includes our efforts to map information by projecting subcategorization frames onto corresponding sentence examples. Additionally, we provide a Python library that includes tools for syntactic parsing and semantic role labeling, designed for those who wish to utilize the Sejong dictionary in developing Korean

```

<?xml version="1.0" encoding="utf-8" standalone="no"?>
<!DOCTYPE superEntry SYSTEM "vv.dtd">
<superEntry>
  <orth>확립하다</orth>
  <entry pos="vv">
    <mnt_grp>
      ...
    </mnt_grp>
    <morph_grp>
      <var type="spr">확립을 하다</var>
      <cntr opt="opt" type="VDelCCon"/>
      <str>N.하</str>
      <org lg="si">確立</org>
      <infl type="yeo"/>
    </morph_grp>
    <sense>
      <sem_grp>
        <sem_class>중요행위</sem_class>
        <trans>establish</trans>
        <trans>settle</trans>
        <trans>fix</trans>
      </sem_grp>
      <frame_grp type="FTR">
        <frame>X=N0-0이 Y=N1-을 V</frame>
        <subsense>
          <sel_rst arg="X" tht="AGT">인간|인간집단</sel_rst>
          <sel_rst arg="Y" tht="THM">사실명제|규범|관습|제도|종교|사조|관계추상적대상|개념|기준</sel_rst>
          <eg>중세는 기독교 정신을 확립한 시기다.</eg>
          <eg>자신의 신념을 확립하지 못한 사람들은 작은 어려움에도 흔들리기 마련이다.</eg>
          <eg>많은 사람들이 사회의 질서를 확립하려는 노력을 기울여 왔다.</eg>
        </subsense>
      </frame_grp>
    </sense>
  </entry>
</superEntry>

```

Figure 1: Example from the Sejong dictionary: 확립하다 *hwaglibhada* (‘establish’)

Frame:	X=N0-0이	Y=N1-을	V
	arg="X" tht="AGT"	arg="Y" tht="THM"	
Sentence:	[많은 사람들이] _{AGT}	[사회의 질서를] _{THM}	[확립하려는] _{TARGET} ...
	<i>manh-eun salamdeul-i</i>	<i>sahoewi jilseoleul</i>	<i>hwaglibhalyeoneun</i> ...
	Many people.nom	social order.acc	establish
	‘Many people have made efforts to establish social order.’		

Figure 2: Example of subcategorization frame to sentence projection

language processing applications. Therefore, our main contribution is two-fold: First, we offer a corpus of sentence examples from the Sejong dictionary with annotated argument-predicate boundaries. Second, we develop two key tools—a web interface and programming libraries—to facilitate access to Sejong dictionary information.

2 Previous work

2.1 Subcategorization frames

Subcategorization refers to the tendency that a lexeme, usually a verb, has its restrictions regarding the syntactic arguments it takes. To address such tendencies, subcategorization frames are provided as the formalization of subcategorization, explicitly describing the arguments the verbs may select. There are two major studies of subcategorization frames for Korean in addition to the Sejong dic-

tionary: (1) manually defined subcategorization frames in the context of tree-adjointing grammars (Han et al., 2000), and (2) automatically extracted subcategorization frames from the treebank (Xia et al., 2000; Park, 2006). Subcategorization frames have been linguistically studied thoroughly (Park, 2002; Kim, 2004), mostly for distinguishing between arguments and adjuncts in Korean within subcategorization frames.

Before presenting our method for improving access to verb subcategorization, we will briefly review previous projects that describe Korean grammar, each specifying a distinct subcategorization frame. Additionally, two key Korean language resources provide verb-related information: the Korean Proposition Bank and the National Institute of Korean Language (NIKL) Semantic Role Labeling (SRL) dataset. To assess and compare the

scope of these resources, we conducted a quantitative analysis by examining the number of verbs and their associated frames, as detailed in Table 1. Korean Proposition Bank (Korean PropBank)¹ has achieved the same level of recognition and development as its English counterpart, Penn’s Proposition Bank (Palmer et al., 2005). As resources, Proposition Banks provide information about the predicate-argument structure of sentences in a corpus. Similarly, Korean PropBank provides annotated sentences from the Penn Korean treebank (Han et al., 2002). The NIKL SRL also provides detailed verb information in addition to the SLR corpus. It refines 1,597 verbs from the previous Sejong dictionary’s entry and adds 2,063 new verbs along with their subcategorization frame information. Interestingly, it follows the annotation scheme of Penn’s Propbank with the Sejong treebank labels. As the name suggests, its primary goal is to provide the SRL annotated corpus for Korean.

	Sejong	K-PropBank	NIKL
# of verbs	15,181	2749	3660
av. # of frames	1.812	1.408	1.593

Table 1: Basic statistics: number of verbs and their average number of frames per verb

2.2 Previous usages of the Sejong dictionary

The Sejong dictionary has been adopted in a limited way. Data and example sentences from the Sejong dictionary were used in the Korean SRL dataset developed by the National Institute of Korean Language, which was later improved through a linguistically-informed annotation strategy (Chen et al., 2024c). Additionally, a small selection of example sentences from the Sejong dictionary was incorporated into Korean FrameNet (Hahm et al., 2018; Chen et al., 2024a), guided by the frames outlined in the dictionary entries within the Sejong dictionary.^{2,3}

3 Developing a Web Interface

We offer a user-friendly interface that allows for easy navigation and access to the Sejong dictionary, catering to individuals with non-computational backgrounds. Visually, words are organized in Korean alphabetical order, enabling users to select each word to examine its linguistic details. A

¹<https://catalog.ldc.upenn.edu/LDC2006T03>

²<https://github.com/machinereading/koreanframenet>

³<https://github.com/jungyeul/k-framenet>

particular piece of critical information that we decided to include is example sentences illustrating the identified argument boundary for each word. To complement this feature, we can expand the service for searching words by their frame, argument, morphological inflection and semantic group. This new systematic approach and visual aid in organizing the Sejong dictionary can help many users understand the relationships between words more intuitively.

Data in the Sejong dictionary is organized so that for each verb lexeme, a set of its morphological, syntactic, and semantic information is provided along with example sentences. The morphological information is usually represented as inflections given the agglutinative feature of Korean, and the dataset specifies whether the inflections are regular. Furthermore, the Sejong dictionary distinguishes different senses for each lexeme. Under each of its possible senses, the dictionary offers both the verb category and a translation of the sense, along with the corresponding frame that indicates the arguments required by the verb.

확립하다

확립을 하다, N.하, 確立_

Sense 1

semantic class: 종료행위
translation: establish, settle, fix

Frame 1

X=N0-이 Y=N1-을 V
arg="X" tht="AGT"
arg="Y" tht="THM"

Example Sentences

- 중세는 기독교 정신을 확립한 시기다.
- 자신의 신념을 확립하지 못한 사람들은 작은 어려움에도 흔들리기 마련이다.
- 많은 사람들이 사회의 질서를 확립하려는 노력을 기울여 왔다.

Figure 3: Interface screenshot of a typical word details page including morphological, semantic, and syntactic frame information along with their associated semantic roles and sentence examples.

Figure 3 is a screenshot of our interface upon displaying the word 확립하다 *hwaglibhada* (‘establish’). We obtain the word’s information from the Sejong dictionary’s XML file. The tags that organize the lexicon information are taken from the files when rendering this screen. For example, the orthographic information of the verb, as shown below,

<orth>확립하다</orth>

corresponds with the title of the page. Other morphological and lexical tags such as these are also displayed near the top as the subtitle of the word’s detail page, suggesting that the verb can be considered as the combination of a nominal morpheme of Chinese origin and the verb-generating derivational suffix 하 (ha):

```
<morph_grp>
<var type="spr">확립을 하다</var>
<str>N.하</str>
<org lg="si">確立_</org>
</morph_grp>
```

Under the tag sense is the semantic information, including the semantic class and its English translation, denoted by <sem_class> and <trans> correspondingly. The subcategorization frame of the instance corresponds with tags under <frame_grp> and <frame>. Further below are denoted by <subsense> and <sel_rst> to display the information with regard to the arguments of the verb. The semantic roles of the required arguments are introduced in <sel_rst>, including the general categories such as AGENT (AGT) and THEME (THM), as well as specific categories such as 인간 *ingan* (‘human’) and 사실명제 *sasilmyeongje* (‘factual proposition’). The example of the sentences with argument boundaries, marked by <eg> in XML, are displayed as the last section on the details page.

In the web interface, we also introduce the distinction of sentence argument boundaries for the sentence examples. These argument boundaries are determined based on the frame information of the provided verb in the Sejong dictionary, as shown in Figure 4. The argument boundaries are presented using brat (Stenetorp et al., 2011)⁴ for the following sentence:

[많은 사람들이]	[사회의 질서를]	[확립하려는]	...
AGENT	THEME	TARGET	
<i>manheun salam-deul-i</i>	<i>sahoe-ui jilseo-leul</i>	<i>hwaglibha...</i>	...
many person-PL-NOM	society-GEC order-ACC	establish	...
('Many people have made efforts to establish order in society.')			

where the frame suggests that the predicate requires two arguments: an agent and a theme. Consequently, these roles are assigned to the detected arguments in the example sentence.

The argument boundaries are detected through a combination of a heuristic-based chunking method and the sentence’s dependency structures. Argument boundaries are determined by the locations of postpositions in the sentence, as postpositions in Korean typically indicate the endpoints of arguments and modifiers, based on our previous

⁴<https://brat.nlplab.org>

work (Chen et al., 2024c). On the other hand, dependency structures assist in identifying argument boundaries when the heuristics are insufficient to derive chunks for certain constituents (Chen et al., 2022). The postpositions of the chunks are further compared with the postpositions given by the frame to determine whether a chunk is an argument or not.

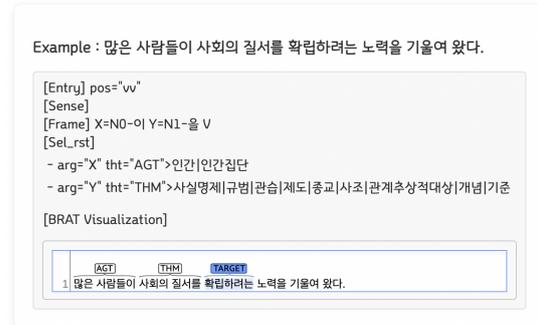


Figure 4: Subcategorization frames of the verb, along with its semantic role information for arguments, including chunked structures based on frame information to identify argument boundaries

The web interface also provides additional functionalities to enable easy navigation for users and linguists, particularly those without a background or expertise in computer science or computing. Accordingly, users can swiftly navigate through the dictionary by searching for specific verbs, frames, arguments, morphological inflection and semantic role groups. After a search, corresponding words will be listed and clicking upon the button for Sejong will guide the user to a word details page, as in Figure 5.

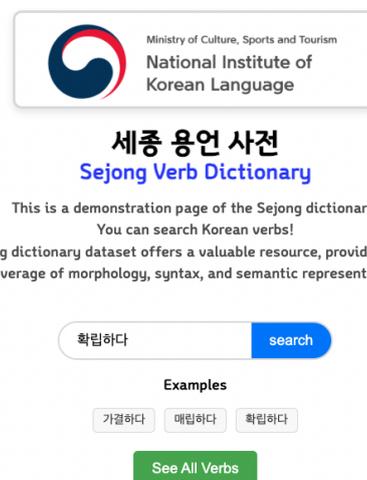


Figure 5: First page of the web interface, which allows users to search for Korean verbs

4 Python Library

This section introduces the `pySejongFrame`, a Python library designed for processing and accessing the 21st-century Sejong Frame Dictionary. With this tool, users can streamline the manipulation of linguistic data for research and development purposes. This library provides various methods for loading and querying the corpus, with support for both direct and lazy loading techniques. Below, we provide a detailed walk-through of the system's initialization and querying capabilities, showcasing its practical use for both corpus-based applications and linguistic research.

4.1 Library initialization

To begin using the `pySejongFrame` library, users must set up the environment by specifying paths for the library and data files. The system assumes that the XML data for the Sejong Frame Dictionary is located in the `share/sejong_framefiles` directory.

```
import sys
import os
from pathlib import PurePath, Path
PATH_THIS = PurePath(os.getcwd()).parent
PATH_PYLIB = Path(PATH_THIS, "python")
PATH_SHARE = Path(PATH_THIS, "share")
sys.path.append(str(PATH_PYLIB.resolve()))
```

Once the environment is configured, users may create an instance of the `SejongFrameCorpusReader`.

4.2 Loading methods

The library offers four different methods for loading the Sejong Frame Dictionary. This allows flexibility depending on the user's requirements. Below, we describe each loading method:

Direct loading using `SejongFrameCorpusReader`

In this approach, the corpus is loaded directly using the `SejongFrameCorpusReader` class.

```
from SejongFrame.reader \
    import SejongFrameCorpusReader
sejong_framenet = SejongFrameCorpusReader(
    Path(PATH_SHARE, "sejong_framefiles"), "*.xml")
```

Lazy loading with `LazyClassLoader` Lazy loading defers the creation of the index until the corpus is actually accessed. This can save time when dealing with large datasets by delaying the initialization process.

```
from SejongFrame.loader \
    import LazyClassLoader
sejong_framenet = LazyClassLoader(
    "sejong_framenet", SejongFrameCorpusReader,
    Path(PATH_SHARE, "sejong_framefiles"), "*.xml")
```

Integration with `nltk` The `pySejongFrame` library can be integrated with `nltk`, enabling users familiar with the `nltk` framework to leverage its tools in conjunction with the Sejong Frame Dictionary.

```
from SejongFrame.nltk.reader \
    import SejongFrameCorpusReader

sejong_framenet = SejongFrameCorpusReader(
    "sejong_framefiles",
    r"*\.\.xml",
    nltk_data_subdir=str(PATH_SHARE)
)
```

4.3 Corpus querying and usage

After loading the Sejong Frame Dictionary, the library allows users to query the corpus for specific linguistic data. The example below demonstrates how to retrieve entries related to the verb `가감하다` ('adjust').

```
gagam = sejong_framenet.entries("가감하다")
print(repr(gagam))
```

Users can also query specific verb forms, such as `수정하다` ('modify') with the tag `vv`, which refers to the verb form:

```
sujeong_vv \
    = sejong_framenet.entries("수정하다.vv")
```

```
print(repr(sujeong_vv))
```

```
# Accessing a specific entry
print(sujeong_vv[2])
```

4.4 Exploring Frame Semantics

The Sejong Frame Dictionary includes rich frame semantics, which allow users to explore semantic classes and frame structures freely. For example, we can examine the semantic group (`sem_grp`) and translation of the word `수정하다` by accessing its specific sense:

```
sense = sujeong_vv['vv.3']['1']
```

```
# Semantic group
print(sense.category)
```

```
# Translation
print(sense.trans)
```

```
# Frames within the sense
print(len(sense.frames), sense.frames)
```

Frames also contain arguments, which can be accessed in the following manner:

```
frame = sense.frames[0]
```

```
# Accessing a specific argument
argY = frame.Y
```

```
# Printing the argument in text form
print(str(argY))
```

4.5 Error handling and best practices

While using the `pySejongFrame` library, users are advised to handle potential errors that may arise due to inconsistent corpus formats or malformed queries. For instance, ensuring that entries exist for the queried lemma or morpheme tag combination can prevent runtime exceptions.

```
try:
    target = sujeong_vv['vv.3']
except KeyError:
    print("Entry not found for the specified key.")
```

4.6 Performance considerations

For large-scale linguistic research, the library’s lazy loading feature (§4.2) is highly recommended. In this way, this method optimizes the system’s performance by postponing indexing until the data is actively queried. Furthermore, the integration with `nltk` allows for compatibility with existing NLP workflows, making the system versatile for different use cases.

4.7 Summary

The `pySejongFrame` library provides a toolset for accessing and analyzing the Sejong Frame Dictionary. With multiple loading options, efficient querying mechanisms, and integration with widely-used NLP libraries, it enables researchers to conduct comprehensive analyses of frame semantics in Korean. The library’s ability to work with large datasets and its support for lazy loading make it an ideal choice for corpus-based linguistic studies.

5 Conclusion and Future Perspectives

Our current contributions, particularly the Python library, have the potential to be applied across various linguistic tasks, including syntactic parsing systems that utilize lexical information, such as lexicalized parsers (Sarkar and Han, 2002; Park et al., 2013), semantic role labeling (Changki et al., 2015; Seo-In, 2019; Chen et al., 2024c,b), and FrameNet parsing (Park et al., 2014; Kim et al., 2016; Hahm et al., 2018; Chen et al., 2024a) for the Korean language. We are committed to broadening access to Korean language data beyond traditional boundaries, offering a unified platform for linguists and computer scientists to explore the rich resources of the Korean language. This platform will be invaluable for a wide range of Korean language processing tasks and future linguistic research.

We are continuously working on integrating additional Korean verb lexicons, such as the Korean

PropBank and FrameNet, into our web service. Our ultimate goal is to develop a comprehensive Korean VerbNet. By systematically comparing verbs and their subcategorization frames across the Sejong verb dictionary, PropBank, and other resources—and resolving any discrepancies—we aim to harmonize frame information across our interfaces. This will offer users a more cohesive and seamless experience when exploring Korean language data, contributing to advancements in Korean linguistic research and natural language processing. Our system, including its Python library with sample verb files, can be accessed at <http://117.17.185.27:8025/>.



Disclaimer

The Sejong Dictionary of the Sejong Project, was initiated by the Ministry of Culture and Tourism in Korea as part of the 21st Century Sejong Project in 1998. The goal of this initiative was to advance Korean language information processing, resulting in the creation of the largest corpus and language resources for the Korean language. Previously, third parties were prohibited from distributing content from the Sejong Project, including derivative works like the Python library. However, after successful negotiations with the National Institute of Korean Language, which holds the project’s copyright, we have secured permission to distribute these materials. This milestone has significantly improved accessibility, allowing for broader use of these valuable linguistic resources. Although the dictionary files required for the Python library must still be obtained directly from the National Institute of Korean Language, the institute will host the web system, ensuring it is available for public use. This arrangement marks an important step forward in making the Sejong Project’s resources more widely accessible to researchers and developers alike.

Limitation

One limitation of this work is that, since the dataset is static, the system may not account for the evolution of the language or new linguistic phenomena without regular updates to both the dataset and system functionality.

Acknowledgments

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant, funded by the Korea government (MSIT) (No.RS-2024-00456709, A Development of Self-Evolving Deepfake Detection Technology to Prevent the Socially Malicious Use of Generative AI) awarded to KyungTae Lim.

References

- Lee Changki, Lim Soo-jong, and Kim Hyun-Ki. 2015. Korean Semantic Role Labeling Using Structured SVM. *Journal of KIISE*, 42(2):220–226.
- Yige Chen, Jae Ihn, KyungTae Lim, and Jungyeul Park. 2024a. [Towards Standardized Annotation and Parsing for Korean FrameNet](#). In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 16653–16658, Torino, Italy. ELRA and ICCL.
- Yige Chen, Eunkyul Leah Jo, Yundong Yao, KyungTae Lim, Miikka Silfverberg, Francis M. Tyers, and Jungyeul Park. 2022. [Yet Another Format of Universal Dependencies for Korean](#). In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 5432–5437, Gyeongju, Republic of Korea. International Committee on Computational Linguistics.
- Yige Chen, Kyuwon Kim, KyungTae Lim, Jungyeul Park, and Chulwoo Park. 2024b. [When the Misidentified Adverbial Phrase Functions as a Complement](#). In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 12326–12336, Miami, Florida, USA. Association for Computational Linguistics.
- Yige Chen, KyungTae Lim, and Jungyeul Park. 2024c. [A Linguistically-Informed Annotation Strategy for Korean Semantic Role Labeling](#). In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 733–738, Torino, Italy. ELRA and ICCL.
- Younggyun Hahm, Jiseong Kim, Sunggoo Kwon, and Key-Sun Choi. 2018. Semi-automatic Korean FrameNet Annotation over KAIST Treebank. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan. European Language Resources Association (ELRA).
- Chung-Hye Han, Na-Rae Han, Eon-Suk Ko, Martha Palmer, and Heejong Yi. 2002. Penn Korean Treebank: Development and Evaluation. In *Proceedings of the 16th Pacific Asia Conference on Language, Information and Computation*, pages 69–78, Jeju, Korea. Pacific Asia Conference on Language, Information and Computation.
- Chung-Hye Han, Juntae Yoon, Nari Kim, and Martha Palmer. 2000. [A Feature-Based Lexicalized Tree Adjoining Grammar for Korean](#). Technical Report IRCS-00-04, Institute for Research in Cognitive Science, University of Pennsylvania.
- Jeong-uk Kim, Younggyun Hahm, and Key-Sun Choi. 2016. [Korean FrameNet Expansion Based on Projection of Japanese FrameNet](#). In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: System Demonstrations*, pages 175–179, Osaka, Japan. The COLING 2016 Organizing Committee.
- Young-hee Kim. 2004. A Criterion for Distinguishing between Arguments and Adjuncts in Korean. *HAN-GEUL*, 266(4):139–167.
- Martha Palmer, Daniel Gildea, and Paul Kingsbury. 2005. [The Proposition Bank: An Annotated Corpus of Semantic Roles](#). *Computational Linguistics*, 31(1):71–106.
- Chulwoo Park. 2002. Criteria for Distinguishing Complements and Adjuncts in Korean. *Korean Journal of Linguistics (EONEOHAG)*, 34(3):75–111.
- Jungyeul Park. 2006. [Extraction automatique d'une grammaire d'arbres adjoints à partir d'un corpus arboré pour le coréen](#). Ph.D. thesis, Université Paris 7 - Denis Diderot, Paris, France.
- Jungyeul Park, Daisuke Kawahara, Sadao Kurohashi, and Key-Sun Choi. 2013. [Towards Fully Lexicalized Dependency Parsing for Korean](#). In *Proceedings of the 13th International Conference on Parsing Technologies (IWPT 2013)*, pages 120–126, Nara, Japan. Association for Computational Linguistics.
- Jungyeul Park, Sejin Nam, Youngsik Kim, Younggyun Hahm, Dosam Hwang, and Key-Sun Choi. 2014. [Frame-Semantic Web : a Case Study for Korean](#). In *ISWC-PD'14: Proceedings of the 2014 International Conference on Posters & Demonstrations Track - Volume 1272*, pages 257–260, Riva del Garda, Italy. International Semantic Web Conference.
- Anoop Sarkar and Chung-Hye Han. 2002. Statistical Morphological Tagging and Parsing of Korean with an LTAG Grammar. In *Proceedings of 6th International Workshop on Tree Adjoining Grammars and Related Frameworks (TAG+6)*, pages 48–56, Venice, Italy.
- Shin Seo-In. 2019. [A study on the semantic role labels for the construction of a semantic role annotated Korean corpus](#). *Korean Semantics*, 66(1):175–204.
- Pontus Stenetorp, Goran Topić, Sampo Pyysalo, Tomoko Ohta, Jin-Dong Kim, and Jun'ichi Tsujii. 2011. [BioNLP Shared Task 2011: Supporting Resources](#). In *Proceedings of BioNLP Shared Task 2011 Workshop*, pages 112–120, Portland, Oregon, USA. Association for Computational Linguistics.

Fei Xia, Chunghye Han, Martha Palmer, and Aravind Joshi. 2000. [Comparing Lexicalized Treebank Grammars Extracted from Chinese, Korean, and English Corpora](#). In *Second Chinese Language Processing Workshop*, pages 52–59, Hong Kong, China. Association for Computational Linguistics.