A Computational Approach to Quantifying Grammaticization of English Deverbal Prepositions

Ryo Nagata, Yoshifumi Kawasaki, Naoki Otani and Hiroya Takamura
Konan University, University of Tokyo, Tokyo University of Foreign Studies, AIST
nagata-dprep @ ml.hyogo-u.ac.jp.

Abstract

This paper explores grammaticization of deverbal prepositions by a computational approach based on corpus data. Deverbal prepositions are words or phrases that are derived from a verb and that behave as a preposition such as *regarding* and *according to*. Linguistic studies have revealed important aspects of grammaticization of deverbal prepositions. This paper augments them by methods for measuring the degree of grammaticization of deverbal prepositions based on non-contextualized or contextualized word vectors. Experiments show that the methods correlates well with human judgements (as high as 0.69 in Spearman's rank correlation coefficient). Using the best-performing method, this paper further shows that the methods support previous findings in linguistics including (i) Deverbal prepositions are marginal in terms of prepositionality; and (ii) The process where verbs are grammaticized into prepositions is gradual. As a pilot study, it also conducts a diachronic analysis of grammaticization of deverbal preposition.

Keywords: deverbal prepositions, grammaticization, prepositionality, word vectors, unidirectionality

1. Introduction

Grammaticization or grammaticalization (Hopper and Traugott, 2003) is a diachronic change of the grammatical category from content words to function words. Among a wide variety of grammatical items undergoing such a change, deverbal prepositions are of current interest. They are words or phrases that are derived from a verb and that behave as a preposition (e.g., *during*, *regarding*, and *according to*, to name a few).

Linguists have provided empirical analyses of grammaticization of deverbal prepositions, including the examination of whether deverbal prepositions are really grammaticized into prepositions or not. Some researchers focus on individual items such as *following* (Olofsson, 1990) and *supposing* (Visconti, 2004). Others (Kortmann and König, 1992; Fukaya, 1997; Hayashi, 2015) analyze a wider variety of deverbal prepositions at a time. They measure the degree of grammaticization as prepositionality by either linguistic tests against instances in a corpus or human judgments of linguistic questionnaires. Both approaches examine whether target deverbal prepositions satisfy certain properties of typical prepositions such as *at* and *on*. In other words, prepositionality is defined as the degree of properties of typical prepositions that deverbal prepositions satisfy. Their findings are summarized as follows:

1. Deverbal prepositions are marginal in terms of prepositionality (Kortmann and König, 1992; Hayashi, 2015); deverbal prepositions have not been fully grammaticized into function words and they fall into a category between verbs and prepositions.
2. The process where verbs are grammaticized into prepositions is gradual (Kortmann and König, 1992; Fukaya, 1997; Hayashi, 2015); some are more like prepositions than others.
3. The categorical change from verbs to prepositions can be examined in terms of an incremental transfer from the verbal to the prepositional pole (Fukaya, 1997; Hayashi, 2015).

Although these previous studies have revealed linguistically important aspects of grammaticization of deverbal prepositions, the analyses are not yet complete. Hayashi (2015) points out that the linguistic analyses, linguistic tests or questionnaires, should be examined with a wider variety of deverbal positions (both types and tokens). It would be ideal to investigate a large number of instances appearing in a corpus. However, it would not be straightforward because the conventional ways of investigations are labor-intensive. Besides, the previous studies disagree with the prepositionality judgements in some deverbal prepositions. For example, Kortmann and König (1992) classifies the deverbal preposition *following* in one of the second lowest degree of the prepositionality while Hayashi (2015) does in the opposite.

To strengthen the previous, linguistic studies, this paper presents a first-ever computational approach to grammaticization of deverbal prepositions solely relying on corpus data. To begin with, it explores several methods for measuring the degree of grammaticization of deverbal prepositions to reveal the one that best correlates with human rankings. All methods are based on contextualized or non-contextualized word vectors. It then examines whether they support the three linguistic

LREC-COLING 2024, pages 211–220
20-25 May, 2024.

findings listed above or not. They are also used to resolve the discrepancies in prepositionality judgements in the previous studies. Finally, as a pilot study, this paper conducts a diachronic analysis of grammaticization of deverbal prepositions by using the best-performing method.

The contributions of this papers are three-fold: (i) the paper presents corpus-based, automatic methods for measuring the degree of grammaticization of deverbal prepositions through prepositionality; experiments show that the best-performing method correlates well with human judgements (as high as 0.69 in Spearman's rank correlation coefficient); (ii) the investigations using the best-performing method support all the linguistic findings, strengthening the previous, linguistic studies in terms of the number of instances examined; and (iii) the paper further shows that the method is effective in analyzing deverbal prepositions; for example, it can be used to resolve discrepancies in prepositionality judgements in the previous studies.

2. Related Work

Grammaticization is the change whereby lexical terms and constructions come in certain linguistic contexts to serve grammatical functions (Hopper and Traugott, 2003). For examples, nouns and verbs may change into grammatical items such as case makers and auxiliaries. There have been a wide variety of linguistic studies on grammaticization in general (e.g., Hopper (1991); Bybee (2006)) including computational approaches (e.g., Bentz and Buttery (2014); Morita (2019)).

Linguists have studied the diachronic development of deverbal prepositions and their pathways of their grammaticization of deverbal prepositions. Some researchers have analyzed individual items including *following* (Olofsson, 1990), *considering* (Kawabata, 2003), *notwithstanding* (Rissanen, 2002), and *supposing* (Visconti, 2004).

Other researchers (Kortmann and König, 1992; Fukaya, 1997; Hayashi, 2015) examine a wider variety of deverbal prepositions at a time. Kortmann and König (1992) analyze 19 deverbal prepositions, proposing five levels of grammaticization as prepositionality as shown in Table 1[1]. Fukaya (1997) conducts a corpus-based analysis for 14 *-ing* deverbal prepositions such as *according to* based on three linguistic tests. Hayashi (2015) proposes two linguistic questionnaires to judge the degree of the prepositionality of deverbal prepositions, which ranges from 0 to 10 as shown in Table 2; eight native speakers of English were involved in the two tests against 37 deverbal prepositions. Table 2 empirically demonstrates that the degree of grammaticization of the deverbal prepositions is gradual.

3. Methods for Measuring Prepositionality

This section describes two types of method for measuring the degree of grammaticization of deverbal prepositions as prepositionality. One is based on non-contextualized word vectors (specifically, word2vec (Mikolov et al., 2013)). There are five variants of this type. The other is based on contextualized word vectors (specifically, the output vectors of the final layer of BERT (Devlin et al., 2019)). The details of the two types are shown in Section 3.1 and Section 3.2.

A common process to the two types is the recognition of target deverbal prepositions in given corpora. A parser (specifically, spaCy[2] is used to recognize target deverbal prepositions; words or phrases satisfying the following two conditions are recognized as target deverbal prepositions: (i) the superficial form is identical to one of the deverbal prepositions shown in Table 1 or Table 2, ignoring upper/lower case; (2) its dependency label is either *prep* (i.e., preposition) or *advcl* (i.e., adverbial clause modifier). Recognized deverbal prepositions are tagged with corresponding dependency labels for the further processing below.

3.1. Non-contextualized Word Vector-based Methods

As introduced in Section 2, the previous, linguistic studies (Kortmann and König, 1992; Fukaya, 1997; Hayashi, 2015) measure the degree of grammaticization of deverbal prepositions as prepositionality. To be precise, they determine the degree by testing whether target deverbal prepositions exhibit certain properties of typical prepositions.

The non-contextualized word vector-based methods exploit this approach, but measure prepositionality directly instead of examining certain properties. Specifically, they measure the prepositionality as the similarity between the deverbal preposition in question and typical prepositions through non-contextualized word vectors. Some of them also exploit the similarity between the deverbal preposition in question and verbs considering that deverbal prepositions are losing their verbality falling into a preposition (Hayashi, 2015) (strictly, in this sense, the dissimilarity is used).

The problem now is how to calculate the similarity between prepositions/verbs and deverbal

[1] In Table 1, *ago*, which is normally classified as a postposition, is treated as a preposition following the convention in Kortmann and König (1992).

[2] spaCy 3.5.1, *en_score_web_sm* (https://spacy.io/)

Level of Prepositionality	Instances
1	facing, lining, preceding, succeeding
2	considering, failing, barring, following
3	according to, allowing for, owning to, notwithstanding
4	during, pending, except, concerning
5	past, ago, bar

Table 1: Kortmann and König (1992)'s Classification of Deverbal Prepositions according to Prepositionality: Higher levels correspond to more advanced prepositionality.

Instance (prepositional score)
past (7.8), during (7.4), following (6.4), starting (6.0), regarding (5.8), according to (5.4), preceding (5.3), succeeding (5.2), including (5.1), pertaining to (5.0), depending on (4.9), owing to (4.4), related to (3.9), given (3.9), respecting (3.9), excluding (3.8), concerning (3.7), pending (3.7), except (3.6), allowing for (3.5), lacking (3.4), barring (3.3), granting (3.3), save (3.3), touching (3.3), confronting (3.2), notwithstanding (3.1), considering (2.8), failing (2.7), covering (2.5), granted (2.3), wanting (2.3), saving (2.1), bar (1.9), facing (1.5), excepting (1.4), bating (1.3)

Table 2: Hayashi (2015)'s Prepositionality Scores of Deverbal Prepositions: Higher scores correspond to higher degrees of prepositionality.

prepositions. This paper explores the following five similarities as the five variants of the non-contextualized word vector-based method where the cosine similarity between two vectors is used as a basis.

1. **Mean preposition**
2. **Most similar preposition**
3. **Mean verb**
4. **Mean preposition and mean verb**
5. **Most similar preposition and mean verb**

Note that for deverbal prepositions consisting of more than one word, their component words are concatenated (e.g., *according_to* for *according to*) before word vectors are obtained. Also note that prepositions and verbs are tagged with their Part-Of-Speech (POS) labels in the given training corpus (pre-processed beforehand).

Mean preposition is the cosine similarity between the word vector of deverbal preposition in question and the mean vector of typical prepositions; 52 major prepositions[3] are considered in this paper. Word vectors are obtained for these 52 prepositions from a given training corpus and then their mean vector is calculated. Prepositions are recognized based on their POS label (i.e., ADP) by using spaCy. Then, the cosine similarity between the mean vector and the word vector of the deverbal preposition in question is calculated as a prepositionality score. By definition of the cosine similarity, the score (and also the following scores) ranges -1 to 1. This score evaluates how similar the deverbal preposition in question is to average prepositions. In other words, this score can be interpreted to be a similarity to typical prepositions.

Instead of taking the mean, **most similar preposition** takes the maximum value of the cosine similarities between the word vector of deverbal preposition in question and the word vectors of the 52 major prepositions. Accordingly, the score evaluates how similar the deverbal preposition in question is to typical prepositions at most.

Mean verb uses the mean vector of verbs. In this paper, the mean is taken over the 200 most frequent verbs (excluding those from which the target deverbal prepositions are derived; e.g., *consider*) in a given corpus. To be precise, occurrences of the infinitive and conjugated forms are counted and the total count is treated as the frequency of the verb in question. For this purpose, verbs are first recognized based on their POS label (i.e., VERB) in the given corpus. Then, word vectors are obtained for the infinitive and conjugated forms for the 200 verbs ignoring lower/upper case. The mean vector for verbs is calculated by taking the mean over all these word vectors. Accordingly, the mean vector reflects the conjugated forms of the 200 most frequent verbs as well as their infinitive forms. Finally, the prepositonality score is calculated based on the cosine similarity between the mean vector and the word vector of the deverbal preposition in question. Unlike **mean preposition** and **most similar preposition**, this cosine similarity corresponds to verbality. Accordingly, the negated cosine is used as a preposition-

[3] *aboard*, *about*, *above*, *across*, *after*, *against*, *along*, *amid*, *among*, *around*, *as*, *at*, *before*, *behind*, *below*, *beneath*, *beside*, *besides*, *between*, *beyond*, *by*, *despite*, *down*, *for*, *from*, *in*, *inside*, *into*, *like*, *near*, *of*, *off*, *on*, *opposite*, *outside*, *over*, *round*, *since*, *through*, *to*, *toward*, *towards*, *under*, *underneath*, *unlike*, *until*, *up*, *upon*, *via*, *with*, *within*, *without*.

ality score[4].

The last two, **mean preposition and mean verb** and **most similar preposition and mean verb** simply take the average of the two scores of **mean preposition** and **mean verb** and of **most similar preposition** and **mean verb**. These scores reflect the Fukaya (1997)'s and Hayashi (2015)'s view that deverbal prepositions are losing verbality and gaining prepositionality.

3.2. Contextualized Word Vector-based Method

Unlike the non-contextualized word vector-based methods above, this method does not compare deverbal prepositions with other words in the given corpus. Instead, the prepositionality[5] is calculated based on all contextualized vectors of the deverbal preposition in question.

The key idea for the score is that function words such as prepositions appear in a wide variety of contexts and that deverbal prepositions should follow this if they are really grammaticized into function words (or prepositions). In vector representation of words, this can be restated that contextualized word vectors for a function word point in a wide variety of directions and that those of deverbal prepositions also exhibit the similar tendency. In contrast, if a word is used in rather fixed contexts, the directions of their contextualized word vectors are concentrated in the vector space; ultimately, if a word is always used in the exact same context, the direction is always identical. The contextualized word vector-based method exploits this idea to calculate prepositionality scores. Fortunately, the concentration of vector directions can be estimated through the *von Mises-Fisher distribution* (Banerjee et al., 2005). It is a probability density function for the random d-dimensional unit vector $\mathbf{x}$. It is defined as

$$f(\mathbf{x}; \boldsymbol{\mu}, \kappa) = z_\kappa \exp\left(\kappa \boldsymbol{\mu}^\mathsf{T} \mathbf{x}\right) . \tag{1}$$

The parameters $\boldsymbol{\mu}$ ($\|\boldsymbol{\mu}\| = 1$) and κ ($\kappa \geq 0$) are respectively the mean direction and concentration parameter. The constant z_κ is the normalization constant depending on κ. It can be regarded as akin to the isotropic Gaussian distribution of the hypersphere. It is commonly used to process directional data.

What is important here is the concentration parameter κ. Under the von Mises-Fisher distribution, the unit vector $\mathbf{x}$ distributes isotropically around the mean direction with κ. In other words, κ represents how concentrated the directions of the unit vectors are, which is exactly what is desired in the key idea above. In the present case, the unit vector $\mathbf{x}$ represents contextualized word vectors for the deverbal preposition in question; note that generally, contextualized word vectors are not unit vectors, and thus they have to be normalized before applying the von Mises-Fisher distribution. Banerjee et al. (2005) show that a simple approximate solution of the maximum likelihood estimate of κ is:

$$\kappa \approx \frac{l(d - l^2)}{1 - l^2}, \tag{2}$$

where l is the norm of the mean vector of d-dimensional unit vectors.

Note that the less variety a word has in terms of the contexts in which it appears, the larger the value of κ becomes. This is the opposite to our demand that the larger variety of contexts, the larger the score is. Then, the prepositionality score is simply defined as $1/\kappa$. This score can be interpreted as a degree of contextual variety.

The procedure for score calculation is as follows:

Input: target deverbal preposition w and corpus C

Output: prepositionality score

1. Vectorize all instances of w in C by a language model
2. Normalize the contextualized word vectors so that their norms equal 1
3. Compute the mean vector of the resulting vectors
4. Estimate κ by Eq. (2)
5. Output $1/\kappa$ as score

Note that for deverbal prepositions consisting of more than one word, the vectors for their first words (e.g., *according* for *according to*) are used to compute their mean vectors.

4. Experiments

4.1. Data

The cleaned version (Alatrash et al., 2020) of COHA (Davies, 2012) is used as a corpus to obtain word vectors required in the proposed methods. COHA is a historical corpus and provides texts published between 1820 and 2019. Texts published between in the 2000s are used as a main corpus. The rest are used to conduct a diachronic analysis. Table 3 shows the sizes of the sub-corpora.

[4]One could consider most similar verbs like **most similar preposition**. Doing so, some deverbal prepositions such as *considering* and *including* have their corresponding verbs while others such as *during* and *past* do not. This brings about an unequal condition in score calculation. In a pilot study we conducted, the most similar verb strategy did not perform well. Considering this, this strategy is not included in this paper.

[5]This method does not directly consider prepositions. In this sense, strictly speaking, the score is not a prepositionality score. However, for compatibility with the non-contextualized word vector-based methods, it will be called prepositionality score.

In COHA, 5% of ten consecutive tokens every 200 are replaced by '@' due to copyright regulations. These sentences containing this special token are excluded when word vectors are obtained from the corpus.

4.2. Experimental Conditions

The 37 and 19 deverbal prepositions in the Kortmann and König (1992)'s and Hayashi (2015)'s studies are the targets in the experiments; the full lists are shown in Table 1 and Table 2 in Section 2. A frequency threshold is set to these targets; those whose frequency is equal to 30 or more in the 2000s sub-corpus are considered in the experiments, resulting in 32 and 16 deverbal prepositions, respectively.

Word vectors are obtained from the 2000s sub-corpus. As a pre-process, the above target deverbal prepositions are recognized and tagged in the sub-corpus as well as the 52 major prepositions and the 200 most frequent verbs by using spaCy with the *en_core_web_sm* model as described in Section 3.

For the non-contextualized word vector-based method, *word2vec* in the *gensim* implementation[6] is used where its hyper parameters are as follows: min count 5; vector dimension 200; window size 5, 10, or 20; the rest are set to the default values. For the contextualized word vector-based method, *bert-large-uncased* in the Hugging Face implementation[7] is used. The output vectors of its final layer corresponding to the target deverbal prepositions are used as contextualized word vectors.

Spearman's rank correlation coefficient is used as performance measure. The target deverbal prepositions, either the Hayashi (2015)'s 32 or Kortmann and König (1992)'s 16 deverbal prepositions, are ranked according to their scores given by the proposed methods and compared to the human rankings to calculate the coefficients. In other words, performance of the proposed methods is measured by how well the rankings by the proposed methods correlate with those by humans.

4.3. Experimental Results

Table 4 shows the results; the window size of word vectors are set to 10 in all non-contextualized word vector-based methods (see Table 5 for the comparison between different window sizes: 5 and 20). For reference, Table 4 also shows the Spearman's rank correlation coefficient between the Hayashi (2015)'s and Kortmann and König (1992)'s studies where 19 deverbal prepositions are involved.

Corpus	# tokens
COHA 1800s	111,048,657
COHA 1900s	262,200,025
COHA 2000s	68,678,659

Table 3: Sizes of Corpora Used to Obtain Word Vectors.

Table 4 shows that methods exploiting the information about prepositions (i.e., preposition vectors) in the score exhibit higher correlation to the human judgements. The results support the linguistic knowledge (Hayashi, 2015; Kortmann and König, 1992) that the degree of grammaticization of deverbal prepositions can be measured by their prepositionality. The values of correlation coefficient for Kortmann are much smaller than for Hayashi. This should be mainly because the Kortmann and König (1992)'s prepositionality consists only of five levels as shown in Table 1; there are many deverbal prepositions in the same rank. This is not the case in the proposed methods and nor in the Hayashi (2015)'s study. Accordingly, the low correlation is also observable in the comparison between the Kortmann and König (1992)'s and Hayashi (2015)'s judgments. To avoid this problem, five deverbal prepositions falling into different prepositionality categories are randomly chosen (specifically, *past*, *during*, *according to*, *preceding*, *considering*). Without ties, the correlation coefficients become much higher as shown in the last column of Table 4 although its p-value is not statistically significant due to a small sample size.

Table 4 also shows that while the verb vectors by themselves are not effective, they contribute to higher correlation in combination with preposition vectors. This agrees with the previous proposal (Hayashi, 2015; Fukaya, 1997) that grammaticization of deverbal prepositions should be measured considering both prepositionality and verbality. This is better exemplified by Fig. 1 where human and predicted scores are plotted for **mean preposition** and **mean preposition and mean verb** with a window size of 10. The left graph shows **mean preposition** already shows a relatively high correlation to the human judgements. The right graph further shows that deverbal prepositions in lower ranks by the human judgments (e.g., *facing*, *saving*, and *wanting*) decrease their prepositionality score when verbality is considered while those in higher ranks (e.g., *past* and *during*) keep their relative positions.

Table 5 shows the comparison between the non-contextualized word vector-based methods with different window sizes (of the word vectors). Ta-

[6] Gensim 4.3.1: `https://radimrehurek.com/gensim/models/word2vec.html`

[7] Hugging Face transformers: `https://huggingface.co/docs/transformers/model_doc/bert`

Method	Hayashi	Kortmann	Kortmann (no ties)
Mean prep.	0.37 (0.03)	0.22 (0.42)	**1.0** (0.1×10^{-23})
Most similar prep.	0.32 (0.08)	0.39 (0.14)	0.80 (0.10)
Mean verb	0.09 (0.64)	-0.15 (0.58)	-0.10 (0.87)
Mean prep. and mean verb	**0.69** (0.1×10^{-6})	0.23 (0.39)	0.70 (0.19)
Most similar prep. and mean verb	0.49 (0.4×10^{-2})	**0.41** (0.12)	0.10 (0.87)
Contextualized word vector-based method	0.14 (0.43)	-0.07 (0.81)	0.80 (0.10)
Hayashi	—	0.19 (0.47)	0.90 (0.04)

Table 4: Spearman's Rank Correlation Coefficient between Human Judgements and Proposed Methods: The numbers in brackets are p-values for the correlation coefficients. Hayashi and Kortmann denote Hayashi (2015)'s and Kortmann and König (1992)'s judgements, respectively. Kortmann (no ties) denote Kortmann and König (1992)'s judgments for only five deverbal prepositions falling into different prepositionality categories (specifically, *past*, *during*, *according to*, *preceding*, *considering*). The window size is set to 10 for all non-contextualized word vector-based methods.

Window size 5			
Method	Hayashi	Kortmann	Kortmann (no ties)
Mean prep.	0.40 (0.02)	0.21 (0.44)	1.0 (0.1×10^{-23})
Most similar prep.	0.23 (0.20)	0.24 (0.38)	0.50 (0.39)
Mean verb	0.15 (0.42)	-0.03 (0.91)	-0.10 (0.87)
Mean prep. and mean verb	0.61 (0.2×10^{-2})	0.26 (0.32)	0.70 (0.19)
Most similar prep. and mean verb	0.38 (0.03)	0.22 (0.42)	-0.20 (0.75)
Window size 20			
Method	Hayashi	Kortmann	Kortmann (no ties)
Mean prep.	0.37 (0.03)	0.25 (0.35)	1.0 (0.1×10^{-23})
Most similar prep.	0.28 (0.12)	0.44 (0.09)	0.80 (0.10)
Mean verb	-0.03 (0.87)	-0.31 (0.24)	-0.30 (0.62)
Mean prep. and mean verb	0.63 (0.1×10^{-3})	0.11 (0.68)	0.70 (0.19)
Most similar prep. and mean verb	0.50 (0.3×10^{-2})	0.22 (0.42)	0.20 (0.75)

Table 5: Comparison between Non-contextualized Word Vector-based Methods with Different Word Window Size (5 and 10): Performance is measured by Spearman's Rank Correlation Coefficient where the numbers in brackets are p-values. Hayashi and Kortmann denote Hayashi (2015)'s and Kortmann and König (1992)'s judgements, respectively. Kortmann (no ties) denote Kortmann and König (1992)'s judgments for only five deverbal prepositions falling into different prepositionality categories (specifically, *past*, *during*, *according to*, *preceding*, *considering*).

ble 5 confirms all settings exhibit a similar tendency to Table 4.

Unlike its counterparts, the contextualized word vector-based does not correlate with the human judgements well. This might be because deverbal prepositions are marginal in terms of prepositionality as Hayashi (2015) points out. In other words, they have not been fully grammaticized yet and have less variety of contexts than function words. One thing we should emphasize here is that it does not exploit the information about typical prepositions nor verbs directly and that it is only based on the concentration parameter κ, which is in turn based on the norm of the mean word vector of the deverbal preposition in question. Nevertheless, it shows a mild correlation to the human judgements, especially to the Kortmann and König (1992)'s judgments with no ties. This suggests that the contextualized word vector-based method might be useful in quantifying grammaticization in general.

5. Discussion

The experimental results in Section 4.3 reveal that the methods based on non-contextualized word vectors are useful for measuring the degree of grammaticization of deverbal prepositions. They have strengthened the previous results in linguistics in a purely corpus-driven way.

They will likely facilitate analyzing deverbal prepositions as will be explored below; the best-performing method (**mean preposition and mean verb** with a window size of 10) will be used in the analyses below, which will be referred to as the proposed method hereafter in this section.

An example would be resolving the discrepancies in prepositionality judgements in linguistics. Specifically, this paper examines the five deverbal prepositions, shown in Table 6, falling in different degrees of prepositionality according to the Ko-

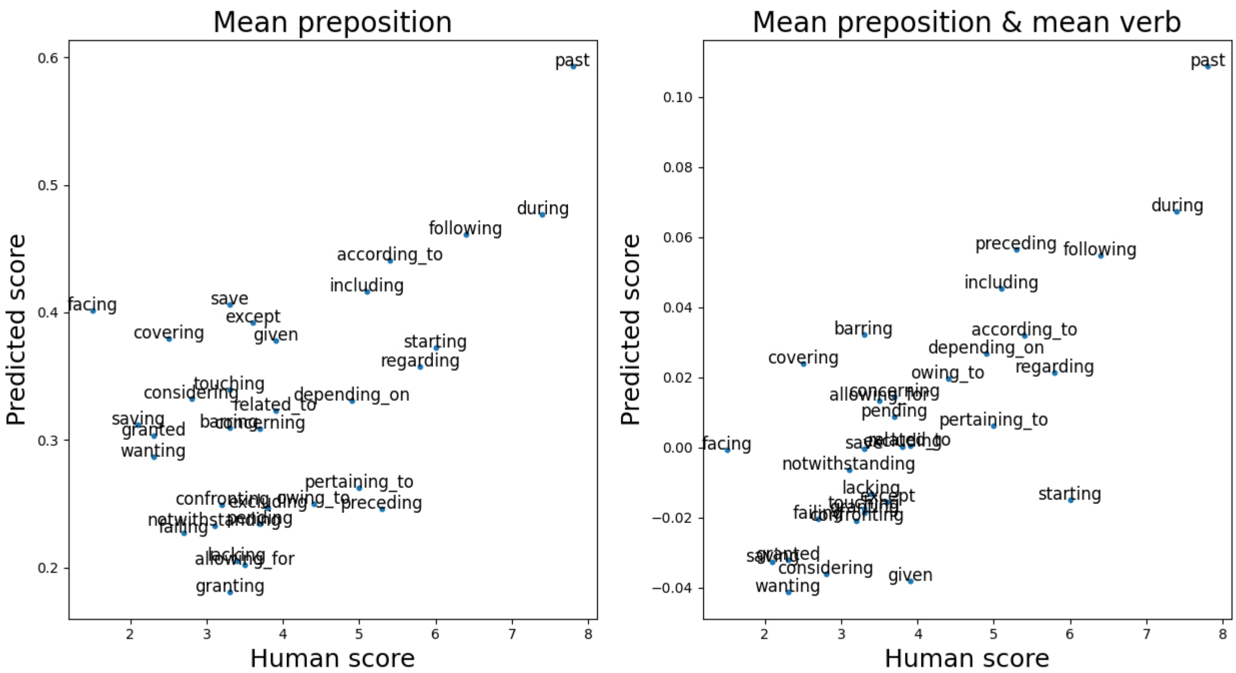

Figure 1: Relationships between Hayashi (2015)'s Judgements and Proposed Methods (Mean preposition (left)/mean preposition and mean verb (right), both window size of 10).

rtmann and König (1992)'s and Hayashi (2015)'s studies.

The largest difference between the two linguistic studies is in *following*. It falls into the second lowest degree of prepositionality in the Kortmann and König (1992)'s study while it is one of the most prepositional ones ranking in third after *past* and *during* in the Hayashi (2015)'s study. The proposed method favors the latter, ranking it in third after *past* and *during* just like in the Hayashi (2015)'s study. One can make a similar argument for *according to*. The proposed method favors the judgement in the Hayashi (2015)'s study. The rest (*pending*, *concerning*, and *except*) fall into the second most prepositional category according to Kortmann and König (1992). In contrast, as shown in Table 6, their scores in Hayashi (2015) are smaller than that of *according to* that falls into the third most category in the former. Namely, the two linguistic studies show a reversal of ranks in those two groups of deverbal prepositions. The proposed method again favors the Hayashi (2015)'s judgements in all three. Besides, it gives the lowest score to *except*, suggesting that *except* is the least prepositional among the three. To sum up, the proposed method (or abstractly speaking, corpus data) favors the judgments of the Hayashi (2015)'s study.

As another application of the proposed method, we now examine the linguistic knowledge that deverbal prepositions are marginal in terms of prepositionality (Hayashi, 2015; Kortmann and König, 1992). To achieve it, the prepositionality scores for the 52 major prepositions are calculated by the proposed method and their distribution is de-

Deverbal preposition	Kortmann	Hayashi
following	2	6.4
according to	3	5.4
pending	4	3.7
concerning	4	3.7
except	4	3.6

Table 6: Prepositionality Scores Measured in Linguistics: Kortmann and Hayashi denote the Kortmann and König (1992)'s and Hayashi (2015)'s scores, respectively. The higher the scores, the more prepositional the targets are.

picted as a histogram together with that of the 32 deverbal prepositions. Note that the preposition in question is left out from the calculation of the mean preposition vector when its score is calculated (i.e., leave-one-out estimate).

Fig. 2 shows the histogram. Fig. 2 reveals that as expected, the distribution for the deverbal prepositions appears in a lower range than that of the typical prepositions. This agrees with the linguistic knowledge. Fig. 2 also shows that a few of the deverbal prepositions, which are *past* and *during*, exhibit as a high score as the typical prepositions. This suggests that some deverbal prepositions such as the two are grammaticized as well as typical prepositions.

The same method can also be used to conduct a diachronic change of deverbal prepositions. Namely, by using the proposed method, we can examine how the degree of the prepositionality of each deverbal preposition varies diachronically. For this purpose, we calculate prepositionality scores from the texts published in each decade in COHA. Then, the results are depicted as a line-

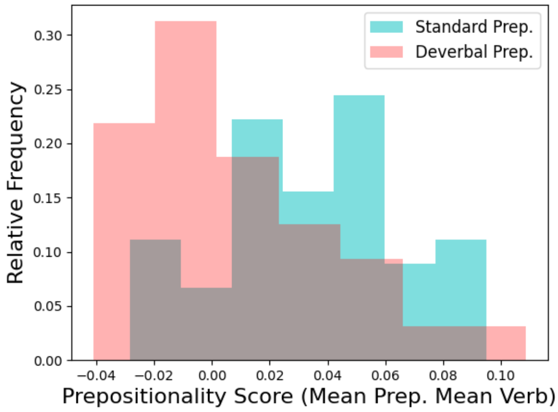

Figure 2: Distributions of Prepositionality Score for Typical and Deverbal Prepositions.

plot graph to reveal their changes. Here, we only target deverbal prepositions appearing 30 times or more in all the decades.

Fig. 3 shows the graph where the horizontal and vertical axes correspond to decades of publication and prepositionality scores. For visibility, the deverbal prepositions are classified into four groups according to the Hayashi (2015)'s prepositionality score; the line for the highest score group (*past*, *during*, *starting*, and *regarding*) is shifted by +0.2×3 along the vertical axis, and then the second by +0.2×2, and so on to avoid too many overlaps of lines.

Fig. 3 shows that some deverbal prepositions in the highest score group (e.g., *past* and *during*) exhibit a mild increase in their prepositonality score. The others exhibit rather globally flat lines. These observations might have something to do with the unidirectionality hypothesis (Hopper and Traugott, 2003) that grammaticization goes only in one direction from content to function words. Having said that, from the results available here, we cannot tell whether the increase in the first group is just a coincidence or a rather observed noise; the increase of prepositionality in other groups is too subtle to be detected; or there are some other explanations for the forms of the lines in Fig. 3.

It would be interesting to investigate further using a wider time range of texts such as Early English Books Online[8] and Eighteenth Century Collections Online[9]. It would be also interesting to use other methods such as the contextualized word vector-based method to draw a time-prepositionality relation.

6. Limitations

One thing that should be noted is that accuracy of recognizing target deverbal prepositions crucially influences performance of the proposed methods. If target deverbal prepositions are not recognized

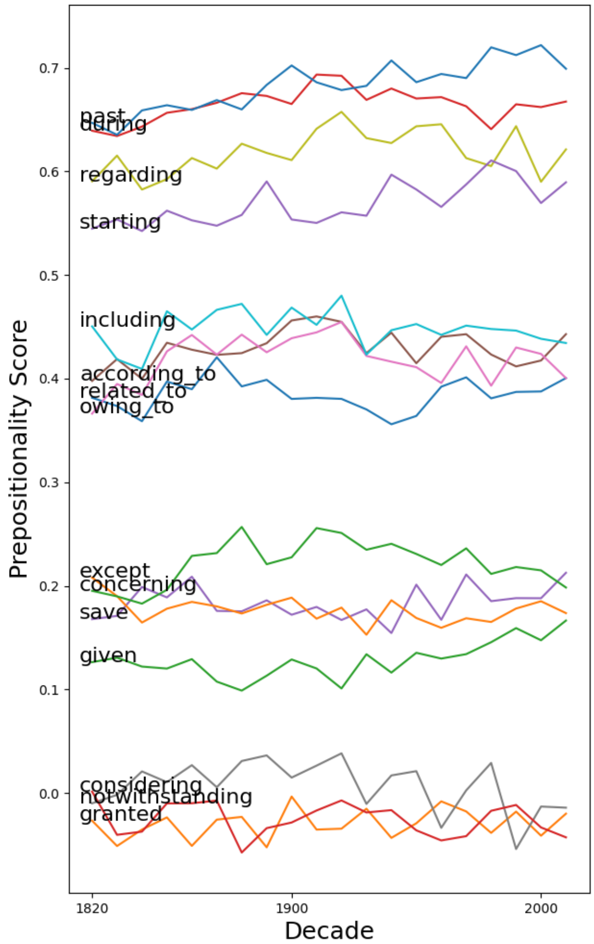

Figure 3: Relationship between Time of Publication and Prepositionality Scores for Unidirectionality Hypothesis for Deverbal Prepositions.

by a parser, they are not considered in the proposed method in the first place.

To make sure that target prepositions were accurately recognized by the parser (i.e., spaCy), we checked its performance for several instances of all target deverbal prepositions. It turned out that the instances were successfully recognized in most of the cases. An exception was the deverbal preposition *save*. Its instances as a deverbal preposition were highly rare in the training corpus. While the word *save* was used as a verb most of the time, spaCy sometimes recognized it as a deverbal preposition. This result suggests that the prepositionality scores estimated by the proposed methods might be unreliable for the deverbal preposition *save* although overall performance (correlation to human) remains almost the same with and without *save*. Generally, the user should be aware of the fact that performance of target recognition influences results obtained by the proposed methods.

The quality of word vectors is another key point in the proposed methods. As explained in Section 3, deverbal prepositions are compared to typ-

[8] https://quod.lib.umich.edu/e/eebogroup/
[9] https://quod.lib.umich.edu/e/ecco/

ical prepositions represented by their mean word vector. This means that it is crucial that their mean word vector should reflect the properties of typical prepositions.

For a sanity check, we retrieved most similar words based on the similarities between each typical preposition and all words in the training corpus. It turned out that 76% of the typical prepositions had at least one of the other typical prepositions within the top three most similar words. Some had other than prepositions such as *difference*, *differences*, and *relationship* for the preposition *between*. These words often co-occur with the preposition in question and the window size is rather large (specifically, ten). As a result, their word embeddings tend to be similar to the preposition in question although these are not major cases. Similarly, we retrieved similar words for the mean vector of the typical prepositions. As a result, 23 (77%) out of the 30 most similar words were typical prepositions (one was the deverbal preposition). The other words include *when*, *where*, and *back*, which describe temporal- or spatial-relations. These results show that the mean word vector represents the typical prepositions at least to some extent. At the same time, we might be able to improve the quality by removing prepositions such as *between* whose most similar words are not prepositions. We can also adjust the hyper-parameters as the window size according to the same criterion. These adaptations might result in a higher correlation with the human judgements.

7. Conclusions

This paper has explored grammaticization of deverbal prepositions based on word vectors. The experiments have shown that the method based on mean vectors for prepositions and verbs (both non-contextualized) exhibits the highest correlation coefficient for human judgments in linguistic studies.

Using the method, this paper has further examined the following three findings known in linguistics:

1. Deverbal prepositions are marginal in terms of prepositionality.
2. The process where verbs are grammaticized into prepositions is gradual.
3. The categorical change from verbs to prepositions can be examined in terms of an incremental transfer from the verbal to the prepositional pole.

It turns out that the method supports all of the three, which augments the previous, linguistic studies in that the number of instances examined are much larger.

This paper has also discussed the applications of the method to grammaticization of deverbal prepositions. It has shown that the proposed method favors the results of Hayashi (2015)'s prepositionality judgements over the Kortmann and König (1992)'s. It has also conducted a diachronic analysis concerning grammaticization of deverbal prepositions, proposing possible explanations for the results. The investigation is still in an early stage and requires further explorations.

Acknowledgements

This work was partly supported by JSPS KAKENHI Grant Number JP23K12152.

8. Bibliographical References

Arindam Banerjee, Inderjit S. Dhillon, Joydeep Ghosh, and Suvrit Sra. 2005. Clustering on the unit hypersphere using von Mises-Fisher distributions. *Journal of Machine Learning Research*, 6(46):1345–1382.

Christian Bentz and Paula Buttery. 2014. Towards a computational model of grammaticalization and lexical diversity. In *Proceedings of the 5th Workshop on Cognitive Aspects of Computational Language Learning (CogACLL)*, pages 38–42, Gothenburg, Sweden.

Joan Bybee. 2006. From usage to grammar: The mind's response to repetition. *Language*, 82:711–733.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Teruhiko Fukaya. 1997. *The Emergence of -ing Prepositions in English: A Corpus-Based Study*, pages 285–300. Taishukan, Tokyo.

Tomoaki Hayashi. 2015. Prepositionalities of deverbal prepositions: Differences in degree of grammaticalization. *Papers in Linguistic Science*, 21:129–151.

Paul Hopper. 1991. *On some Principles of Grammaticalization*, pages 17–35. John Benjamins Publishing Company.

Paul J. Hopper and Elizabeth Closs Traugott. 2003. *Grammaticalization*, second edition. Cambridge University Press, New York.

Tomohiro Kawabata. 2003. On the development of *Considering*: The prepositional conjunctive and adverbial usages. *Studies in Modern English (The Twentieth Anniversary Publication of the Modern English Association)*, pages 138–152.

Bernd Kortmann and Ekkehard König. 1992. Categorial reanalysis: The case of deverbal prepositions. *Linguistics*, 30:671–698.

Christian Mair. 2004. Corpus linguistics and grammaticalisation theory: Statistics, frequencies, and beyond. *Corpus Approaches to Grammaticalization in English*, pages 121–150.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems 26*, pages 3111–3119.

Junya Morita. 2019. Grammaticalization in derivational morphology: Verification of the process by innovative derivatives. In *Proceedings of the Second International Workshop on Resources and Tools for Derivational Morphology*, pages 121–130, Prague, Czechia.

Arne Olofsson. 1990. A participle caught in the act. On the prepositional use of following. *Studia Neophilologica*, 62(1):23–35.

Matti Rissanen. 2002. *Despite or notwithstanding? On the development of concessive prepositions in English*, pages 191–203. Gunter Narr Verlag, Germany.

Jacqueline Visconti. 2004. *Conditionals and subjectification: Implications for a theory of semantic change*, pages 169–192. John Benjamins Publishing Company, Amsterdam/Philadelphia.

9. Language Resource References

Reem Alatrash, Dominik Schlechtweg, Jonas Kuhn, and Sabine Schulte im Walde. 2020. CCOHA: Clean corpus of historical American English. In *Proc. of the 12th Language Resources and Evaluation Conference*, pages 6958–6966.

Mark Davies. 2012. Expanding horizons in historical linguistics with the 400-million word corpus of historical American English. *Corpora*, 7(2):121–157.