# 「論文」

# Cues to identifying support verb constructions: A corpus-based study of [Verb + EXAMINATION]

Taishi CHIKA and Kazuho KAMBARA

# Abstract

This paper describes a corpus survey of [Verb + EXAMINATION] to explore the formal specifications of support verb constructions (SVCs). Previous studies described the properties of SVCs by focusing on the apparent semantic and syntactic paucity of support verbs (Brugman, 2001; Newman, 1996; Wierzbicka, 1982) and the process of argument transfer triggered by the complements (Grimshaw & Mester, 1998; Grimshaw, 1990). However, these verb-centered approaches face the issue of ambiguity between light and heavy senses in context (e.g., *make<sub>light</sub> a diagnosis, make<sub>heavy</sub> a certificate*), and coverage of low frequency support verbs (e.g., *sustain an injury*). To address these issues, we point out the need for formal specifications of SVCs targeted on their compliments. Our corpus analysis of [Verb + EXAMINATION], in which the deverbal noun *examination* possesses its own argument structure, revealed the types of verbs preferred in SVCs and the grammatical properties of *examination* (e.g., the occurrence of an *of*-phrase).

# 1. Introduction

We typically have the following options when referring to the event of "inspecting someone or something to determine their nature or condition, or testing someone's knowledge or proficiency by requiring them to answer questions or perform tasks" (*New Oxford American Dictionary*, s.v. *examine*, v, 3rd ed.):

- (1) a. A doctor *examined* me and said I might need a cesarean.
  - b. The colleges examined candidates.
- (2) a. A doctor made an *examination* of the need of a cesarean.
  - b. The colleges conducted an *examination* of candidates.

From a syntagmatic viewpoint, both examples (1–2) appear to realize the standard transitive verb construction (i.e., [Verb + Obj]). However, in (1), the "examining" event is expressed as the main verb *examine*, whereas in (2), the object noun "examination" (rather than the verb "made" or "conducted") signifies its event. In (2a), the verb *made* taking *examination* as its complement does not literally refer to any actual process of creation (or "making"). Rather, the lexical meaning of the main verb becomes bleached. Traditionally, such verbs are referred to as **light verbs** (Jespersen, 1940; Wierzbicka, 1982) or **support verbs** (Fillmore et al., 2003; Fujii & Uegaki, 2008).

**Support verb constructions (SVCs)**, also known as **light verb constructions** (LVCs)<sup>\*1</sup>, constitute a subclass of transitive verb constructions (Fujii & Uegaki, 2008). Previous studies have focused on the behavior of typical support verbs (e.g., *do, make, have*), although there are several explanations that differ in the extent to which verbs contribute to the semantic and syntactic properties of SVCs. This approach, which we refer to as the **verb-centered approach**, focuses on typical support verbs and aims to figure out the process of argument transfer (Grimshaw & Mester, 1988; Grimshaw, 1990), the syntactic and semantic contribution of support verbs (Brugman, 2001; Newman, 1996; Wierzbicka, 1982), and the collocation (Giparaite, 2023).

However, two issues must be addressed to elucidate the linguistic knowledge that allows speakers to use SVCs. First, "How do we differentiate light senses of verbs from heavy ones?" The verb-centered approach often assumes *a priori* that the verbs in question are support verbs. However, given that verbs used as support verbs can also function as regular transitive verbs, which Brugman (2001) calls using them in a heavy sense, they possess inherent ambiguity. Second, some lexical items function as support verbs only in combination with specific complements (e.g., *sustain an injury*), so when investigating them, there is a risk of excluding verbs that are less frequently used as support verbs.

To address these issues of ambiguity and coverage, verbs should not be described as distinct lexical items but as parts of constructions—conventionalized associations of meaning and form (Goldberg, 2006; Taylor, 2012; Hoffmann, 2022). Support verbs can then be detected by analyzing their complements (cf. Langer, 2005). For a verb to occur in an SVC as a constructional unit, an analysis of semantic properties is needed along with investigation of formal characteristics to differentiate SVCs from regular transitive verb constructions within the configuration of constructions, as in Figure 1.



Figure 1. The configuration of regular transitive verb constructions and support verb constructions

In this paper, we demonstrate the need for formal specifications of SVCs, focusing on complements, particularly event nouns that possess their own argument structure (e.g, *examination*), to address the issues inherent in the verb-centered approach. Specifically, we try to answer the following questions: (1) Does association strength (collostructional strength) impact the likeliihood of a verb to be considered a light or heavy verb?

(2) Does the presence of an *of*-phrase following the complement exhibit relatively strong predictive power for SVCs?

## 2. SVCs as a construction

Section 2 provides an overview of previous research on SVCs and points out the methodological and empirical issues inherent in the verb-centered approach. We highlight the need for cues to distinguish verbs with light senses from those with heavy senses, and SVCs from regular transitive verb constructions.

#### 2. 1. Verb-centered approaches to SVCs

In traditional English grammar, support verbs are supposed to lack independent meaning, with their constructions expressing events through the complement (Jespersen, 1940). SVCs are not only observed in English, but widespread among several languages (e.g., Japanese, Korean, German, and Russian). The meaning of the verb phrase relies heavily on the nominal complement, which also determines the argument structure of the verb (cf. Grimshaw & Mester, 1988; Grimshaw, 1990). A representative study of SVCs by Grimshaw and Mester (1988) discussed the Japanese support verb *suru* (trans.: "do"), as in *shuppatsu-o suru* (trans.: "departure"). The authors analyzed

*suru* as lacking  $\theta$ -role assignment capacity and therefore having an incomplete argument structure, proposing the process of argument transfer through which the support verb inherits the argument structure of its complement, enabling it to function within the construction (Grimshaw & Mester, 1988).

Nevertheless, such a process alone fails to explain the preference for certain combinations (e.g., *have a drink*, *\*have an eat*) and misses the subtle nuances imposed on SVCs but lacking from their regular transitive construction counterparts. Cognitive linguists adopt the gradient view that the light and heavy sense of a particular verb are not strictly categorical (e.g., Brugman, 2001; Newman, 1996; Wierzbicka, 1982). This perspective highlights the fuzzy nature of support verbs, supported by the constraints on several properties support verbs can take such as manner of action, aspect, and valency. For instance, Wierzbicka (1982) pointed out that while it is possible to say "have a drink," meaning *to drink something*, it sounds odd or unacceptable to say "have a study" (or *work*, or *practice*) to express the action in question. Based on those observations, she postulates a prototypical condition under which the support verb *have* is applicable to SVCs, namely as an "AIMLESS OBJECTLESS INDIVIDUAL ACTIVI-TY WHICH COULD CAUSE ONE TO FEEL GOOD" (Wierzbicka, 1982, p. 762).

In addition, support verbs can impose aspectual constraints on events. For instance, the verb *shower* in (3a) lacks a specific endpoint (i.e., atelic), making it incompatible with adverbial phrases like [in + TIME], which indicate the completion of an event within a limited duration of time (i.e., telic). However, Brugman (2001, p. 556) reports that when the verb *shower* is replaced with the SVC *take a shower*, the sentence in question becomes acceptable.

(3) a. Ashley showered { for / in } 10 minutes.

b. Ashley took a shower { #for / in } 10 minutes.

Furthermore, some support verbs retain the valency of their heavy-sense counterparts. Newman (1996) pointed out that SVCs headed by the verb *give* typically demand a dative phrase (*to NP*), just as the heavy sense of *give* does. This observation calls into question Grimshaw and Mester's (1988) view that support verbs do not possess an independent argument structure.

(4) a. <sup>#</sup>John gave a presentation.

b. John gave a presentation to his students.

In the following discussion, we refer to the methodology applied in the studies

reviewed in this section as the "verb-centered approach" because the explanations rely heavily on the properties of support verbs in SVCs.

## 2.2. SVCs as constructional units

As seen in Section 2.1., verb-centered approaches take the inventory of support verbs as granted. The following problems are inherent in verb-centered approaches: (i) the ambiguity between "light" uses of verbs (e.g., "Alice <u>made</u> an appointment") and "heavy" uses of verbs (e.g., "Alice <u>made</u> a breakfast"), and (ii) a relatively low coverage of SVCs. Previous studies have not addressed these issues.

Analysts must differentiate between light and heavy sense in context to explicitly describe the linguistic knowledge that allows speakers to use SVCs. It is worth noting that most verbs used as support verbs in SVCs can also appear in regular transitive verb constructions, which Brugman (2001) calls the *heavy* sense.

(5) a. A doctor *made*<sub>*light*</sub> an early diagnosis.

b. A doctor *made*<sub>heavy</sub> a medical certificate.

- (6) a. Alice *sustained*<sub>light</sub> injury.
  - b. Alice *suffered*<sub>light</sub> a loss.
  - c. Alice wage<sub>light</sub> war.

(cf. Fillmore et al., 2002, p. 790)

In (5a), the verb *made* functions as a support verb, inheriting the semantics and argument structure of the complement *diagnosis*. In contrast, in (5b), *made* conveys a heavy sense (literally "make"). Verbs in (6) are examples of lexical items that are typically interpreted as having a heavy sense but peculiarly function as support verbs with a very limited set of complements. When determining whether a given verb in [Verb + Obj] is used as a support verb or a regular transitive verb, and when extending the scope of investigation beyond common support verbs, we must, at the very least, refer to its complement. The verb-centered approach often assumes *a priori* that the verbs in question are support verbs. However, this approach ignores the potential ambiguity of verbs, thus posing the risk of excluding verbs that are less frequently used as support verbs from the scope of investigation. To address this issue, verbs should not be reduced to distinct lexical items but rather described as parts of constructions—conventionalized associations of meaning and form (Goldberg, 2006; Taylor, 2012; Hoffmann, 2022).

As we have seen, Brugman (2001) emphasizes the continuity between support verbs and heavy-sense verbs in terms of semantic contribution, and seems reluctant to establish a clear delimitation between SVCs and regular transitive verb constructions. Interestingly, she makes the following remarks regarding the characterization of SVCs as constructions.

There are certainly reasons to talk about an LVC (taking even the verb head as variable rather than specified), given the common semantic relationships associated with schema extraction and their consequent properties — we can say with some assurance that, fuzzy as they are, there are some Aktionsart properties common to all LVCs by contrast with their monomorphemic paraphrases. (Brugman, 2001, p. 576)

Following this approach, it is possible to postulate a set of constructions subsumed under the support verb construction, as shown in Figure 2. In the configuration, the subschema (e.g., [take + OBJ]) inherits the abstract specification of superschema (SVC) and elaborates the semantic constraints (e.g., aspect, manner of action) that each support verb imposes on their complements. Linguistic knowledge is constructed in a bottom-up fashion and conceptualized as an extensive inventory of actual usage patterns (Taylor, 2012) when adopting the usage-based model (Langacker, 2000), a view that aligns closely with the principles of construction grammar (Hoffmann, 2022).

If speakers' linguistic knowledge of SVCs constitutes an inheritance structure (as in Figure 2), verb-centered approaches only deal with a handful of SVC subclasses. Analyzing SVCs with "major" support verbs could lead to misguided generalizations. To address this issue, analysts should treat the verb slots as variables rather than constants (cf. Uchida, 2010). In this way, they should be able to observe the ambiguity of light and heavy senses of collocating verbs and mediate the coverage.



To support this claim, we extracted instances from a corpus where the deverbal noun examination, which has its own argument structure, functions as the complement of a verb, conducting both quantitative and qualitative analyses of  $[Verb + EXAMINA-TION]^{*2}$ . Through the analyses, we found at least two cues to identify the SVCs in this syntactic environment: (i) verbs with high association strength (collostructional strength) and (ii) an *of*-phrase following EXAMINATION (e.g., "the examination of old age and society").

## 3. Methods

To observe the formal environment of SVCs, we used Sketch Engine (Kilgarriff et al., 2004, 2014) to extract all instances of the noun *examination* that occurred as an object of a verb. The initial query yielded 1,117 cases. We excluded cases where the noun *examination* was not realized as an object of a verb, resulting in 1,036 cases and a type frequency of 231.

We annotated the instantiation of SVCs based on the realizations of relevant semantic roles<sup>\*3</sup>. Since characterizing SVCs can be challenging, we employed the manageable semantic role-based characterization of SVCs as defined in (7), similar to that employed in FrameNet (Fillmore et al., 2002, 2003). We also defined EXAMIN-ING as in (8) to annotate the distribution of *examination*-related semantic roles.

- (7) Semantic role-based characterization of SVCs: The construction containing *examination* is an instance of SVCs if and only if the noun in the subject position realizes at least one semantic role of EXAMINING regardless of the collocating verbs (cf., "Alice {passed, conducted} an examination").
- (8) EXAMINING: <Examiner> assesses the <Attribute> of <Examinee>
  - a. [<Examinee> Alice] passed an examination.
  - b. [<Examiner> Alice] conducted an examination.

Then, for each case, we annotated the following formal features to identify the formal environment that SVCs prefer. First, we annotated the voice of the construction containing *examination* as is\_passive (9a). Most transitive constructions can be realized in passive or active voice, making them candidates for the crucial formal SVC environment. We also annotated the realization forms of noun phrases by coding the grammatical number as noun\_is\_singular (9b) and the presence of any article (9c).

- (9) Formal environments of SVCs:
  - a. is\_passive: 1 iff the case in question is realized in passive voice, 0 otherwise.
  - b. noun\_is\_singular: 1 iff the noun examination is realized as a singular noun, 0 otherwise.
  - c. **has\_article:** 1 iff the noun *examination* co-occurs with an (indefinite or definite) article, 0 otherwise.
  - d. **coll\_strength:** Pearson residuals between expected and observed frequency of verbs.

To compute the collocational strength, we performed collostructional analysis using Coll.analysis 4.1. (Gries, 2024). The term "collostructional analysis" refers to a family of collocational analyses that can accurately capture the collocational strength between grammatical constructions and words (Stefanowitsch & Gries, 2003, 2005; Gries & Stefanowitsch, 2004a, b; Gries, 2019, 2023). While many association measures are currently available in corpus linguistics, employing Pearson residuals was proposed to measure the degrees of collocational preference (Gries, 2023). Pearson residuals refer to the difference between observed and expected frequency in the form of a cross-tabulation table. Using this approach, analysts can capture the words' preference (or repulsion) in the construction. Pearson residuals were employed to determine collostructional strength.

We performed logistic regression analysis (Gries, 2021; Levshina, 2015; Speelman, 2014) to explore the extent that predictors in (9) contribute to discriminating SVCs from non-SVCs. Logistic regression analysis is a type of regression analysis using categorical response variables (i.e., every sentence in question is either an SVC or not). Regression analysis can reveal differences in the data and predict the variables contributing to the distribution of response variables. Performing logistic regression analysis allows analysts to determine the contribution of predictors in classifying constructions.

Moreover, regression analysis provides a formula that predicts the distribution of its response variable (i.e., isSVC). This allows analysts to compute the extent of correct data predictions using a confusion matrix consisting of the frequency of predicted and actual instances. For instance, the contingency table shown in Table 1 reveals 20 misclassified items (10 false positives and 10 false negatives). Analysis of a confusion

matrix can be used to evaluate the performance of a constructed model.

All annotations were carried out by the authors. We used R (R Core Team, 2024) to perform the computation and a family of ggplot2 to visualize the results (Wickam et al., 2024).

Table 1. A fictitious distribution of predicted and actual frequency of isSVC

	isSVC (Predicted)	¬ isSVC (Predicted)
isSVC (Actual)	90	10
¬ isSVC (Actual)	10	90

# 4. Results

This section reports the result of our corpus study. As a result, we revealed that the interpretation of SVC is likely to be realized when the verb slot of SVC is filled with verbs with high association strength. We report quantitative and qualitative results of our study. All the codes and data used in this study are available on the Open Science Framework (OSS).

## 4.1. Descriptive statistics

#### 4.1.1. Types of verbs

Of 231 verb types, 99 were realized as SVCs and 164 types were not. While the realization of SVC was not mutually exclusive in some verbs (e.g., allow), token



Figure 3. Distribution of raw frequency and isSVC

frequency of SVC was higher than the transitive uses (759 cases realized as SVC and 277 cases realized as non-SVC). Figure 3 shows the distribution of logged frequency on the y-axis and the types of constructions in which verbs occur.

Given that the raw frequencies do not necessarily convey accurate association (Kambara & Chika, 2023; Kambara et al., in press), the association strength of verbs and EXAMINATION was computed using Pearson residuals as collostructional strength. Figure 4 shows the distribution of Pearson residuals in SVC and non-SVC constructions with the boxplot under each collocate. The height of the boxplot in Figure 4 suggests that the distribution of Pearson residuals is spread more widely in non-SVCs because that type includes low frequency verbs. In contrast, values of Pearson residuals in SVCs show a relatively more even spread across the y-axis. This result suggests that a high association strength between the noun and verbs is a strong cue for identifying an SVC.

We can deduce that lexical items functioning as support verbs, occurring in SVCs, tend to exhibit relatively high association strength. Additionally, verbs used as heavy sense, occurring in regular transitive verb constructions, generally show low collostructional strength with EXAMINATION, except for a few outliers.



Figure 4. The distribution of Pearson residuals in SVCs

## 4.1.2. Grammatical features of EXAMINATION

In addition to the types of verbs, we also recorded the morpho-syntactic environment of the noun *examination*, focusing on three variables: has\_article, noun\_is\_ singular, and collocates\_with\_of. The raw frequencies of these three variables are summarized as cross-tabulations in Table 2 and visualized as three distinctive mosaic plots in Figure 5.

	isSVC	¬ isSVC
has_article	559	176
-has_article	200	101
noun_is_singular	533	146
¬noun_is_singular	226	131
collocates_with_of	175	1
-collocates_with_of	584	276

Table 2. Raw frequency of each morpho-syntactic variable



Figure 5. Morpho-syntactic environments of examination

In the mosaic plots, the widths of the bars represent the proportional distribution of the variable on the x-axis and within each of the (stacked) bars, the heights indicate the proportional distribution of the levels of the variable on the y-axis (Gries 2021, p. 123). The plot indicates two types of features that most instances of *examination* are likely to realize: (i) as singular and with an article, and (ii) the presence of an *of*-phrase as shown in (10).

- (10) a. [...] his research team began the electrical examination of acupuncture points of human system [...] [CB9 1436]
  - b. This project conducts an examination of old age and society between 1918 and 1948 [...] (HJ0 3670)

## 4.2. Inferential statistics

Based on the descriptive statistics, we constructed a statistical model using a logistic regression analysis using collocates\_with\_of and collostructional strength. The

effects of each predictor are shown in Figure 6. The constructed binomial model, represented in Table 3, was statistically significant with each of the predictors significantly contributing to the discrimination of the presence of SVC. The plot and table show the presence of the preposition *of* and collostructional strength contribute to the realization of SVCs.

A notable finding from Table 3 is that combinations where an *of*-phrase follows almost certainly instantiate SVCs. In contrast, at least within the scope of the features annotated in this paper, other grammatical features were not interpreted as crucial factors in determining whether [Verb + EXAMINATION] instantiates SVCs or regular transitive verb constructions, that is, whether the verb in question is a support verb or not.

Table 4 shows the confusion matrix of isSVC. The accuracy of the model was 83.5%. Given that the SVC proportion was high in the observed data, we opted to set the baseline by computing the proportion of maximum value, which was 74% (= 0.7326). The accuracy exceeded the baseline, and we concluded that the model made a fairly "good" classification. Nagelkerke's Pseudo R<sup>2</sup> was calculated at 0.548, suggesting that the constructed model was partly effective. The C score was 0.917, showing outstanding discrimination of the model (Levshina, 2015, p. 256). These results suggest that SVC classification is a fairly easy task when collostructional strength is taken into account, and that the presence of the preposition *of* can help



Figure 6. Effect plot of formal predictors

listeners readily identify the construction.

	Estimate	Std. Error	z value	$\Pr(> z )$	
(Intercept)	-1.077	0.132	-8.1250	0.000	***
collocates_with_of TRUE	5.248	1.009	5.200	0.000	***
coll_strength	0.050	0.004	12.525	0.000	***

Table 3. Coefficient table of the constructed model

Table 4. Confusion matrix of isSVC				
	isSVC (Predicted)	¬ isSVC (Predicted)		
isSVC (Actual)	647	112		
¬ isSVC (Actual)	59	218		

## 5. Discussion

The overall findings of the result in Section 4 can be summarized as follows:

- (11) Overall findings:
  - a. When verbs with high collostructional strength collocate with EXAMI-NATION, the verb phrase [Verb + EXAMINATION] generally instantiates SVCs (Figure 4).
  - b. Among the grammatical properties of *examination*, the occurrence of an *of*-phrase is the strongest predictor for [Verb + EXAMINATION] instantiating SVCs (the left panel of Figure 6).
  - c. Based on findings (11a-b), it is relatively straightforward to predict which constructions [Verb + EXAMINATION] instantiate (the right panel of Figure 6, Table 4).

In this section, we discuss the qualitative results and the implications of describing SVCs as a type of constructional unit, based on our findings.

Regarding (11a), it is possible that verbs with high collostructional strength form

a cluster, sharing certain semantic features and providing information such as [Verb<sub>i</sub> + EXAMINATION]  $\leftrightarrow$  <examining in the manner/aspect specified by  $i > *^4$ . Given that SVCs are commonly used in everyday language, they might appear to exhibit high productivity (i.e., creativity). However, our findings suggest the opposite—that their productivity is relatively low. In other words, this SVC applies to a relatively limited group of verbs.

Figure 7 displays the ten verbs with the highest and lowest association strengths with SVCs. Some verbs in Figure 7 have not been previously described in SVC studies (e.g., *resit, conduct,* and *pass* in [Verb + EXAMINATION]). While this finding supports our claim in Section 2.2 that the verb-centered approach alone results in low coverage, many verbs with low collostructional strength (e.g., *commission, prove, achieve*) were also found in SVCs. Thus, caution is necessary when predicting that verbs with low collostructional strength with EXAMINATION instantiate regular transitive verb constructions rather than SVCs.

In (11b), the complement following a support verb is generally characterized as possessing an argument structure and functioning as an input for argument transfer (Grimshaw, 1990), as reviewed in Section 2.1. When distinguishing SVCs from regular transitive verb constructions within a constructional network, as represented in Figure 1, it may be effective to specify the slot for the argument instantiated as an *of*-phrase. However, while this specification suggests that the presence of an *of*-phrase serves as a sufficient condition for SVCs, the absence of *an* of-phrase does not necessarily imply that the given environment is not an SVC. Based on the frame-semantic identification criteria presented in (7–8), an *of*-phrase appears to fulfill one of the frame elements within the EXAMINING frame (e.g., <Examiner>, <Examinee>). Therefore, to further examine the relationship between the presence of an of-phrase and whether the given environment instantiates SVCs, additional analysis focusing on the frame elements within the *of*-phrase is necessary.

Finally, as a key finding from the logistic regression analysis conducted in Section 4.2, (11c) indicates that the features annotated in this study provide a relatively accurate model for determining whether [Verb + EXAMINATION] instantiates SVCs. However, as discussed in relation to (11a), this predictive power is likely offset by the idiomaticity of the construction, suggesting that the constructional knowledge enabling the use of SVCs may be more specific than the sketch presented in Figure 2. Future



Figure 7. Verbs with the highest and lowest association strength with SVCs

research should explore the appropriate level of abstraction for constructional units.

## 6. Conclusion

This paper presented a corpus-based survey of [Verb + EXAMINATION] to explore the formal specifications of SVCs. To describe the linguistic knowledge that enables speakers to use SVCs, we must address two issues: the contextual ambiguity between support verbs and their heavy-sense counterparts, and the limited coverage of low-frequency support verbs. We argue for the need for formal specifications of complements to address the limitations inherent in the verb-centered approach.

Our results in Section 4 showed (i) verbs with high collostructional strength, and (ii) the grammatical properties of EXAMINATION (e.g., the occurrence of an *of*-phrase) when [Verb + EXAMINATION] instantiates an SVC. Furthermore, we found that verbs such as *resit, conduct,* and *pass*, which exhibit high collostructional strength with *examination*, can function as support verbs—types of verbs that have not been addressed in previous studies on SVCs. These findings indicate the importance of examining formal properties of complements in addition to semantic constraints when describing SVCs as constructions, which Brugman (2001) emphasized in relation to aspectual constraints.

Another limitation of this study is that the scope of analysis was restricted to EXAMINATION. First, it is necessary to examine whether the deverbal noun EXAMI-NATION can represent the behavior of SVCs where deverbal nouns of the V-*tion* type function as complements. Additionally, further investigation is required to identify the formal characteristics of SVCs when deverbal nouns other than the *-tion* type (e.g., *make a catch, take a test*) occur as complements.

## Notes

- \*1 According to Fujii and Uegaki (2008), SVCs and LVCs share many semantic and syntactic properties. These authors distinguish between LVCs and SVCs from the perspectives of Frame Semantics and Construction Grammar. Specifically, they classify constructions like *make a complaint*, where the verb's semantic contribution is minimal and the construction exhibits high generality, as LVCs. In contrast, constructions like *lodge a complaint*, where the verb contributes to the overall meaning of the phrase and displays a higher degree of idiomaticity, are categorized as SVCs. In the following discussion, the broader term SVCs and support verbs will be used.
- \*<sup>2</sup> In the following discussion, EXAMINATION in [Verb + EXAMINATION] represents any NPs headed by *examination*.
- \*<sup>3</sup> We also considered the results of Kambara (2021) during the annotation process. While Kambara (2021, p. 154) presents 28 verbs used as LVCs, we included all verbs extracted from the corpus.
- \*4 To represent the constructional schema, we divide [Verb + EXAMINATION] into a phonological pole and a semantic one. The index *i* indicates the correspondence relation between the phonological and semantic poles.

## Acknowledgements

Portions of this work were supported by JSPS KAKENHI Grant Number 24K16143. We thank the comments from the two anonymous reviewers. Usual disclaimers apply.

#### Data availability

All codes and data are available on OSF (Open Science Framework): 10.17605/ OSF.IO/4TAEK

## Dictionaries

New Oxford American Dictionary. 3rd ed. 2010.

## References

- Fillmore, C. J., Baker, C. F., & Sato, H. (2002). Seeing arguments through transparent structures. In *Proceedings of Third International Conference on Language Resources and Evaluation (LREC 2002)* (pp. 787–791).
- Fillmore, C. J., Johnson, C. R. & Petruck, M. R. L. (2003). Background to FrameNet. *International Journal of Lexicography*, 16(3), 235–250.
- Fujii, S. & Uegaki, W. (2008). Shien doushi koubun no bunseki: Ko-pasu ni motoduku koubun riron teki apuro-chi [An analysis of Support Verb Constructions: A constructionist approach based on corpus]. *Proceedings of the thirteenth annual meeting of the association for natural language processing* (pp. 943–946) [Written in Japanese].
- Giparaitė, J. (2024). A corpus-based analysis of light verb constructions with MAKE and DO in British English. *Kalbotyra*, *76*, 18–41. https://doi.org/10.15388/Kalbotyra.2023.76.2
- Goldberg, A. E. (2006). *Constructions at works: The nature of generalization in language*. Oxford University Press.
- Gries, S. T. (2019). 15 years of collostructions: Some long overdue additions/corrections (to/of actually all sorts of corpus-linguistics measures). *International Journal of Corpus Linguistics*, 24(3), 385–412.
- Gries, S. T. (2021). *Statistics for linguists with R: A practical introduction*. De Gruyter Mouton.
- Gries, S. T. (2023). Overhauling collostructional analysis: Towards more descriptive simplicity and more explanatory adequacy. *Cognitive Semantics*, *9*(3), 351–386.
- Gries, S. T. (2024a). Frequency, dispersion, association, and keyness: Revising and tupleizing corpus-linguistic measures. John Benjamins.
- Gries, S. T. (2024b). *Coll.analysis 4.1. A script for R to compute and perform collostructional analyses.* https://www.stgries.info/teaching/groningen/index.html.
- Gries, S. T. & Stefanowitsch, A. (2004a). Co-varying collexemes in the into-causative. In M. Achard & S. Kemmer (Eds.), *Language, culture, and mind* (pp. 225–236). CSLI.
- Gries, S. T. & Stefanowitsch, A. (2004b). Extending collostructional analysis: A corpus-based perspective on alternations. *International Journal of Corpus Linguistics*, 9(1), 97–129.
- Grimshaw, J. (1990). Argument structure. MIT Press.
- Grimshaw, J. & Mester, A. (1988). Light verbs and θ-marking. Linguistic Inquiry, 19, 205–232.
- Hoffmann, T. (2022). *Construction grammar: The structure of English*. Cambridge University Press.
- Jespersen, O. (1940). A modern English grammar on historical principles: Part V., Syntax.

Copenhagen.

- Kambara, K. (2021). Fure-mu imiron ni motoduku meishi no imi bunseki [Frame semantics analysis of nominal meanings] [Doctoral dissertation, Graduate School of Human and Environmental Studies, Kyoto University]. https://doi.org/10.14989/doctor.k23581.
- Kambara, K. & Chika, T. (2023). Toward a corpus-based identification of nominal relationality and uniqueness: A constructionist approach. In *Proceedings of the 37th Pacific Asia Conference on Language, Information and Computation* (pp. 661–669).
- Kambara, K., Takahashi, N., & Chika, T. (in press). Quantifying the degrees of relationality: A collostructional approach. In Proceedings of the 38th Pacific Asia Conference on Language, Information and Computation.
- Kilgarriff, A., Rychly, P., Smrz, P., & Tugwell, D. (2004). The sketch engine. In G. Williams and S. Vessierssier (eds.), *Proceedings of 11th EURALEX International Congress* (pp. 105–116).
- Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., ... & Suchomel, V. (2014). The sketch engine: Ten years on. *Lexicography*, 1(1), 7–36.
- Langacker, R. W. (2000). A dynamic usage-based model. In *Grammar and conceptualization* (pp. 1–63). De Gruyter Mouton.
- Langer, S. (2005). A formal specification of support verb constructions. In Langer, S. & Schnorbusch, D. Semantik im lexikon (pp. 179–202). Narr Dr. Gunter.
- Levshina, N. (2015). *How to do linguistics with R: Data exploration and statistical analysis.* John Benjamins.
- Newman, J. (1996). GIVE: A cognitive linguistic study. John Benjamins.
- R Core Team (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/.
- Stefanowitsch, A. & Gries, S. T. (2003). Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics*, 8(2), 209–243.
- Stefanowitsch, A. & Gries, S. T. (2005). Covarying collexemes. *Corpus Linguistics and Linguistic Theory*, *1*(1), 1–43.
- Speelman, D. (2014). Logistic regression: A confirmatory technique for comparisons in corpus linguistics. In D. Glynn and J. A. Robinson (eds.), *Corpus methods for semantics: Quantitative studies in polysemy and synonymy* (pp. 487–533). John Benjamins.
- Taylor, J. R. (2012). *The mental corpus: How language is represented in the mind*. Oxford University Publishing.
- Uchida, S. (2010). On the lexicographic descriptions of event nouns: An insight from frame semantics. *The Journal of Institute for Language and Education Research*, *27*, 411–426.
- Wickham, H., Çetinkaya-Rundel, M., & Grolemund, G. (2023). R for data science: Import, tidy, transform, visualize, and model data. O'Reilly Media.
- Wierzbicka, A. (1982). Why can you *have a drink* when you can't *\*have an eat*?. *Language* 58 (4), 753–799.

(近 大志 京都大学)(神原 一帆 情報通信研究機構/立命館大学)