

TRANSLATING LIGHT VERBS IN LEGAL CONTEXTS: ENGLISH-CROATIAN PERSPECTIVES

Jana Kegelj

Faculty of Maritime Studies, University of Rijeka, Studentska 2, Rijeka 51000, Croatia

Abstract

Light verb constructions (LVCs), such as ‘make a decision’ or ‘give a smile’, are verbo-nominal structures in which the verb has lost its meaning to a certain degree and the predicative noun determines the meaning of the construction. The association between the verb and the noun is quite strong. LVCs usually contain high-frequency verbs, such as make, take, do, give, in their so called ‘light’ meaning.

This paper focuses on a cross-linguistic comparison of the verbs make, take and give within LVCs in a corpus of English legal texts and a parallel corpus of their translations into Croatian. The corpus-based analysis of LVCs also considered the translation strategies used in the language pair. Since LVCs are often the result of the process of nominalization in different registers, it is expected that they will be quite frequent in the legal register, which shows a tendency towards nominalization. The study also explores the frequency and valency of the verbs make, take and give in the English legal corpus using word frequency lists and statistical measures of association between words. It then considers the translation equivalents in the Croatian corpus. The study contributes to the understanding of LVCs in this particular register, as well as the issues of translating such constructions.

Keywords: *light verb constructions, legal register, corpus analysis, parallel corpora, cross-linguistic study*

1. INTRODUCTION

Light verbs, also called thin verbs [1], function verbs [2], support verbs [3], operator verbs [4], vector verbs [5], and light verb constructions (LVCs), also referred to as stretched verb constructions [1], complex predicates [6], complex predicate constructions [6], support verb constructions [3] and verbo-nominal constructions [7], were the focus of many theoretical analyses, particularly in German and French linguistic circles, for several decades. The term was used to describe functional verbs with a very general meaning combining with a noun and forming a lexicalized predicate, such as ‘make a complaint’, ‘take a decision’, ‘give assistance’. The LV in these structures contributes to the meaning of the entire unit to a small extent, i.e. the meaning of the verb is semantically reduced [8: 63], while the meaning of the whole unit is derived from the noun. Owing to their complexity, variability and wide use, they have been the subject of studies that have analysed LVCs from many aspects, e.g. cross-linguistic studies, diachronic studies, translation studies, semantic analysis, discourse analysis, second language acquisition, etc. In recent years, the study of LVCs was directed specifically towards automatic identification of LVCs in large corpora to cater for the needs of natural language processing and large language models (e.g. [9], [10], [11]). While there are a few corpus based studies that focus on LVCs (cf. [12], [13], [14]), analyzing their construction and frequency, the studies focusing on parallel corpora and translation strategies used are few, particularly the ones in the English-Croatian language pair. This study focused on the verbs *make*, *take* and *give* in their so called ‘light’ meanings in a parallel English-Croatian corpus of maritime legal texts. More specifically, the frequency and valency of each LV in the corpus was compared, possible translation equivalents investigated and translation strategies identified. The aim was to investigate the use of LVCs containing *make*, *take* and *give* within one register: the legal register, and the ways these were translated to Croatian to identify the similarities and differences between languages and possible challenges they present for the process of translation.

Maritime legal English is a specific sub-variety of the umbrella term Maritime English, which includes legal texts regulating various legal aspects of maritime venture, shipping business, ship management, marine insurance, liability, safety at sea, work conditions on board, ship’s ownership and registration,

marine pollution, etc. [15]. The language of maritime legal English shares some features of legal English, which mainly refers to greater number of nouns than verbs, frequent use of conjunctions, passive voice, impersonal style, limited use of tenses, long complex sentences, frequent use of conditional clauses, latinisms, clichés, repetitions, while its style is described as analytical, schematic, objective, logical, explicit, specific, unemotional, stylistically unmarked, impersonal, redundant, with intertextual references and formulaic expression. The same features can also be found in the Croatian legal register and they represent potential universal features of the register.

When it comes to translating prescriptive legislative texts of this kind, [16: 8] claims that the language of international conventions is deliberately ambiguous, vague and unclear, which is why translators need to be particularly cautious and refrain from explaining or clarifying because such expressions result from the specific procedure of international negotiations. Translators use various strategies in the process of translation in order to achieve functional equivalence. The strategy applied in a particular circumstance depends of the type of text being translated, concepts, purpose of the text and the target audience [17]. This might include making the translation more explicit or more implicit than the original, simplifying, adjusting to the typical patterns of the target language or adapting to the target culture, adding or omitting to fit the target language, etc. In the process, the source language (SL) might influence the translation so examples of SL interference and appearance of unique linguistic elements or untypical multi-word combinations are not uncommon.

As research has shown, LVCs tend to be challenging for non-native speakers of English as they frequently underuse, overuse or misuse LVs. This occurs because it is often not easy to determine which LV to use, e.g. we can say ‘make a decision’ but we cannot say ‘do a decision’. Some studies established that advanced level speakers tend to use LVCs more frequently than speakers at lower proficiency levels [18] and also that the use of LVCs varies across regional varieties of English [19: 151–152]. Nonetheless, LVCs have been shown to be more frequent than their full verb counterparts, despite the law of language economy [20: 580]. This might be due to the fact that LVCs show greater flexibility in modification and compositional structure, i.e. they can be modified by various elements and their structure can be discontinued within a sentence [21: 3]. Therefore, they are considered to be somewhere in between purely compositional language and fixed idioms (Figure 1).

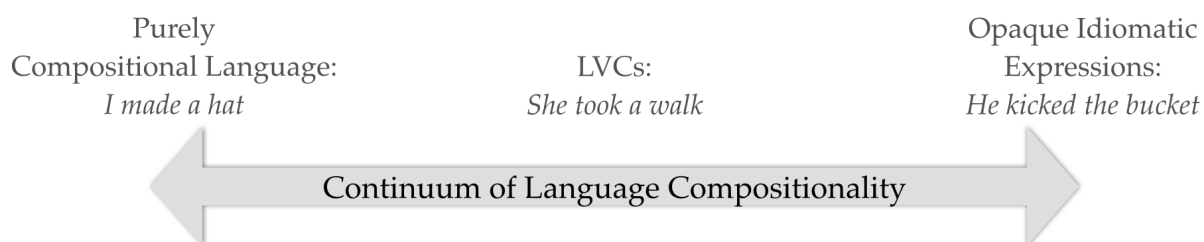


Fig. 1. Continuum of Language Compositionality [21].

Both the features of LVCs and translation challenges were taken into consideration in formulating the hypotheses. The first hypothesis was that, as the legal register in both languages shows a tendency towards nominalization, the English LVCs will be translated by the Croatian equivalent LVCs more frequently. The second hypothesis was that the English LVs under study will be translated by a few of the most frequent Croatian LVs, typical for the register. The third hypothesis was that there will be some overlap in the Croatian equivalent LVs, i.e. that the same Croatian LVs will be used for the translation of English LVs.

2. MATERIALS AND METHODS

2.1. Materials

The study was conducted on two parallel corpora of texts from the area of maritime legal English, MarLaw EN (553,120 tokens) and MarLaw CRO (498,389 tokens). The corpora consist of 14 international conventions and five international regulations in the English corpus and their translations into Croatian in the Croatian corpus. As such, the texts belong to the category of legal prescriptive texts, that are public and normative. EU maritime regulations were not selected for the corpora as the results of previous research have shown that EU English may be regarded as a variety of British legal English (cf. [22], [23], [24], [25]) and also due to the fact that EU regulations are composed based on specific guidelines and translated according to special criteria and instructions. For these reasons, the language of EU legal documents in English may contaminate the results of the legal documents originally written in English. The international conventions and regulations selected for the corpus regulate various legal aspects of shipping, including ownership, liability, safety, insurance matters, rights and obligations of different stakeholders in maritime ventures. The Republic of Croatia as a maritime-oriented country is a signatory of all the selected conventions and has implemented many of the regulations in its national legislation. The authors of the original texts are legal experts, while the translators are unknown, presumably native speakers of Croatian. The translations are formally published on the relevant web pages of the Croatian Ministry of Justice and Public Administration and Ministry of the Sea, Transport and Infrastructure.

The corpus compilation was conducted in several stages. First, the planning stage which involved defining the type of corpora and the relevant criteria for the selection of texts for the corpora to make it suitable for the study. The criteria included defining the register and field from which to collect the texts, selecting texts that have been translated into Croatian, using whole texts not just excerpts, and omitting EU legislation. This was followed by collecting relevant texts, filtering them according to the previously defined criteria and finally, preparing the texts for uploading into the online corpus analysis tool, Sketch Engine [26].

2.2. Methods

In the corpus-based analysis of LVCs it was necessary to triangulate the methods in order to validate the obtained results. The analysis in Sketch Engine using advanced features of corpus query language (CQL) showed that the identification of LVCs requires additional qualitative insight owing to the nature of LVCs. The problem of identification resulted from the fact that the individual elements of LVCs are often disjoined by other linguistic units, e.g. “**make** appropriate **recommendations** to the Council”, “**make** any **recommendations** which it may find appropriate”. Another problem is that LVCs may appear in both active and passive form (e.g. “additional **entries** in the radiolog should be **made** in accordance with...””) owing to which CQL does not provide adequate results in concordances. On the other hand, CQL may provide hits that are not LVCs but only the ‘heavy’ use of an LV and a noun appearing in the same context (e.g. “they will make the decision known to the public”). Therefore, after obtaining the results by CQL, the hits had to be filtered in the concordance to obtain the final results. The results from the concordance were then further cross-checked with the results of the Word Sketch tool in Sketch Engine. This is a special feature offered by Sketch Engine that provides a word’s collocates organized into categories, and in this case the focus was on verb + noun structures. The results of Word Sketch were analyzed according to a set of specific criteria: the noun denotes an action, event or state, the noun carries the meaning of the whole, noun and verb together denote one action and behave as a single semantic unit, the structure can be paraphrased by a simple verb (Figure 2).

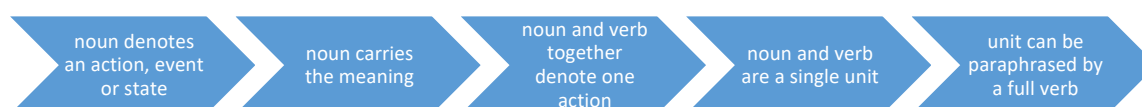


Fig. 2. Process of identifying LVCs.

The study focused specifically on the analysis of LVCs with the verbs *make*, *take* and *give* in the corpus MarLaw EN. Using the above-described methods, the frequencies of LVCs containing *make*, *take* and *give* were found along with the different nouns that form a part of their LVCs. To get a deeper insight into the valency and strength of the connection between the light verbs and their nouns, the association measure logDice and log likelihood were obtained for each LVC. As logDice provides typical collocations, not just words that frequently co-occur, this was considered as benchmark, while log likelihood served as a check of statistical significance of each LVC. The second part of the study included the analysis of translations of English LVCs into Croatian by examining the corpus of parallel Croatian translations in Sketch Engine.

3. RESULTS

After making some adjustments to CQL, the following formula was used for searching LVs in the corpus MarLaw EN `[lemma="(light verb)"][[tag="N.*"]]` and the final results for frequencies were obtained after manual filtering. The results, shown in Table 1, indicate that the verb *take* is the most frequent in the corpus, but it has the least number of noun collocates. The verb *make* is the most valent of the three.

Table 1. Frequency and number of collocates for the verbs *make*, *take* and *give*.

Verb	Absolute freq.	Relative freq. (per million)	No of collocates
make	136	245,88	28
take	552	997,98	16
give	99	178,98	26

After analyzing the frequencies, the study focused on the statistical significance of the obtained results for the collocations to see whether the strength of association between the LV and the noun are significant and not due to chance. In this sense, the logDice values and log likelihood values were considered to identify those structures which represent typical collocations for the corpora under study. The results, shown in Figure 3, represent the logDice values for collocations with *make*, *take* and *give* that exceed the threshold set at 7.00, which is considered significant. Another criterion for including the LVC in the results was its log likelihood value that was greater than 20, as this shows that the probability of the LVCs appearing by chance is less than 1 %.

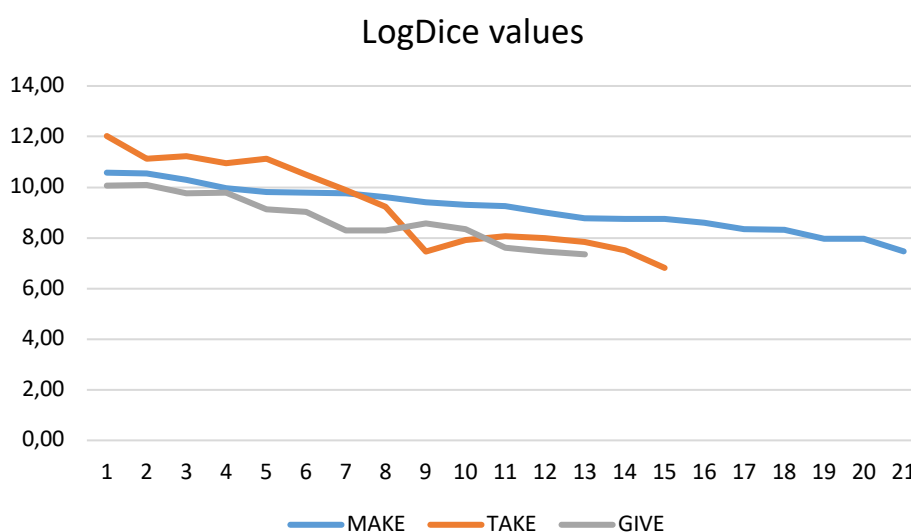


Fig. 3. logDice values for collocations with *make*, *take* and *give*.

These results complement the results of LV frequencies, showing the strength of association that each LV makes with their collocates. Within the MarLaw EN corpus, the verb *take* makes the strongest associations, while the verb *make* has a greater number of strong associations. The verb *give* is the least valent and least productive, as many of the associations it makes with nouns are rather weak (>7). However, when we take a look at the different meanings of their LVCs, we can see that the verb *give* shows a greater variability in meanings than the other two verbs, particularly the verb *take* (Table 2).

Table 2. Various meanings of LVCs with the verbs *make*, *take* and *give*. Source: author.

VERB	MEANING	EXAMPLE
MAKE	write/submit (a document)	<i>make a submission; make an application</i>
	perform sth	<i>make a departure, make a seizure</i>
	arrange	<i>make arrangements</i>
	decide	<i>make a nomination; make an award</i>
TAKE	consider	<i>take into account; take into consideration</i>
	act	<i>take measures; take action</i>
GIVE	publish	<i>give publicity</i>
	notify	<i>give warning; give information</i>
	respect	<i>give regard; give attention</i>
	guarantee	<i>give assurance</i>
	help	<i>give assistance; give support</i>
	train	<i>give training</i>

The second part of the study focused on the translation equivalents of LVCs in the Croatian parallel corpus MarLaw CRO. First, the frequencies of nouns and verbs in the two corpora were extracted to analyze the tendency of a corpus towards nominalization as the use of LVCs is usually connected to that. In that sense, the Croatian corpus showed a greater tendency towards nominalization than the English corpus, with a noun to verb ratio of 3,54, compared to the noun-verb ratio in the English corpus of 2,58. Subsequently, the concordance lines containing the LVCs identified in the previous stage of the study were analyzed in parallel with the Croatian translations. All translations were identified and categorized either as translation into an LVC in Croatian, translation into a full verb or translation into a noun. The results are shown in Figure 4.

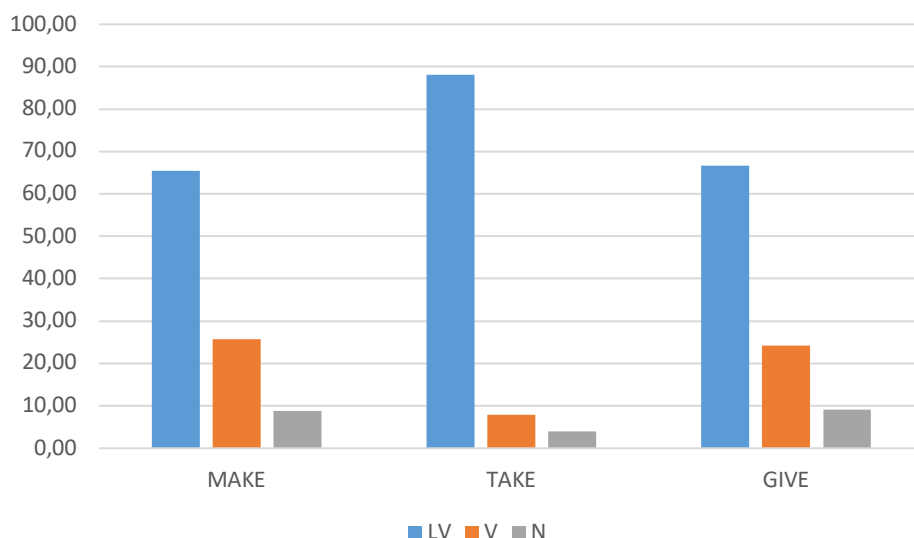


Fig. 4. Translation outcomes by verb (LV, V, N).

The results show that the greatest part of English LVCs was translated by an LVC in Croatian, while the verbs were the second choice for translators. In this sense, the verbs *make* and *give* show similar tendencies while the verb *take* differs in the sense that it was more frequently translated by a Croatian LVC than the other two verbs.

4. DISCUSSION

The results of the corpus analysis indicate that while the verb *take* is the most frequent of the three LVs, it has the least number of collocates. However, the strength of the bonds it makes with its collocates is rather high. When taking a look at the translations into Croatian, it can be seen that *take* is more frequently translated by a Croatian LVC than the other two verbs. This might be due to the greater strength of association that it has with its collocates so its LVCs are perceived as fixed units of language. It could also be attributed to the universal feature of the register, i.e. similar use of language structures within the same register, even though the language is different. The same applies to the fact that LVCs are more frequently translated as LVCs in Croatian, indicating a tendency towards nominalization in the same register in English and in Croatian.

The translations of LVCs into Croatian show a variety of solutions. In the majority of cases (81.4%), the LVCs are translated by a Croatian equivalent LVC. However, in 18.4% of cases they are translated by a Croatian full verb or a noun. The meaning of the full verb is much clearer and it is retrieved faster than the combination of the light verb and a noun, which makes these expression in translation easier to understand, particularly in cases where the LVC in the original is discontinued by other elements. Such cases might be interpreted as instances of simplification in translation, or explicitation as the translated expression more explicitly states the meaning.

Further analysis of translations shows that in Croatian LVCs there is a variety of verbs used, as shown in Figure 5.

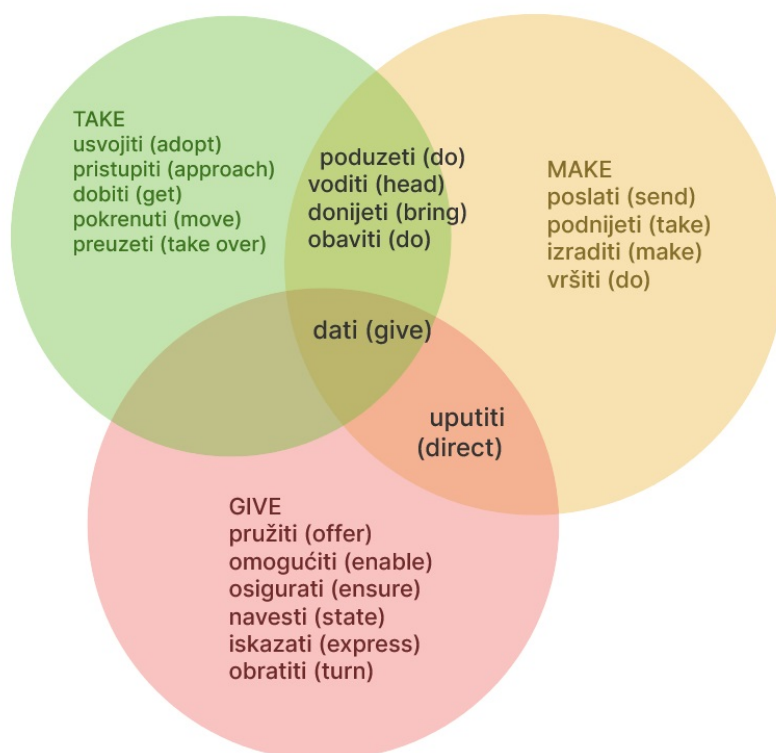


Fig. 5. Venn diagram of Croatian translations for LVCs with *make*, *take* and *give*

This suggests a greater variation in LVC combinations in Croatian, which is further corroborated by frequencies of light verbs in the two corpora: *make*, *take* and *give* together appear 4213 times, while all the Croatian light verbs together have a frequency of 2785. It is further interesting to notice the light verb translations which overlap. The Croatian verb *dati* (give) is a common equivalent for all three English verbs. However, the English verb *give* has the least number of Croatian equivalents in common with the other two light verbs, while the verbs *take* and *make* have several common equivalents in Croatian.

5. CONCLUSIONS

Owing to the tendency of the register towards nominalization, LVCs represent one of the typical features of legal registers across languages. The LVCs in both English and Croatian legal corpora, analyzed in this study, showed shared structural features and high productivity. Among the three verbs, the verb *take* is the most frequent in the English legal corpus but *make* is the most productive in terms of valency. The verb *give* shows the widest range of semantic variability despite having weaker associations overall. As the majority of English LVCs were translated using equivalent LVCs in Croatian, this indicates a cross-linguistic preference for maintaining the nominal structure in translation. This confirms the first hypothesis about the legal discourse tending to preserve nominalization to maintain a degree of formality and information density. However, as Croatian translations of English LVCs displayed a broader range of light verb choices, the second hypothesis was not confirmed. It seems that Croatian has a variety of verbs that may be used in LVCs, which needs to be taken into consideration when translating. The Croatian verb “*dati*” may serve as equivalent for different English light verbs, and could therefore be considered typical of the register, but it still is not as frequent in the Croatian maritime legal register as the English light verbs are in the English maritime legal register. This suggests a higher degree of flexibility or fluidity in Croatian LVC constructions compared to their English counterparts. The third hypothesis was partially confirmed as the LVCs containing the verbs *make* and *take* have some common

equivalents in Croatian, but the LVCs with the verb *give* were translated with a different set of light verbs in Croatian. This might be due to the intrinsic semantic differences between the verbs in English. Besides maintaining the LVC structure, which shows a structural consistency of the legal language, the translations demonstrated a certain degree of simplification and explicitation in cases when the English LVCs were translated using Croatian full verbs of nouns, which indicates some nuancing in linguistic decisions involved in cross-linguistic translation of formulaic expressions. This aspect would require further study into the effects of using full verbs to see whether there are any legal nuances lost or added in translation. Other possible future research might focus on the patterns of LVC usage in other specialized registers (e.g. administrative, medical, academic) to determine whether the structural patterns and the translation strategies are register-specific. Also, in light of recent technological advancements, it should be analyzed how well legal LVCs are handled by machine translation tools in English-Croatian language pair.

ACKNOWLEDGMENTS

The research was financed by University of Rijeka, Research Project ZIP (project code UNIRI-ZIP-2103-14-22).

REFERENCES

1. Allerton, DJ 2002, *Stretched Verb Constructions in English*, Routledge.
2. Nickel, G 1968, 'Complex Verbal Structures in English', *International Review of Applied Linguistics*, vol. 6, pp. 1-21.
3. Gross, G 1989, *Les constructions converses du français*, Librairie Droz, Geneve-Paris.
4. Rassi, A, Santos-Turati, C, Baptista, J, Mamede, N and Vale, O 2014, 'The fuzzy boundaries of operator verb and support verb constructions with dar "give" and ter "have" in Brazilian Portuguese', *Proceedings of the Workshop on Lexical and Grammatical Resources for Language Processing*, vol. 39, pp. 92–101.
5. Nespital, H 1989, 'Verbal Aspects and Lexical Semantics in Indo-Aryan Languages: The Typology of Verbal Expression ("Compound Verbs") and their Relation to Simple Verbs', *Studien zur Indologie und Iranistik*, vol. 15, pp. 159-195.
6. Goldberg, AE 1996, 'Words by default: optimizing constraints and the Persian complex predicate', *Proceedings of the Berkeley Linguistics Society*, vol. 22, no. 1, pp. 132-146. <https://doi.org/doi:10.7282/T3F47RC6>.
7. Cetnarowska, B 2012, 'On English verbo-nominal constructions and different construals of situations', in B Cetnarowska and O Glebova (eds.), *Image, imagery, imagination in contemporary English studies*, Wydawnictwo AJD, Częstochowa, pp. 189–207.
8. Fleischhauer, J 2021, 'Light Verb Constructions and Their Families – A Corpus Study on German stehen unter-LVCs', *Proceedings of the 17th Workshop on Multiword Expressions*, Bangkok, Thailand (viewed 6 August 2021), pp. 63–69.
9. Maldonado, A, Han, L, Moreau, E, Alsulaimani, A, Chowdhury, KD, Vogel, C, & Liu, Q 2017, 'Detection of Verbal Multi-Word Expressions via Conditional Random Fields with Syntactic Dependency Features and Semantic Re-Ranking', *Proceedings of the 13th Workshop on Multiword Expressions (MWE 2017)*, 114–120. <https://doi.org/10.18653/v1/w17-1715>.
10. Nagy, I, Rácz, A, Vincze, V 2020, 'Detecting light verb constructions across languages', *Natural Language Engineering*, vol. 26, pp. 319–348, doi:10.1017/S1351324919000330.

11. Tan, KSN, Lim, TM, Tan, CW, Chew, WW 2022, 'Automatic identification of light verb constructions: a review', *IEM Journal Special Edition: International Conference on Digital Transformation and Applications (ICDXA 2021)*, pp. 95-100.
12. Algeo, J 1995, 'Having a Look at the Expanded Predicate', in B Aarts and CF Meyer (eds.), *The Verb in Contemporary English: Theory and Description*, CUP, Cambridge, pp. 203-217.
13. Leech, G, Hundt, M, Mair, Ch and Smith, N 2009, *Change in Contemporary English*, Cambridge University Press, Cambridge.
14. Ronan, P and Schneider, G 2015, 'Determining light verb constructions in contemporary British and Irish English', *International Journal of Corpus Linguistics*, vol. 20, no. 3, pp. 326-354. DOI: <https://doi.org/10.1075/ijcl.20.3.03ron>.
15. Bocanegra-Valle, A 2011, 'The language of seafaring: Standardized conventions and discursive features in speech communications', *International Journal of English Studies*, vol. 11, no. 1, pp. 35-53. <https://doi.org/10.6018/ijes/2011/1/137091>.
16. Šarčević, S 2000, 'Legal Translation and Translation Theory: a Receiver-Oriented Approach', in *La Traduction Juridique, Histoire, theorie(s) et pratique. Actes*, Universite de Geneve, Ecole de Traduction et d'Interpretation/ASTTI, Berne Geneve, pp. 329-47.
17. Gémard, JC 2006, 'Traduction spécialisée et droit. Langages du droit, styles et sens', in M. Gotti, & S. Šarčević, *Insights into Specialized Translation*, pp. 79-106, Peter Lang, Bern.
18. Gilquin, G 2019, 'Light verb constructions in spoken L2 English: An exploratory cross-sectional study', *International Journal of Learner Corpus Research*, vol 5, no. 2, pp. 181-206.
19. Smith, A 2009, 'Light verbs in Australian, New Zealand and British English', in P Peters, P Collins and A Smith (eds.), *Comparative studies in Australian and New Zealand English: Grammar and beyond*, John Benjamins, Amsterdam, pp. 139-55.
20. Bonial, C and Pollard, KA 2020, 'Choosing an event description: What a PropBank study reveals about the contrast between light verb constructions and counterpart synthetic verbs', *Journal of Linguistics*, vol. 56, pp. 577-600. <https://doi.org/10.1017/S0022226720000109>.
21. Bonial, C 2021, 'Take a look at this! Form, function, and productivity of English light verb constructions', *Colorado Research in Linguistics*, vol. 25.
22. Trosborg, A 1997, *Rhetorical Strategies in Legal Language: Discourse Analysis of Statutes and Contracts*, Gunter Narr Verlag, Tübingen.
23. Koskinen, K 2000, 'Institutional Illusions. Translating in the EU Commission', *The Translator*, vol. 6, no. 1, pp. 49-65.
24. Tosi, A 2005, 'EU Translation Problems and the Danger of Linguistic Devaluation', *International Journal of Applied Linguistics*, vol. 15, no. 3, pp. 384-388.
25. Biel, Ł 2014, 'The textual fit of translated EU law: a corpus-based study of deontic modality', *The Translator*, vol. 20, no. 3, pp. 332-355.
26. Kilgariff, A, Rychly, P, Smrz, P and Tugwell, D 2004, 'ITRI-04-08 The Sketch Engine', *Proceedings of Euralex*, Lorient, France, July 2004, pp. 105-116.