

---

# A corpus linguistic study of constructional equivalence for the Indonesian translation of ROB and STEAL based on the *Open Subtitles Parallel Corpus*

Gede Primahadi Wijaya Rajeg\*, I Made Rajeg

*Bachelor of English, Udayana University, Bali, Indonesia*

---

## ABSTRACT

Catford's classic idea in translation theory indicates the measurability of translation equivalence. Following up on this idea, this paper offers a case study to measure the translation equivalence of English verbal near-synonyms ROB and STEAL (R&S), especially the equivalence at the constructional level. Adopting a quantitative corpus linguistic method and the Construction Grammar approach, we analyse random usage samples of R&S from English-Indonesian parallel corpus for the degree of constructional equivalence along two dimensions: (i) the profiled participant roles and (ii) the grammatical construction types of these verbs. We discover that the Indonesian translations maintain a high degree of equivalences along these dimensions, albeit with few variations. This suggests that the translators attempt to be as faithful as possible to the source texts. Furthermore, our study reveals the translation norms/typicality in how the constructional profiles of the near-synonyms R&S are translated into Indonesian. The paper generally seeks to demonstrate how such a central notion as equivalence in translation studies can be investigated using parallel corpora and the quantitative corpus linguistic method.

## ARTICLE INFO

*Paper type:*  
Research Article

*Article history:*  
Received : 30/12/2021  
Revised : 26/04/2022  
Accepted : 26/05/2022

*Keywords:*

- construction grammar
- constructional equivalence
- parallel corpus
- quantitative corpus linguistics
- translation study
- translation equivalence
- 

---

## 1 Introduction

The recent explosion of digital data in all fields has positively impacted the broader field of linguistics, especially on the availability of large collection of digitalised samples of textual data, which is also referred to as the language corpus (or corpora, for plural). Nowadays, there are two major types of language corpora as far as the language is concerned. The first type is monolingual corpora that contain texts from a single language (e.g., British English language corpora). The second one is bi-/multilingual, parallel corpora containing “original texts in language A vs. their translations in language B” (Mikhailov & Cooper, 2016, p. 5) (e.g., corpus of an original novel and its translations in different languages).

The past two decades have seen a rapid growth of corpus-based translation studies as a branch of translation studies (Baker, 1993; Hu, 2016). Corpus-based translation studies share the same philosophy with corpus linguistics as it embraces empiricism and descriptive focus over the prevailingly predominant paradigm of prescriptivism in translation studies. The prescriptive approach of translations “assumes the primacy of the source text and argues that the target text should seek to be as equivalent to the source text as possible” (Hu, 2016, p. 1; Zanettin, 2014, p. 180). In doing so, the prescriptive approach “relies heavily on intuition, anecdotal evidence, or a small number of samples” (Hu, 2016, p. 1). The corpus-based approach, in contrast, focuses on “describing the features of translation and translational norms” to reveal “the nature of translation and the interrelationship between translation

---

\* E-mail address: [primahadi\\_wijaya@unud.ac.id](mailto:primahadi_wijaya@unud.ac.id) (Rajeg), [made\\_rajeg@unud.ac.id](mailto:made_rajeg@unud.ac.id) (Rajeg)

and social culture, based on statistical analysis of a wealth of corpus data” (Hu, 2016, p. 1).

In this paper, we embrace the quantitative, empirical foundation of corpus linguistics to investigate translation equivalence and norm (§1.1 below) in the samples of Indonesian translated texts from English movies and TV subtitles (§2.1). As a case study, we focus on a pair of synonyms of dispossession verbs in English, namely ROB and STEAL (R&S), and the Indonesian equivalence of the verbs’ constructional properties (i.e., profiled participant roles and construction types; see §1.2 for further details on these points).

### 1.1 *A brief overview on the translation equivalence and norm*

Equivalence is a fundamental notion in translation theories (Baker, 1993; Kenny, 2001). From the perspective of the equivalence-based theory of translation, equivalence is understood as “the relationship between a source text (ST) and a target text (TT) that allows the TT to be considered as a translation of the ST in the first place” (Kenny, 2001, p. 77). In this paper, following Toury (1980, p. 39, as cited in Kenny, 2001, p. 79), we view equivalence as an empirical category that could be established after the translation: “actual relationships between actual utterances in two languages (and literatures), recognised as TTs and STs – which are subject to direct observation” (Kenny, 2001, p. 79).

With this view of equivalence as an empirical category and the use of parallel translation corpus, we can measure and potentially establish equivalence after the fact based on the evidence available in the corpus. Furthermore, Toury (as cited in Kenny, 2001) notes that the study of equivalence should not focus on (i) whether the two texts are either equivalent or not, which is essentially a prescriptive view (Hu, 2016, pp. 5, 140), but (ii) the kind and degree of equivalence. The parallel corpus allows us to inspect not just a single or two examples, but a set of samples of utterances that can be analysed both qualitatively and quantitatively, allowing us to measure the degree of equivalence between the utterances in ST and TT.

The view that equivalence could be measured in probabilistic terms goes back to Catford’s (1965) classic idea. The probability of using a given item (in comparison to the other items) as the translation of X can be recast as a translation rule or “norm” (Toury, 2000). Following Baker (1993, p. 239), we view translation norm as “typicality” emerging from works in corpus-based lexicography. Translation norms and equivalence can be identified by inspecting not just an individual translation of a linguistic unit, but a set of corpora of the source and the translated texts. Investigating a set of samples enables us to record and quantify which pattern of translation is opted for more frequently (i.e., used repeatedly) in preference to the other patterns for a given unit of analysis in each culture (Baker, 1993, p. 240).

Furthermore, statistical method derived from large corpus underlies the development of data-driven, machine translation (Kenny, 2020). The quantitative information concerning the constructional patterns (e.g., argument structures) of a verb derived from a parallel corpus can be used as input-datasets for building translation models by the computational machine learning techniques implemented in statistical and neural-network machine translation (Kenny, 2020, pp. 307–308). The “machine” (i.e., computational method) learns the translation model by means of the distributional patterns/tendency (e.g., argument structure, construction types) that a given verb appear in the corpus. This approach is in line with the usage-based approach to meaning (originally promoted in Wittgenstein, 1953; and Firth, 1957) that views the meaning of a linguistic form is constructed based on its morphosyntactic, semantic, and lexical co-occurrence properties (see Stefanowitsch, 2010, pp. 368–370, for the overview). This idea of meaning as co-occurrence networks of linguistic forms is adopted in Catford’s proposal for the theory of meaning in translation called “the total network of relations entered into by any linguistic form” (Catford, 1965, p. 35). One example of the “relations” of a linguistic form relevant in this study is the “formal relations” of the form, such as the grammatical construction of the verb.

To summarise, corpus-based, quantitative data derived from a sample of translated texts in a parallel corpus allows us to measure such central notions as equivalence and norms in translation studies. This paper is an attempt to provide a modest example of the study of equivalence and norms at the level of grammatical, syntactico-semantic construction of verbal synonyms.

### 1.2 *Overview on ROB and STEAL*

Dispossession verbs namely ROB and STEAL (hereafter R&S) have attracted recent interests in

the context of cross-linguistic studies of their syntactico-semantic behaviours, especially in English (Goldberg, 1995; Glynn, 2004; Thorgren, 2005; Stefanowitsch, 2011; Dux, 2011), Spanish and French (Enghels & Wylin, 2015; Fernández-Martínez & Faber, 2020), and German (Dux, 2018). These verbs represent three-participant events (cf. below) that can involve alternation phenomena, such as in the domain of transfer (i.e., the ditransitive vs. prepositional dative alternations) (Enghels & Wylin, 2015).

Construction Grammar (CxG) theory (Goldberg, 1995, 2013; Hilpert, 2020) has incorporated a central apparatus in Frame Semantics theory (Fillmore, 2014) to provide a unified model in capturing the meaning of verbs; this apparatus is the concept of *semantic frame*. A semantic frame is a rich, encyclopaedic knowledge of “situation type describable in terms of the kinds of relations, situations or sub-events ‘evoked’ in the minds of those who know the language” when they use or hear lexical items belonging to a given frame (Fillmore, 2014, p. 126). In Frame Semantics, a semantic frame consists of the so-called *Frame Elements*, which represent semantic participant roles in the situation related to the frame (Fillmore, 2014, p. 123; Thorgren, 2005, pp. 3–4). From the perspective of CxG, therefore, the meaning of verbs is defined relative to the evoked semantic frames and frame elements.

In the case of R&S, following what is presented in the *FrameNet* repository, ROB evokes the so-called *Robbery* frame while STEAL evokes the *Theft* frame (but see Thorgren, 2005, pp. 6–7 that groups R&S under the broader domain of POSSESSION since the two verbs evoke a situation of taking someone’s possession without permission). Even though R&S belong to different semantic frames, their meanings involve the same configuration of participant roles: THIEF (or PERPETRATOR), GOODS, and TARGET (i.e., VICTIM or locational SOURCE) (Goldberg, 1995; Stefanowitsch, 2011). Goldberg (1995, pp. 44–45) further notes that R&S differ in terms of which of these participant roles are lexically profiled (i.e., obligatorily expressed) in the utterances with R&S. ROB profiles the TARGET and syntactically realise this role as the direct object, while STEAL profiles the GOODS (Goldberg, 1995, pp. 45–46; Dux, 2011, p. 22; Stefanowitsch, 2011, p. 263). Further qualitative difference between R&S proposed is that ROB exerts a high degree of negative affectedness on the TARGET/VICTIM while STEAL does not (Goldberg, 1995, p. 46).

This paper will further assess the proposed difference in profiling of R&S using different genres of data from the previous studies (see §3.2). These constructional profiles of R&S will then be taken as the linguistic domain for which the verbs’ equivalence will be measured in the Indonesian translations (§3.2.1 and §3.3), following Toury’s suggestion to focus on measuring the degree of equivalence.

## 2 Methods

### 2.1 The corpus data

The data for this study comes from the *OpenSubtitles v2018* (hereafter OSub) parallel corpora (Lison & Tiedemann, 2016), especially the English-Indonesian sub-corpus. The whole OSub parallel corpora are built out of a large, open-source database of movie and TV subtitles. The latest version of OSub (v2018) contains over two billion sentences across sixty-two languages. The English-Indonesian sub-corpus used in this study consists of 9.7 million aligned sentences between English (as the source texts) and its Indonesian translation (as the target texts). The sentence in each language holds in total 72.8 million word-tokens for English and 60.9 million word-tokens for Indonesian. We downloaded the English-Indonesian corpus file in .tmx format (1.31 GB) the content of which is shown in Figure 1.

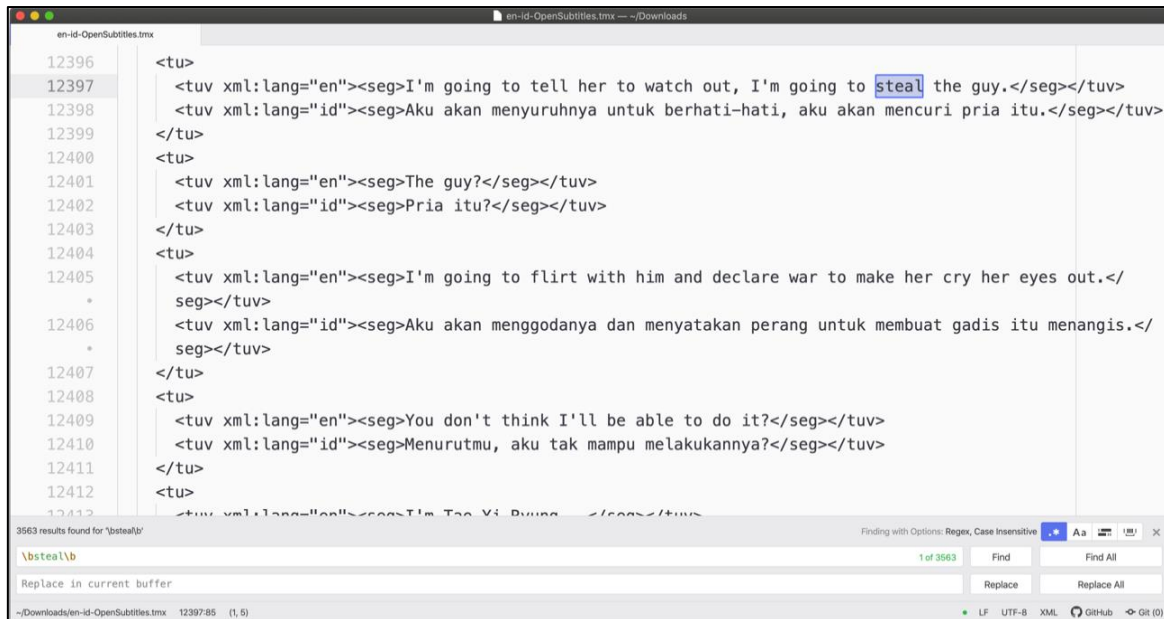


Figure 1 Snippet of the English-Indonesian OSub corpus in .tmx format.

Figure 1 shows that (i) each English sentence appears immediately above the Indonesian translation and (ii) each sentence is tagged with language identifier (i.e., `<tuv xml:lang="en">` for the English sentences and `<tuv xml:lang="id">` for the Indonesian translations). We designed a programming script with R (R Core Team, 2020; see Rajeg et al., 2021b, for the R script) to separate the English and Indonesian sentences in the .tmx file into two plain text files (.txt). These plain text files then became the input data for generating a parallel concordance (i.e., parallel Keyword-in-Context display) (see Figure 4). The R programming script maintains the alignment between the source English sentences and their target Indonesian translations in the two files (see Figure 2).

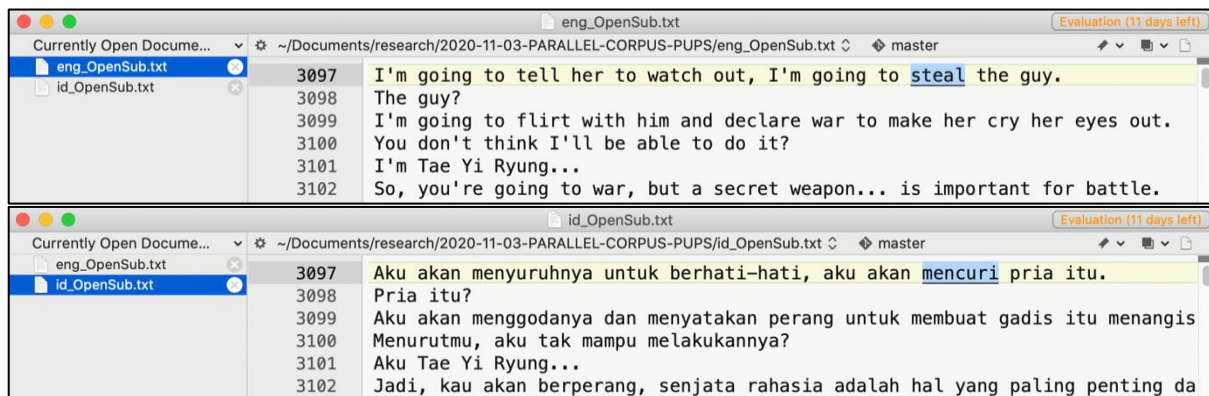


Figure 2 Snippet of the two plain-text files for the English and Indonesian sentences split up from the .tmx file format.

One can see from Figure 2 that, for example, the English sentence number 3097 in the `eng_OpenSub.txt` file (i.e., *I am going to steal the guy*) corresponds to its Indonesian translation in sentence number 3097 in the `id_OpenSub.txt` file (i.e., *aku akan mencuri pria itu*). This is what is meant by sentence-aligned parallel corpora. From these two files, we retrieved sample sentences for various inflectional forms of ROB and STEAL (R&S) and their corresponding Indonesian translations.

## 2.2 Data retrieval

The data for this study represent the English sentences (and their Indonesian translations) containing the various inflectional verbal forms referring to the lemma (i.e., abstract form of) R&S

(Figure 3). The occurrences of the verb forms referring to ROB amount to 1,870 tokens, which is highly significantly lower than the total occurrences of the verb forms for STEAL, amounting to 10,715 tokens ( $X^2_{\text{goodness of fit}} = 6216,5$ ;  $df = 1$ ;  $p < 0.001$ ). The ratio of ROB to STEAL is 1:6.

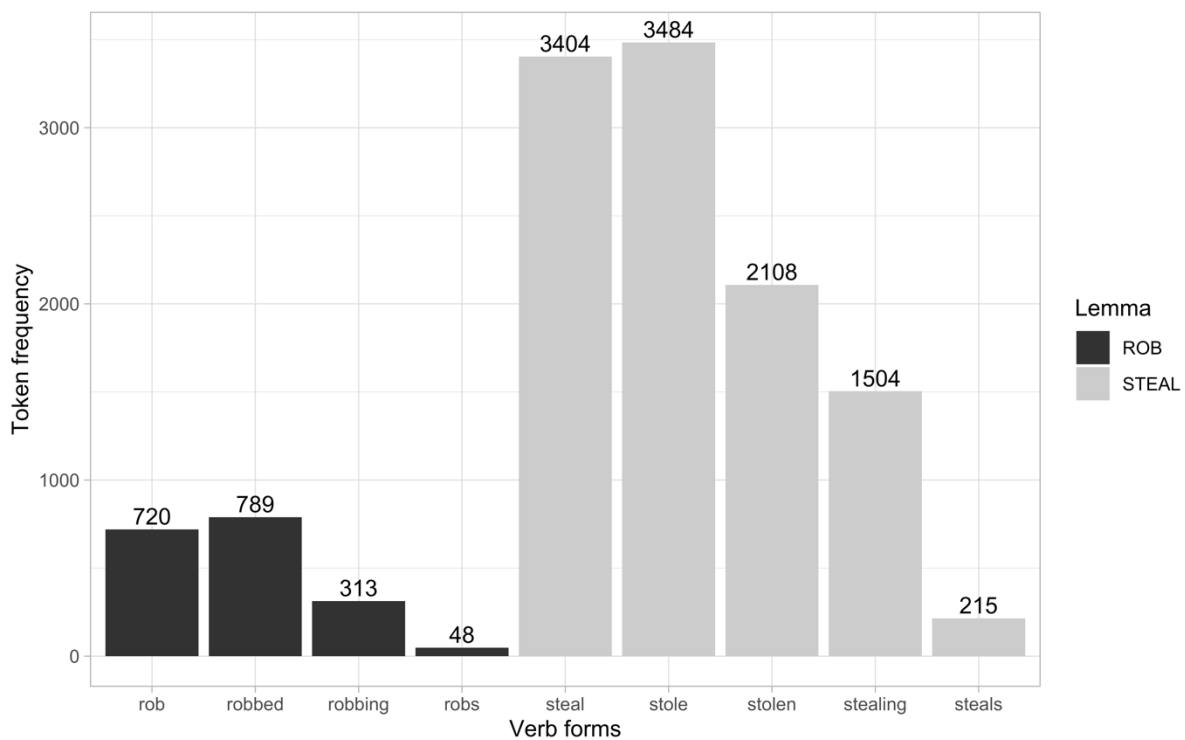


Figure 3 Token frequencies of the inflectional forms of ROB and STEAL in the English-Indonesian OSub sub-corpus.

We analysed two sets of data: (i) one is a random sample of 150 sentences with forms referring to the lemma ROB and the corresponding 150 sentences of their Indonesian translations (300 sentences in total for ROB), and (ii) another random sample of 150 sentences with forms referring to the lemma STEAL and the corresponding 150 sentences of the Indonesian translations (300 sentences in total for STEAL). In total, we analysed 600 sentences from the two languages.

The concordance (Keyword-in-Context display) technique in corpus linguistics is used to bring the data in a more accessible, tabular format. We generated parallel concordances for the R&S samples whereby their forms become the central keywords surrounded by their preceding (left) and following (right) contexts; the Indonesian translations are represented in a separate column (see Figure 4). We used the `para_conc()` function from the `paracorp` R package (Rajeg, 2021a) designed to create parallel concordances from parallel corpus inputs. The output of `para_conc()` is a tab-separated plain text file that can be imported into spreadsheet software.

CASES	LEFT	NODE	RIGHT	TRANSLATION
1	Because everybody else	robs	banks.	Karena orang lain merampas bank.
3	Mr. A always	robs	in disguise..	Mr. A selalu ngerampok dengan menyamar..
6	Don't confuse that with	robbing	, which means the removal of	Jangan bingung bahwa dengan merampok, ya
8	And you're accusing them of	robbing	?	Dan kalian menuduh mereka merampok?
10	Well, like	robbing	a bank?	Yah, seperti merampok bank?
11	g about a bank robbery and actually	robbing	a bank.	Tapi ada sedikit perbedaan antara mendengar
12	We're	robbing	a casino?	Kita akan merampok sebuah kasino?
13	Got sent to the rock for	robbing	a grocery store.	Dikirim untuk merampok sebuah toko kelontor
14	Bad enough you got us	robbing	a military base, now we got to	Cukup buruk kau buat kita merampok pangkal
18	He was just	robbing	banks then.	Saat itu ia hanya merampok bank.
20	I think it's a miracle she's not	robbing	banks with the Symbionese Li	Aku pikir adalah sebuah keajaiban dia tidak
23	And all I was doing was	robbing	Clem of whatever life she was	Dan yang aku lakukan adalah merampas Clem
25	Somebody's	robbing	from us?	Seseorang merampas uang kami?
26	4, and that I'm somehow deliberately	robbing	him of glory.	Dia pikir blokade adalah bagian dari imajinas
29	Yo, I'm not	robbing	no banks with you guys, all rig	Yo, aku tak akan merampok bank bersama kal
38	We	robbing	the bank.	Kami merampok bank.
47	"- Are you	robbing	us?	"- Apakah Anda merampok kami?

Figure 4 Snippet of the parallel concordance for ROB generated by the `para_conc()` function in the `paracorp` R package (Rajeg, 2021a) and viewed in the LibreOffice Spreadsheet.

The parallel concordance is not the results of data analyses. It is only a facet of the data collection procedures. The concordance organises the data in a format that facilitates qualitative and quantitative analyses in relation to the research problems, which are the topic of the following sub-sections.

### 2.3 Aspects of data analysis

In this sub-section, we outline the qualitative and quantitative aspects of the data analyses on the parallel concordances for R&S and their Indonesian translations.

#### 2.3.1 Annotated qualitative variables

The qualitative aspect of the analyses includes manual annotations of the syntactic, semantic, and lexical features for each use of R&S in the concordance samples (see Table 1 for the summary). At the lexical level, we identified the Indonesian words and word classes used as the translation of R&S in each sample sentence; this is not discussed in more detail here (§3.1) but in another publication (Rajeg et al., 2021a). Moreover, we manually extracted the syntactic collocates of R&S and their translations. These collocates represent lexical items filling in the participant roles of R&S in the constructions; these roles are THIEF, GOODS, and TARGET; the TARGET role can be HUMAN/VICTIM or LOCATION (Goldberg, 1995, p. 46; Stefanowitsch, 2011, p. 260; Dux, 2011, pp. 18, 22).

At the constructional level, we focused on two main analytical aspects. The first one is the overt expression of the three participant roles of R&S (and their translations) in the samples (§3.2). If each participant role was linguistically encoded (i.e., explicitly mentioned) and profiled (i.e., mapped onto the core syntactic arguments) in the sentence, it was coded as TRUE, otherwise, it was coded as FALSE (Stefanowitsch, 2011, pp. 262–263). This analysis relates to the research problem concerning the profiled participant roles of R&S and the degree to which such profiling is maintained in the Indonesian translations (see §3.2.1).

The second aspect of the constructional analysis is the type of constructions of R&S (and their Indonesian translations) (§3.3). The construction type, in this case, captures the mapping of the participant roles, especially GOODS and TARGET, onto the core arguments of the two verbs, especially the direct object (in an active sentence) and subject (in a passive sentence) arguments. The focus on these two participant roles and their mappings onto the two arguments in two different grammatical voices is motivated by previous works on R&S arguing that the two verbs differ regarding which participant (esp. GOODS and TARGET) is profiled (i.e., mapped onto the direct object argument) and which one can be unexpressed explicitly in syntax (Goldberg, 1995; Stefanowitsch, 2011; Dux, 2011; Glynn, 2004). Consider the following attested examples of R&S in active sentences found in British English online writings (e.g., mainly online news and blogs).

- (1) a. Cloughie *robbed* [the rich]<sub>TARGET</sub> of [their trophies]<sub>GOODS</sub> ...<sup>1</sup>  
 b. two others then assaulted and *robbed* [a bus driver]<sub>TARGET</sub> while they were getting off ...<sup>2</sup>
- (2) a. Verstappen, 24, *stole* [the world championship]<sub>GOODS</sub> from [rival Lewis Hamilton]<sub>TARGET</sub> ...<sup>3</sup>  
 b. Brazen valet *stole* [the cars]<sub>GOODS</sub> he cleaned.<sup>4</sup>
- (3) ... a surgeon *robbed* [Robert Burn's grave]<sub>TARGET</sub> and *stole* [his skull]<sub>GOODS</sub> in the name of science.<sup>5</sup>

In (1)a, the TARGET role of ROB is profiled as it is mapped onto the direct object (a core argument), while the GOODS is expressed in an adjunct prepositional phrase and can be left out completely as shown in (1)b. Similarly, in (2), STEAL profile the GOODS as it is mapped onto the direct object ((2)a) while the TARGET can be optional ((2)b). Example (3) provides a clear contrast between ROB and STEAL in terms of which participant role they tend to profile, given these verbs now co-occur in the same sentence: ROB profiles the locational TARGET (*Robert Burn's grave*), with the human TARGET is expressed in possessive/genitive 's modifier (i.e., *Robert Burn's*), while STEAL profiles the GOODS (*skull*) associated with the locational TARGET. In the database, we labelled these construction types as "TARGET-OBJECT" and "GOODS-OBJECT" (for the active voice usage of R&S), and "TARGET-SUBJ-PASS" and "GOODS-SUBJ-PASS" (for the passive voice usage of R&S). We are interested in measuring the degree to which these constructions for the TARGET and GOODS are preserved in the Indonesian translations. In a way, this aspect is like what Baker (2017) calls "equivalence above the word level".

The final level of analysis is semantics and related to the two roles of interest (GOODS and TARGET), namely (i) the animacy of the TARGET, and (ii) the semantic types of the GOODS and TARGET roles. In terms of the animacy of the TARGET, we determined as to whether the TARGET is (i) animate and/or sentient, hence showing a more specific role of VICTIM, or (ii) location, thus LOCATION TARGET. The motivation behind this analysis is Goldberg's (1995, pp. 46–47) proposal that ROB semantically exerts a higher degree of negative affectedness on the TARGET, suggesting the higher likelihood of animate and/or sentient TARGET. As for the semantic types coding, we adopted the categorisation proposed by Fernández-Martínez and Faber (2020). This semantic annotation is motivated by the hypothesis that the GOODS of ROB tend to be of high value and precious, which is not always the case for the GOODS of STEAL (Dux, 2018, p. 325). The semantic analyses will not be discussed in this paper but are part of the upcoming one related to this project (see §4 below).

Table 1 Summary of the qualitative variables manually annotated in the sample.

No	Qualitative variables	Notes
1	TRANSLATION	Words for the Indonesian translations for R&S; NA when it is not translated
2	LEX_THIEF_ENG	Lexical item for the collocates of R&S filling the THIEF role
3	LEX_TARGET_ENG	Lexical item for the collocates of R&S filling the TARGET role
4	LEX_GOODS_ENG	Lexical item for the collocates of R&S filling the GOODS role
5	LEX_THIEF_IDN	Lexical item for the THIEF collocates of the Indonesian translation
6	LEX_TARGET_IDN	Lexical item for the TARGET collocates of the Indonesian translation
7	LEX_GOODS_IDN	Lexical item for the GOODS collocates of the Indonesian translation
8	WORD_CLASSES	Word classes of the Indonesian translations of R&S (e.g., VERB, NOUN)
9	ENG_THIEF	Is the THIEF role explicitly mentioned in the English data? TRUE/FALSE
10	ENG_TARGET	Is the TARGET role explicitly mentioned in the English data? TRUE/FALSE
11	ENG_GOODS	Is the GOODS role explicitly mentioned in the English data? TRUE/FALSE

<sup>1</sup> <http://nottslit.blogspot.com/2015/08/nottingham-rebel-city.html> (last accessed 14 December 2021).

<sup>2</sup> <https://www.glasgowlive.co.uk/news/glasgow-news/govan-teen-led-cops-12-22438578> (last accessed December 2021).

<sup>3</sup> <https://www.dailymail.co.uk/sport/formulaone/article-10303789/F1-news-LIVE-Mercedes-plan-appeal-Max-Verstappen-crowned-champion.html> (last accessed 14 December 2021).

<sup>4</sup> <https://www.cambs.police.uk/news-and-appeals/Brazen-valet-stole-the-cars-he-cleaned> (last accessed 14 December 2021).

<sup>5</sup> <https://www.thecourier.co.uk/fp/news/scotland/1861537/robert-burns-the-remarkable-night-a-surgeon-robbed-the-national-bards-grave-and-stole-his-skull/> (last accessed 14 December 2021).

No	Qualitative variables	Notes
12	IDN_THIEFS	Is the THIEF role explicitly mentioned in the Indonesian data? TRUE/FALSE
13	IDN_TARGET	Is the TARGET role explicitly mentioned in the Indonesian data? TRUE/FALSE
14	IDN_GOODS	Is the GOODS role explicitly mentioned in the Indonesian data? TRUE/FALSE
15	CxN_ENG_AV	Construction types of R&S in Active voice (e.g., TARGET-OBJ, GOODS-OBJ)
16	CxN_ENG_PASS	Construction types of R&S in Passive voice (e.g., TARGET/GOODS-SUBJ-PASS)
17	CxN_IDN_AV	Construction types of R&S's translations in Active voice
18	CxN_IDN_PASS	Construction types of R&S's translations in Passive and Objective voices
19	ANIMACY_ENG_TARGET	(In)animacy of the TARGET role of R&S in the English source text
20	ANIMACY_IDN_TARGET	(In)animacy of the TARGET role of R&S in the Indonesian target text
21	SEM_THIEF_ENG	Semantic type of the THIEF collocates of R&S
22	SEM_TARGET_ENG	Semantic type of the TARGET collocates of R&S
23	SEM_GOODS_ENG	Semantic type of the GOODS collocates of R&S
24	SEM_THIEF_IDN	Semantic type of the THIEF collocates of the Indonesian translation
25	SEM_TARGET_IDN	Semantic type of the TARGET collocates of the Indonesian translation
26	SEM_GOODS_IDN	Semantic type of the GOODS collocates of the Indonesian translation

In the parallel concordance table, each of these variables were annotated in separate columns. The fully annotated database would then become the input for the statistical analyses discussed next.

### 2.3.2 Quantitative, statistical analyses

The statistical analyses include both descriptive and analytic/inferential statistics (Gries, 2009, Chapters 3 and 4). As we primarily deal with qualitative/categorical variables, the descriptive statistics can be (i) frequency count of one categorical variable of interest (univariate statistics) (Gries, 2009, pp. 96–97) or (ii) co-occurrence frequency count between two categorical variables (bivariate statistics) (Gries, 2009, pp. 127–130). The inferential statistics include significance tests for two scenarios: (i) univariate categorical distribution, such as the *Chi-Square Goodness-of-Fit* test (Gries, 2009, p. 125) or the *Multinomial/Binomial* tests (Gries, 2009, pp. 41–42, 158) when the pre-requisites for Chi-Square test are not met; and (ii) bivariate categorical distribution, such as the *Chi-Square Test for Independence* (Gries, 2009, pp. 165–175); when the pre-requisites for Chi-Square test are not met, the *Fisher-Yates Exact* (FYE) test are used (Stefanowitsch & Gries, 2003, p. 218; Gries, 2009, p. 177; Stefanowitsch, 2013, p. 305; for FYE application with Indonesian data, see Rajeg & Rajeg, 2019).

As mentioned in §1, this paper measures the degree of the constructional equivalence of R&S in Indonesian, involving equivalences at the profiled participant roles (§3.2 and §3.2.1) and construction types (§3.3). To measure the degree of equivalence for the profiled participant roles (variables 9-14), for instance, we cross-tabulate how many times does the profiling of a role (e.g., THIEF) in English is maintained in its Indonesian translation (i.e., ENG\_THIEF: TRUE and IDN\_THIEF: TRUE) vs. the de-profiling of the role in the translation (i.e., ENG\_THIEF: TRUE and IDN\_THIEF: FALSE), and vice versa (e.g., the de-profiling of a role in English but profiled in Indonesian). These distributions were evaluated for statistical significance (using the Chi-Square Test for Independence or the FYE test) and for the size of the effect (Cramer's *V* or Phi-Coefficient).

For the second aspect, namely measuring the equivalence at the construction-type level (see variables 15-18), the same approach is adopted as in the profiled participant roles. For instance, given the use of the TARGET-OBJECT construction for ROB in English (cf. example (1)a), to what extent would this same construction be used in the Indonesian translation compared to the other construction types? To answer this question, we run a series of FYE tests á la “Constructional Analysis” (Stefanowitsch, 2013; Stefanowitsch & Gries, 2003) for each construction-type co-occurrence in English and Indonesian to determine which construction-type co-occurrence is strongly preferred between the source English text and the Indonesian translations of R&S.

The complete annotated databases, and the R programming codes for the statistical analyses and



visualisations, are available open access (see Rajeg, 2021b, for the download URL). The following R packages are also used for the study: `tidyverse` (Wickham et al., 2019), `vcd` (Meyer et al., 2006), `readxl` (Wickham & Bryan, 2019), and `RColorBrewer` (Neuwirth, 2014).

### 3 Results and discussion

#### 3.1 Lexical equivalence of ROB and STEAL

This subsection briefly reports on the norms for the Indonesian lexical equivalences for word forms based on the lemma ROB (Table 2) and STEAL (Table 3), which have been part of a related publication (see Rajeg et al., 2021a, for further discussion on the choice of certain equivalences). The numbers inside the parentheses are the token frequencies of a given Indonesian form as the translation of the English form (e.g., *merampok* is used in 45 occurrences for the translation of the form *rob*).

Table 2 Indonesian lexical equivalence for word forms of the lemma ROB

English form	Indonesian translation
rob	<i>merampok</i> (45); <i>mencuri</i> (4); <i>rampok</i> (2); <i>merebut</i> (1); <i>merampas</i> (1)
robbed	<i>dirampok</i> (28); <i>merampok</i> (26); <i>mencuri</i> (4); <i>perampok</i> (1); <i>merampas</i> (1); <i>mengambil</i> (1); <i>kerampokan</i> (1); <i>kemalingan</i> (1); <i>dicopet</i> (1)
robbing	<i>merampok</i> (14); <i>merampas</i> (2)
robs	<i>ngerampok</i> (1); <i>merampas</i> (1)

Table 3 Indonesian lexical equivalence for word forms of the lemma STEAL

English form	Indonesian translation
steal	<i>mencuri</i> (44); <i>pencurian</i> (1); <i>merenggut</i> (1); <i>merebut</i> (1); <i>dicuri</i> (1); <i>curi</i> (1)
stealing	<i>mencuri</i> (15); <i>pencurian</i> (1); <i>pencuri</i> (1); <i>curi</i> (1)
steals	<i>mencuri</i> (4)
stole	<i>mencuri</i> (36); <i>curi</i> (5); <i>merampas</i> (1); <i>mengambil</i> (1); <i>menculik</i> (1); <i>dicuri</i> (1)
stolen	<i>dicuri</i> (14); <i>mencuri</i> (3); <i>curian</i> (3); <i>curi</i> (2); <i>hilang</i> (1)

Overall, we have found that the word forms of the lemma ROB in English are predominantly translated into Indonesian with the different forms of the lemma *RAMPOK* (compared to the other attested lemmas) in the sample ( $p_{\text{goodness-of-fit}} < 0.001$ ) (see Rajeg et al., 2021a, pp. 083–3 for details); meanwhile, the word forms of the lemma STEAL are more often translated into Indonesian with the different forms of the lemma *CURI* ( $p_{\text{goodness-of-fit}} < 0.001$ ) (see Rajeg et al., 2021a, pp. 083–084, Gambar 3, for details). These statistical tendencies suggest that the norm, or the default choice, of the equivalences for ROB and STEAL appears to be lexically determined (i.e., a given lemma such as ROB is strongly associated with the lemma *RAMPOK* rather than with the other lemmas in the Indonesian translation).

#### 3.2 Profiled participant roles of ROB and STEAL

We adapted Stefanowitsch's (2011, p. 262) corpus-based approach to investigate the profiled participant roles of R&S. If a given role is explicitly mentioned in the sentence and mapped onto the core syntactic argument of the verbs (i.e., the roles are lexically profiled in the sense of Goldberg, 1995, pp. 44–45), that role received TRUE coding, otherwise it was coded as FALSE. Consider the following examples.

- (4) [I]<sub>THIEF</sub> don't *rob* [banks]<sub>TARGET</sub>.
- (5) Got sent to the rock for *robbing* [a grocery store]<sub>TARGET</sub>.
- (6) Fucking *robbed* [the store]<sub>TARGET</sub> for [sixty some odd dollars]<sub>GOODS</sub>.

In (4), the THIEF (*I*) and the locational TARGET (*banks*) are explicitly mentioned and profiled; hence the THIEF and TARGET were coded as TRUE while the GOODS received FALSE coding. In contrast, example (5) does not explicitly mention the THIEF and the GOODS, but only the locational TARGET (*grocery store*), which is also profiled at the same time as the direct object of ROB. Example (6) also omits the THIEF but mentions the GOODS (*dollars*) in a non-core argument slot (i.e., prepositional phrase) and the locational TARGET (*store*). We calculated the frequency of explicit mention for each role in the core syntactic arguments. Which participant role is strongly profiled by R&S? Is there an association between a certain role and one of the two verbs, as suggested in the previous studies (Goldberg, 1995; Stefanowitsch, 2011)? Figure 5 summarises the results of this analysis.

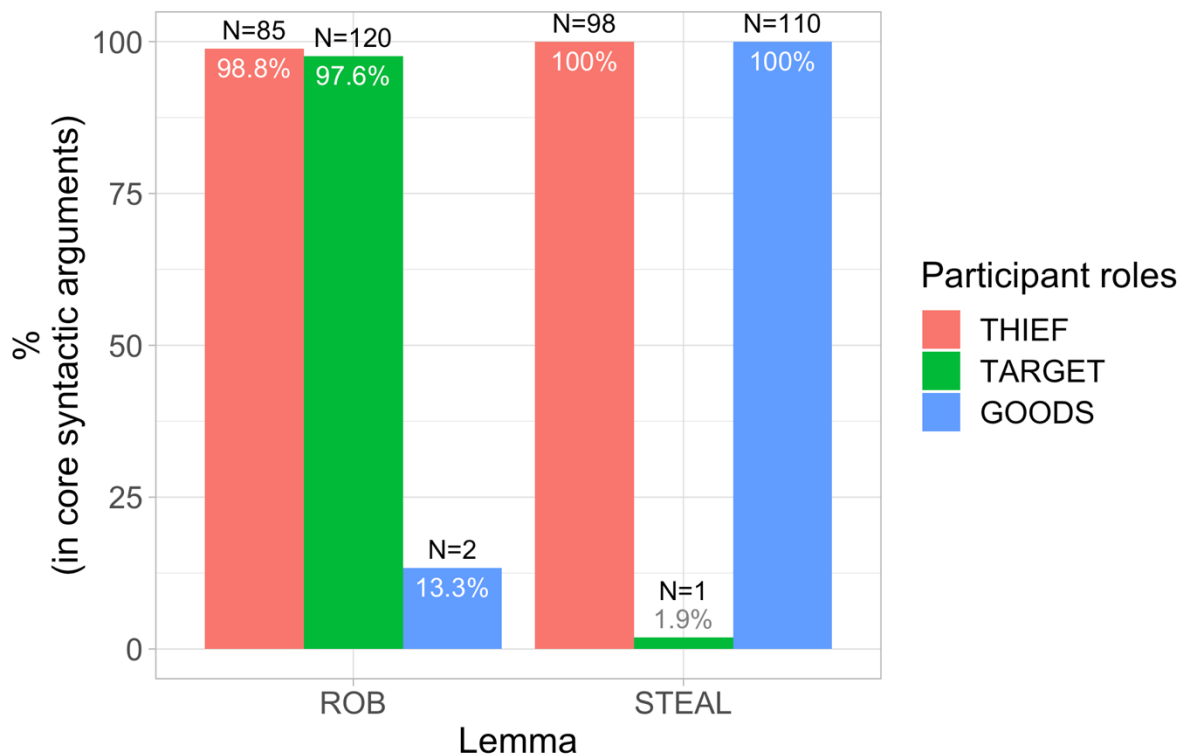


Figure 5 Distribution of the profiled participant roles (y-axis) by verbs (x-axis).

The results shown in Figure 5 are in line with Goldberg's (1995) introspective proposal and Stefanowitsch's (2011) preliminary corpus-based findings on the profiled participant roles for R&S in English. The Chi-Squared test for Independence indicates that there is a highly significant asymmetry ( $X^2=222.09$ ;  $df=2$ ;  $p_{\text{two-tailed}} < 0.001$ ) with strong/robust effect size (Cramer's  $V=0.73$ ) in the profiling of the participant roles of R&S. The specific effects are shown by the stark differences in the profiling of the TARGET (profiled by ROB) and the GOODS (profiled by STEAL); the THIEF, meanwhile, is nearly always profiled by R&S. The association plot in Figure 6 visualises more intuitively the direction and strength of the association between the participant roles and R&S.

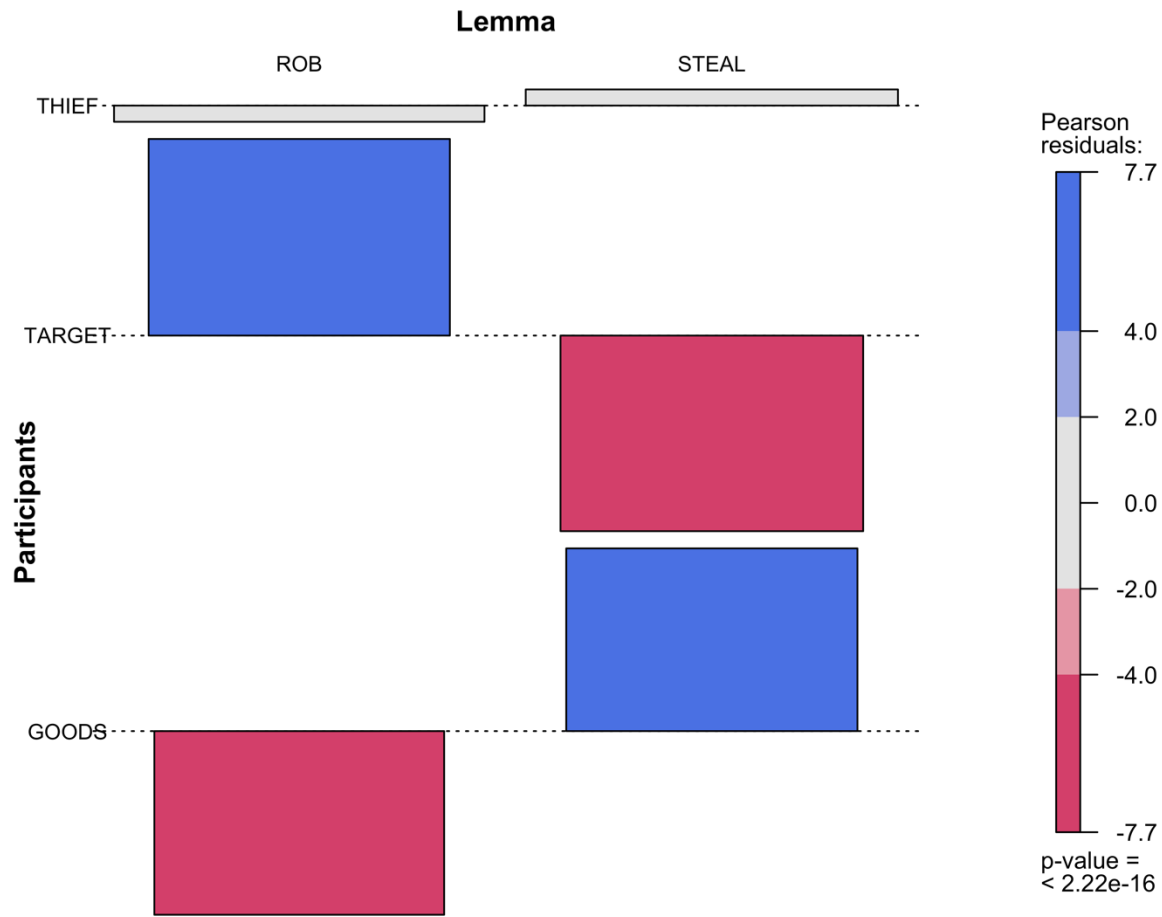


Figure 6 Association plot for the profiled participant roles (rows) and the verb lemmas (columns).

The rising, bluish rectangles indicate positive associations while the falling, reddish rectangles indicate negative associations, the darker the shadings, the stronger the effects. The prominent effects are shown by the strong preferences of ROB to profile the TARGET, and of STEAL to profile the GOODS; these provide further empirical evidence based on movie/TV subtitles corpus for the distinct profiled participant roles of R&S as proposed by Goldberg (1995) and Stefanowitsch (2011). The following sub-sections discuss the degree of equivalence for the profiled participant roles and their Indonesian translations.

### 3.2.1 Degree of equivalence for the profiled participant roles of ROB and STEAL

In §3.2 above we have presented the degree to which a participant role in the R&S semantic frame is profiled, and found strong, positive associations between the TARGET and ROB, and between the GOODS and STEAL (Figure 6). The THIEF role does not show any strong preference as it is nearly equally frequently profiled by R&S. In this sub-sub-section, we present the degree of equivalence for the profiled participant roles of R&S in their Indonesian translations. A participant role is considered equivalence in its profiling if that given role is explicitly mentioned in the core argument functions in the English source text as well as in the Indonesian translations. The greater the frequency of the profiling of a participant role in English and its corresponding Indonesian translation, the greater the constructional equivalence of that participant role. Such a high degree of equivalence may also suggest a low degree of translation shift and loss of information at the level of the participant role of the translated verbs. To illustrate this idea, consider the following examples.

- (7) Eng: [You]<sub>THIEF</sub> stole [him]<sub>GOODS</sub> from [Scott Fischer]<sub>TARGET</sub>  
 Idn: [kamu]<sub>THIEF</sub> **menculik**=**[nya]**<sub>GOODS</sub> dari [Scott Fischer]<sub>TARGET</sub>  
 2SG abduct=3SG from NAME
- (8) Eng: [I]<sub>THIEF</sub>'m somehow deliberately robbing [him]<sub>TARGET</sub> of [glory]<sub>GOODS</sub>.  
 Idn: [gue]<sub>THIEF</sub> entah bagaimana sengaja **merampok** [kemuliaan]<sub>GOODS</sub>  
 1SG unsure how intentionally rob glory

Example (7) represents a constructional equivalence for the participant roles of STEAL. It is because all core argument participant roles of STEAL (the THIEF and GOODS) are preserved and profiled in the Indonesian translation, even though in the Indonesian translation, the GOODS is realised as an encliticised core argument with the third person suffix *-nya*. Moreover, the TARGET role of STEAL expressed in oblique preposition phrase in (7) is also mapped onto the same function in the Indonesian translation. Semantically, the use of STEAL in (7) is metaphorically interpreted as an abduction and thus rendered into *menculik* ‘abduct’ in Indonesian, given that the TARGET is understood as human.

Example (8), in contrast, represents a loss of participant role information. The TARGET, which is explicitly mentioned as the direct object of ROB in the English source text, is absent in the Indonesian translation. With respect to (8), note that *RAMPOK*, which is the norm for the Indonesian lexical equivalence for ROB (Rajeg et al., 2021a), occurs in a different construction pattern whereby it is the GOODS, rather than the TARGET, that is profiled onto the direct object core argument, which is the constructional profile associated with STEAL (see §3.3 for the evidence).

### 3.2.1.1 Degree of equivalence for the THIEF role

First, we look at the degree of equivalence for the profiling of the THIEF role, shown in Figure 7; the left panel is the data for ROB and the right one is for STEAL. Overall, R&S demonstrate a high degree of equivalences in the profiling and de-profiling of the THIEF role. For ROB, 98.8% of the total 85 cases where the THIEF is profiled in the English source text are also profiled in the Indonesian translation (cf. examples (7) and (8)). We can also look at this issue from a different perspective, namely, when THIEF is not profiled. When the THIEF is de-profiled in the English source text, the THIEF is also de-profiled in the Indonesian translations in 94.6% of the 37 cases (see (9)). These correlations are highly significant ( $p_{\text{Fisher-Yates Exact}} < 0.001$ ) and represent robust effects ( $\phi=0.94$ ).

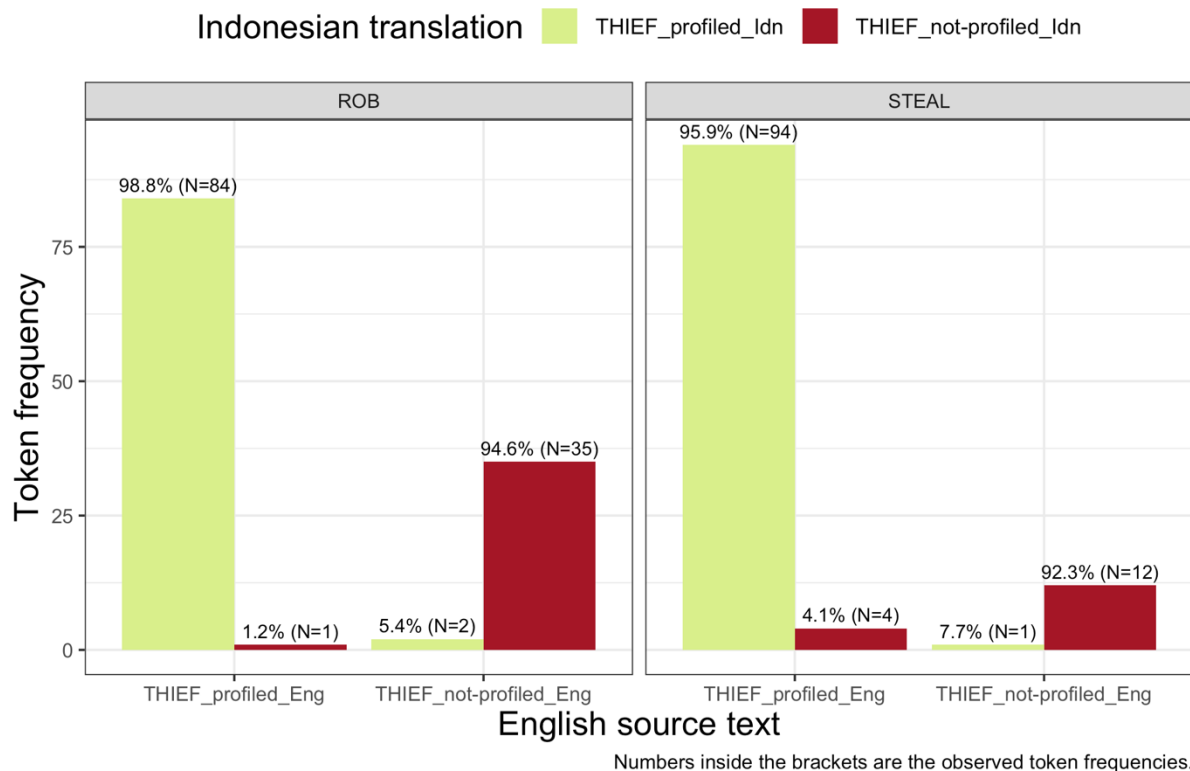


Figure 7 Degree of equivalence for the (de-)profiling of the THIEF role.

- (9) Eng: [I]<sub>TARGET</sub> got *robbed*!  
 Idn: [*Aku*]<sub>TARGET</sub> *di-copet*!  
 1SG PASS-pickpocket

The translation of ROB in (9) is a less prominent verb form other than RAMPOK, namely *copet* ‘to pickpocket’. It might be opted for when the robbery happens as the TARGET, especially VICTIM, is on the street rather than inside a static location. In Indonesian, the event of *copet* happens on the street.

Next, we also identify a high degree of equivalence for the translation of the THIEF role of STEAL. The profiling and the de-profiling, of this role is highly significantly ( $p_{\text{Fisher-Yates Exact}} < 0.001$ ) and robustly ( $\phi=0.81$ ) maintained in more than 90% of all the cases in the translations. The results for THIEF with R&S may suggest that the Indonesian translation attempts to be as faithful as possible as to whether the THIEF role is (de-)profiled in the English source texts.

### 3.2.1.2 Degree of equivalence for the GOODS role

The contrasting distribution in the (de-)profiling of the GOODS in Figure 8 below may not be surprising given the patterns shown in Figure 5 and Figure 6. That is, the GOODS role is strongly profiled by STEAL rather than by ROB. Here, we look at the extent to which these constructional profiling patterns of the GOODS are rendered in the Indonesian translations.

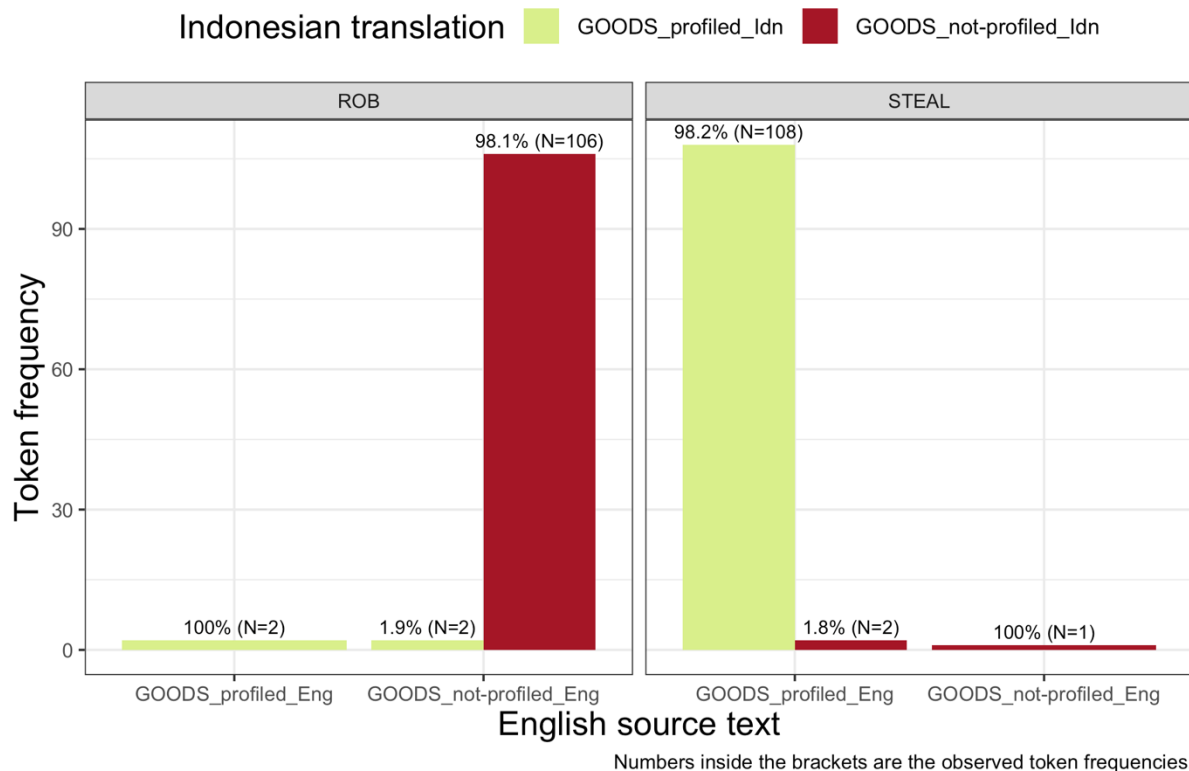


Figure 8 Degree of equivalence for the (de-)profiling of the GOODS role.

From the total 108 cases when the GOODS of ROB is not profiled in the English source text, the Indonesian translations preserve this de-profiling in 98.1% of the cases (see the right maroon bar in the ROB panel) rather than explicitly profiling the GOODS role (1.9%) into the direct object slot (see example (10) below). In example (10), the GOODS (i.e., *uang* ‘money’) is explicitly added and profiled in the direct object slot in the Indonesian translation, meanwhile in the English source text, it is absent (but may already be understood from broader contexts and influenced by the constructional pattern of ROB which rarely profiles the GOODS; see Figure 5).

(10)Eng: [Somebody]<sub>THIEF</sub>'s *robbing* from [us]<sub>TARGET</sub>?  
 Idn: [*seseorang*]<sub>THIEF</sub> *merampas* [*uang*]<sub>GOODS</sub> [*kami*]<sub>TARGET</sub>?  
 someone rob money 1PL.POSS.INCL

On a very few occasions when the GOODS is overtly encoded and profiled by ROB (N=2), the Indonesian translations follow this 100%. Overall, we identify a highly significant ( $p_{\text{Fisher-Yates Exact}} < 0.001$ ) and robust ( $\phi=0.7$ ) correlation between the (non-)profiling of the GOODS in the English source text and in the Indonesian translations, suggesting a high degree of equivalence at the constructional, non-profiled participant level of analysis.

For STEAL, we also identify a significant ( $p_{\text{Fisher-Yates Exact}} < 0.05$ ) and robust ( $\phi=0.57$ ) pattern of correspondence in the profiling of GOODS in the source text and in the Indonesian translations. In 98.2% of the total 110 cases when the GOODS is profiled in the English source text, it is also profiled in the Indonesian translations. As in the results for THIEF in §3.2.1.1 above, the Indonesian translators seem to be faithful with the English source text in terms of the (de-)profiling of the GOODS role.

### 3.2.1.3 Degree of equivalence for the TARGET role

Lastly, we discuss the degree of equivalence for the TARGET role. Like for the THIEF and GOODS roles, we found a high degree of equivalence in the (de-)profiling of the TARGET role in the English source text and in their Indonesian translations.

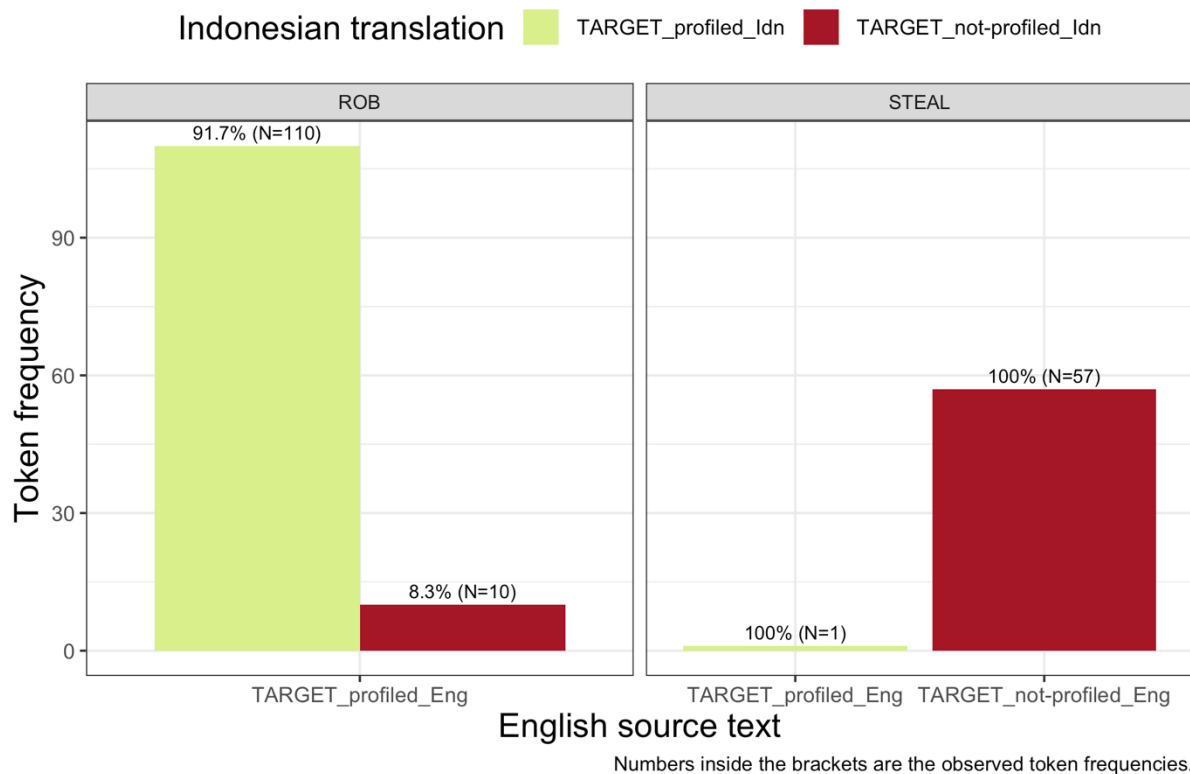


Figure 9 Degree of equivalence for the (de-)profiling of the TARGET role.

Focusing on the panel for ROB on the left, we can see that the TARGET is always profiled in the English source text (cf. examples (1) and (3)), reflecting the profiled participant role of ROB (Figure 5). The Indonesian translations significantly maintain this profiling of the TARGET in 91.7% of the cases ( $p_{\text{Binomial}} < 0.001$ ), suggesting a high degree of constructional equivalence and an attempt to be as faithful as possible to the source text. Furthermore, a significant ( $p_{\text{Fisher-Yates Exact}} < 0.05$ ) and perfect correlation ( $\phi=1$ ) is shown in the (de-)profiling of the TARGET of STEAL. Given these conditions, the Indonesian translations of the TARGET of STEAL demonstrate a high degree of equivalence.

### 3.3 Construction types of ROB and STEAL and their degree of equivalence in Indonesian

This sub-section discusses the degree of equivalence at the construction-type level (see variables 15-18 in Table 1). We focus on the construction types related to the TARGET and GOODS roles that exhibit strong and highly significant interactions with R&S in the English source texts. For instance, given the use of the TARGET-OBJECT construction for ROB in English, to what extent would this same construction be used in the Indonesian translation compared to the other construction types? To answer this question, we run a series of Fisher-Yates Exact tests à la “Collostructional Analysis” (Stefanowitsch, 2013; Stefanowitsch & Gries, 2003) for each construction-type co-occurrence in English and Indonesian. As a result, we can determine which construction-type co-occurrence is strongly maintained between the English source text and the Indonesian translations of R&S (cf. Janda & Lyashevskaya, 2013, for a similar approach to different object of study). Following the standard procedure of multiple significance testing (McDonald, 2014, p. 257), we adjusted the significance level using Holm’s method (Gries, 2009, pp. 242–243, 249) and will discuss the construction-type co-occurrences (between the English source and the Indonesian translations) that are significant at the adjusted level of  $p_{\text{Holm}} < 0.001$ .

### 3.3.1 Constructional equivalence for ROB

In §3.2 we have shown that ROB and STEAL differ with respect to which non-THIEF roles are profiled (i.e., mapped onto the core, syntactic argument). Now, we will specify the construction type for the profiling of these non-THIEF roles and the extent to which a given construction type is rendered into the same construction type in the Indonesian translations. Consider the data for ROB in Figure 10. The construction types for the verbal use of ROB in the sample are shown in the horizontal axis (*x*-axis) while the distribution of the construction types in the Indonesian translation is shown by the bar height (i.e., vertical, *y*-axis) and the different colours.

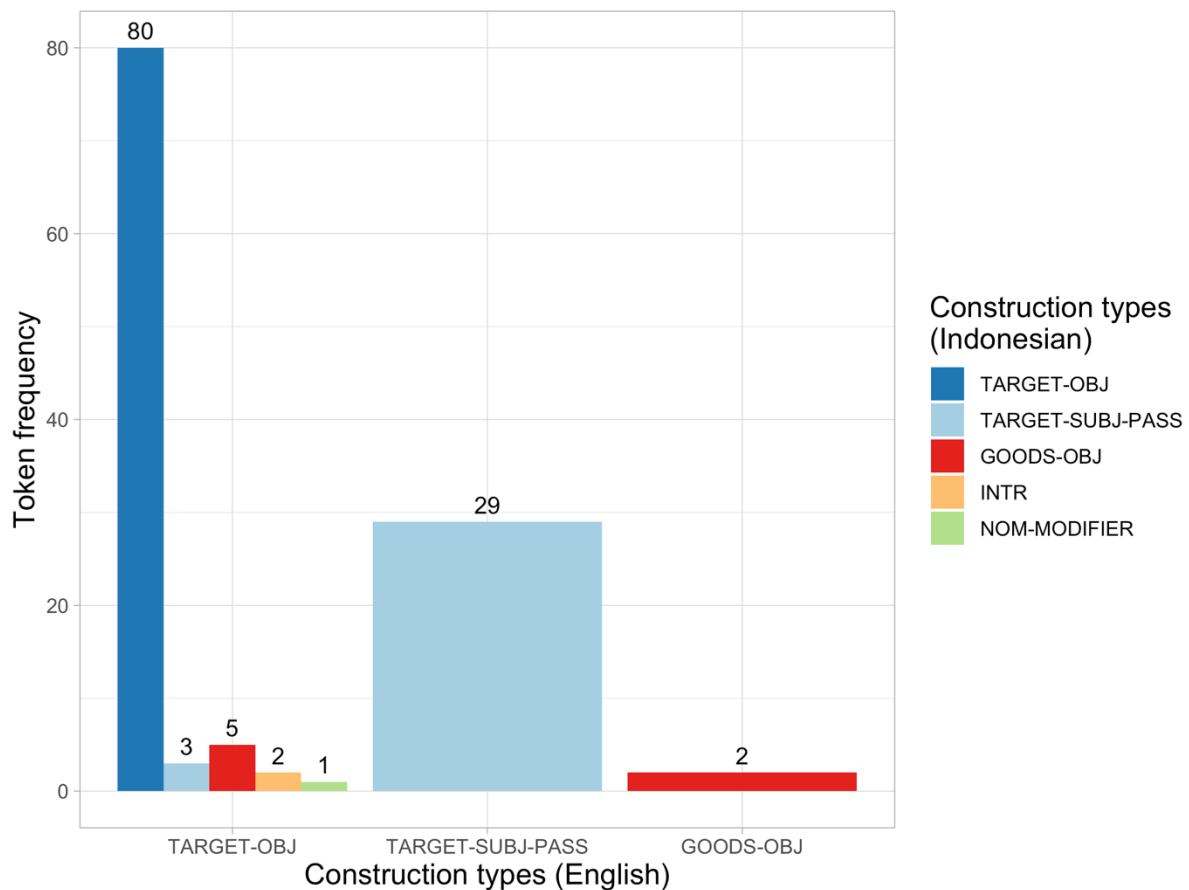


Figure 10 The construction types pertaining to the TARGET and the GOODS roles of ROB in English (*x*-axis) and the distribution of the construction types in the Indonesian translations (*y*-axis and the colour of the bars).

It can be observed that the predominant construction type for ROB in the Active Voice (AV) is the TARGET-OBJ(ECT) ( $N = 91$ ) (cf. example (1)a), followed by the TARGET-SUBJ-PASS ( $N = 29$ ) (example (9)) in the Passive (PASS) Voice. When ROB is used in the TARGET-OBJ construction, there are five different construction types used in the Indonesian translations (cf. the different bar colours for the TARGET-OBJ axis in Figure 10). The norm is that the Indonesian translations would highly significantly ( $p_{\text{Holm}} < 0.001$ ) and strongly ( $\phi = 0.81$ ) preserve the same construction. This suggests a high degree of constructional equivalence for the TARGET-OBJECT construction of ROB in the Indonesian translations.

It is interesting to note that three out of the five occurrences of the GOODS-OBJ construction in the Indonesian translations of ROB from the TARGET-OBJ construction involve three different verbal lemmas, namely *CURI* (11), *RAMPAS* (12), and *AMBIL* (13). These lemmas are not the norm for the lexical equivalence of ROB in Indonesian, namely the lemma *RAMPOK* (Rajeg et al., 2021b). Presumably, these different lemmas exhibit different constructional profile from ROB/*RAMPOK* so that different construction type is used in the Indonesian translations. However, this constructional shift



from the TARGET-OBJ to the GOODS-OBJ construction is not statistically significant. Qualitatively, such shift may be felt necessary by the translators and suitable in the Indonesian contexts.

- (11)Eng: Rivals *robbed* [his bank]<sub>TARGET-OBJ</sub> last year.  
 Idn: *Saingan-nya mencuri [uang-nya] tahun lalu.*  
 rival-3SG.POSS steal money-3SG.POSS year pass
- (12)Eng: Vronsky has *robbed* [me]<sub>TARGET-OBJ</sub> of [my cloak]<sub>GOODS-PP</sub>  
 Idn: *Vronsky telah merampas [jubah] [ku]*<sub>TARGET-POSS</sub>  
 NAME PERF seize/rob cloak-1SG.POSS
- (13)Eng: You forgot they *robbed* [me]<sub>TARGET-OBJ</sub> of [mine]<sub>GOODS-PP</sub>.  
 Idn: *Kau lupa mereka mengambil [pikiran] [ku]*<sub>TARGET-POSS</sub>.  
 2SG forget 3PL take mind-1SG.POSS

Another qualitatively interesting constructional shift, though not statistically significant, is the shift from the TARGET-OBJ (English) into the TARGET-SUBJ-PASS construction (Indonesian). In all three cases, such a shift happens when the TARGET-OBJ construction of ROB in the source text occurs in the Object Relative Clause, namely that the direct object of ROB is the head of a noun phrase modified by the relative clauses headed by ROB. Example (14) illustrates this phenomenon.

- (14)Eng: And [the guy]<sub>TARGET-OBJ-REL</sub> you *robbed*  
 Idn: *Dan [cowok] yang di-rampok*  
 and guy REL PASS-rob

Then, the TARGET-SUBJ-PASS construction in the English sample is always translated with the same construction. The Fisher-Yates Exact test indicates that this is a highly significant ( $p_{Holm} < 0.001$ ) and a very robust pattern ( $\phi=0.94$ ) of constructional equivalence for ROB in Indonesian with respect to the TARGET-SUBJ-PASS construction.

### 3.3.2 Constructional equivalence for STEAL

Figure 5 in §3.2 has demonstrated that STEAL always profiles the GOODS, rather than the TARGET. Figure 11 below visualises the distribution of the construction types in the Indonesian translations of STEAL, especially constructions pertaining to the GOODS and the TARGET roles.

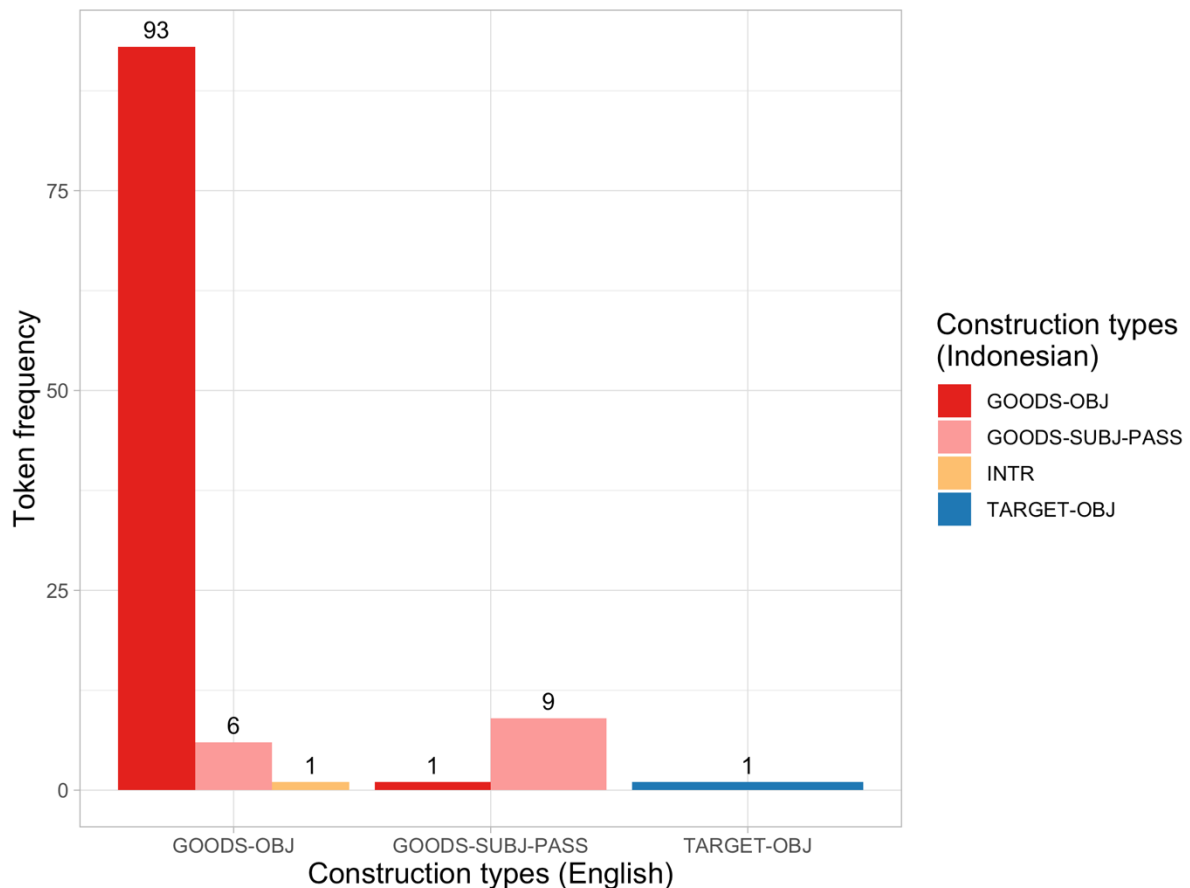


Figure 11 The construction types pertaining to the GOODS and the TARGET roles of STEAL in English (x-axis) and the distribution of the construction types in the Indonesian translations (y-axis and the colour of the bars).

The predominant construction type for STEAL, namely the GOODS-OBJ(ECT) (N = 100) (see example (15)), is highly significantly ( $p_{\text{Holm}} < 0.001$ ) and strongly ( $\phi=0.7$ ) maintained in 93% of all the cases in the Indonesian translations. This also suggests that there is a high degree of constructional equivalence for this construction.

- (15)Eng: ... someone *stole* [a radio]<sub>GOODS-OBJ</sub>  
 Idn: ... *seorang mencuri* [radio]<sub>GOODS-OBJ</sub>  
 someone steal radio

In the remaining seven percent of the cases, the GOODS-OBJ construction type underwent constructional shifts as it is translated using different construction types, namely GOODS-SUBJ-PASS (16) and the intransitive (17) constructions.

- (16)Eng: [The Chemical]<sub>GOODS-OBJ-REL</sub> they stole ...  
 Idn: [*Bahan kimia*]<sub>GOODS-SUBJ-OBJECTIVE-VOICE</sub> *yang mereka curi* ...  
 chemicals REL 3PL OV.steal

- (17)Eng: And you *steal* [things]<sub>GOODS-OBJ-REL</sub>  
 Idn: *Dan kau mencuri?*  
 And you steal

As in the case of ROB, the switch from the AV pattern GOODS-OBJ in English into the PASS GOODS-SUBJ-PASS in Indonesian (6 cases in total), is mostly triggered when the English source text appears in Object Relative Clause construction (4 cases) (16); the other two are interrogative sentences (18). Thus, it is interesting to see similar specific pattern of translation with STEAL.

- (18)Eng: What did he *steal*?  
 Idn: *Apa yang dia curi?*  
 what REL 3SG OV.steal

In example (18), the interrogative pronoun *what* represents the direct object of STEAL linked to the GOODS role. Our use of PASS in the GOODS/TARGET-SUBJ-PASS construction glosses over two passive-like constructions in Indonesian, the canonical passive with the verbal prefix *di-* (14) and the so-called Objective Voice (OV) (Arka & Manning, 2008) with bare, unprefixed verb (16). We did this since we primarily focus on the syntactic function of the GOODS and the TARGET in these two passive-like constructions, namely as the grammatical subject of the clause. From our Indonesian native speaker intuition, the specific case of translating the GOODS/TARGET-OBJ in the English Object Relative Clause into GOODS/TARGET-SUBJ-PASS sound natural and can be also considered equivalent for the structure of the target, Indonesian language.

As a matter of completion, the minor construction GOODS-SUBJ-PASS is also highly significantly ( $p_{Holm} < 0.001$ ) and robustly ( $\phi=0.7$ ) preserved in the Indonesian translations in 90% of the 10 cases of the construction attested in the English sample. In sum, there is an overall high degree of constructional equivalence in the Indonesian translation for the construction types of STEAL.

#### 4 Conclusion

This paper has addressed one aspect of the study of equivalence, namely the degree of translation equivalence. We have integrated (i) a classic yet important idea in translation theory that equivalence is measurable in probabilistic terms, (ii) recent developments in quantitative corpus linguistic method, and (iii) Construction Grammar theory, into a study of equivalence beyond the word level, namely constructional equivalence for the verbal near-synonyms ROB and STEAL. We have discovered an overall high degree of constructional equivalence for the predominant constructional profiles of R&S in English and the way they are translated into Indonesian (§3). More specifically, we have shown these in the (de-)profiling of the participant roles (§3.2.1) and the construction types representing the mapping of the participant roles onto the core syntactic functions (§3.3). Such a high degree of equivalence could be due to the translator is unconsciously confined to the form and meaning of the source language.

We have also shown that the quantitative analyses on the qualitative data allow us to measure certain norms, or typical options, in how the constructional properties of R&S are translated into Indonesian, at least based on data from movies/TV subtitles genre. Furthermore, the publicly shared, annotated database of this study may offer one example to translation researchers on how to handle the parallel concordance data and to document usage properties of the source and the target texts. Our next step from this project is to compare the constructional profile of the Indonesian translation of R&S in the translated Indonesian text vs. the constructional profile found in the original, non-translated Indonesian text. Do the constructional profiles that we found in the translated text resemble those produced NOT in the context of translation from English? The answer to this question could provide a glimpse of the naturalness of the Indonesian translation norms of R&S (cf. §3.1) in their constructional profiles, given the constructional profiles of the same verbs in the non-translated Indonesian corpus.

#### Acknowledgement

The research reported in this paper was funded by the Faculty of Humanities, Udayana University under the grant scheme *Penelitian Unggulan Program Studi* (PUPS) (Grant Number: 023.17.2.677526/2021). The official website of this project can be accessed via <http://bit.ly/gpwr-rs21>. We thank our student members, Putu Dea Indah Kartini and I Gede Semara Dharma Putra, for their aid with some aspects of the data annotation and for being cooperative co-presenters/co-authors in a different paper. We are grateful to (i) two anonymous reviewers for their constructive feedback to improve the paper, as well as (ii) Utei Charaleghy Pamphila and Ni Made Ayu Widiastuti for the discussion and suggestion. None of them is held accountable for any remaining flaws in the paper.

## References

- Arka, I. W., & Manning, C. D. (2008). Voice and grammatical relations in Indonesian: A new perspective. In P. Austin & S. Musgrave (Eds.), *Voice and grammatical relations in Austronesian languages* (pp. 45–69). CSLI.
- Baker, M. (1993). Corpus Linguistics and Translation Studies: Implications and Applications. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and technology: In honour of John Sinclair* (pp. 233–250). Benjamins.
- Baker, M. (2017). *In other words: A coursebook on translation* (Third edition). Routledge.
- Catford, J. C. (1965). *A linguistic theory of translation*. Oxford University Press.
- Dux, R. (2011). *A frame-semantic analysis of five English verbs evoking the Theft frame* [Master's thesis, University of Texas]. <https://repositories.lib.utexas.edu/handle/2152/ETD-UT-2011-05-3114>
- Dux, R. (2018). Frames, verbs, and constructions: German constructions with verbs of stealing. In H. C. Boas & A. Ziem (Eds.), *Constructional Approaches to Syntactic Structures in German* (pp. 367–405). De Gruyter. <https://doi.org/10.1515/9783110457155-010>
- Engiels, R., & Wylin, K. (2015). Expressing the source of dispossession acts in French and Spanish: A contrastive study of *voler* and *robar*. *Languages in Contrast*, 15(1), 102–124. <https://doi.org/10.1075/lic.15.1.06eng>
- Fernández-Martínez, N. J., & Faber, P. (2020). Who stole what from whom? A corpus-based, cross-linguistic study of English and Spanish verbs of stealing. *Languages in Contrast*, 20(1), 107–140. <https://doi.org/10.1075/lic.19002.fer>
- Fillmore, C. J. (2014). Frames, constructions, and FrameNet. In T. Herbst, H.-J. Schmid, & S. Faulhaber (Eds.), *Constructions Collocations Patterns* (pp. 121–166). DE GRUYTER.
- Firth, J. R. (1957). *A synopsis of linguistic theory 1930-55. 1952-59*, 1–32.
- Glynn, D. (2004). Constructions at the crossroads: The place of construction grammar between field and frame. *Annual Review of Cognitive Linguistics*, 2(1), 197–233.
- Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.
- Goldberg, A. E. (2013). Constructionist approaches. In T. Hoffmann & G. Trousdale (Eds.), *The Oxford Handbook of Construction Grammar* (pp. 15–31). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780195396683.013.0002>
- Gries, S. Th. (2009). *Statistics for linguistics with R: A practical introduction*. Mouton de Gruyter.
- Hilpert, M. (2020). Constructional approaches. In B. Aarts, J. Bowie, & G. Popova (Eds.), *The Oxford handbook of English grammar* (First edition, pp. 106–123). Oxford University Press.
- Hu, K. (2016). *Introducing Corpus-based Translation Studies* (1st ed.). Springer Berlin Heidelberg : Imprint: Springer. <https://doi.org/10.1007/978-3-662-48218-6>
- Janda, L. A., & Lyashevskaya, O. (2013). Semantic profiles of five Russian prefixes: *Po-*, *s-*, *za-*, *na-*, *pro-*. *Journal of Slavic Linguistics*, 21(2), 211–258. <https://doi.org/10.1353/jsl.2013.0012>
- Kenny, D. (2001). Equivalence. In M. Baker (Ed.), *Routledge Encyclopedia of Translation Studies* (pp. 77–80). Routledge.
- Kenny, D. (2020). Machine Translation. In M. Baker & G. Saldanha (Eds.), *Routledge encyclopedia of translation studies* (Third edition, pp. 305–310). Routledge.
- Levshina, N. (2015). *How to do Linguistics with R: Data exploration and statistical analysis*. John Benjamins Publishing Company.
- Lison, P., & Tiedemann, J. (2016). Opensubtitles2016: Extracting large parallel corpora from movie and tv subtitles. *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, 923–929.
- McDonald, J. H. (2014). *Handbook of Biological Statistics* (Third). Sparky House Publishing. <http://www.biostathandbook.com/HandbookBioStatThird.pdf>
- Meyer, D., Zeileis, A., & Hornik, K. (2006). The Strucplot framework: Visualizing multi-way contingency tables with *vcd*. *Journal of Statistical Software*, 17(3), 1–48.
- Mikhailov, M., & Cooper, R. (2016). *Corpus linguistics for translation and contrastive studies: A guide for research*. Routledge.
- Neuwirth, E. (2014). *RColorBrewer: ColorBrewer Palettes* [R package version 1.1-2]. <https://CRAN.R-project.org/package=RColorBrewer>
- R Core Team. (2020). *R: A language and environment for statistical computing* [Manual]. <https://www.R-project.org/>
- Rajeg, G. P. W. (2021a). *paracorp: A concordancer for parallel, bilingual corpora* [R package v0.0.1]. Universitas Udayana. <https://doi.org/10.17605/OSF.IO/HV9CU>
- Rajeg, G. P. W. (2021b). *Constructional equivalence in the Indonesian translations of ROB and STEAL* (0.0.1)

- [Computer software]. <https://doi.org/10.17605/OSF.IO/PZC8Y>
- Rajeg, G. P. W., & Rajeg, I. M. (2019). Analisis Koleksem Khas dan potensinya untuk kajian kemiripan makna konstruksional dalam Bahasa Indonesia. In I. N. Sudipa (Ed.), *ETIKA BAHASA Buku persembahan menapaki usia pensiun: I Ketut Tika* (Vol. 1, pp. 65–83). Swasta Nulus. <https://doi.org/10.26180/5bf4e49ea1582>
- Rajeg, G. P. W., Rajeg, I. M., Kartini, P. D. I., & Putra, I. G. S. D. (2021a). Pemanfaatan Bank-Data Digital Dwibahasa dalam Kajian Terjemahan: Studi Kasus Padanan Bahasa Indonesia untuk Verba Sinonim Bahasa Inggris ROB & STEAL. *Prosiding Seminar Nasional Sains Dan Teknologi (SENASTEK) 2021*, 0831–0835. <https://doi.org/10.6084/m9.figshare.17078369>
- Rajeg, G. P. W., Rajeg, I. M., Kartini, P. D. I., & Putra, I. G. S. D. (2021b). *Material pendukung untuk MODEL KAJIAN TERJEMAHAN BERBASIS BANK DATA TERJEMAHAN DIGITAL INGGRIS-INDONESIA DAN IMPLIKASI PEDAGOGISNYA*. <https://doi.org/10.17605/OSF.IO/Y6ESA>
- Stefanowitsch, A. (2010). Empirical cognitive semantics: Some thoughts. In Dylan Glynn & Kerstin Fischer (Eds.), *Quantitative methods in cognitive semantics: Corpus-driven approaches* (pp. 355–380). Mouton de Gruyter.
- Stefanowitsch, A. (2011). Cognitive linguistics meets the corpus. In M. Brdar, S. Th. Gries, & M. Ž. Fuchs (Eds.), *Cognitive Linguistics: Convergence and Expansion* (pp. 257–289). John Benjamins Publishing Company.
- Stefanowitsch, A. (2013). Collostructional analysis. In T. Hoffmann & G. Trousdale (Eds.), *The Oxford Handbook of Construction Grammar* (pp. 290–306). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780195396683.013.0016>
- Stefanowitsch, A., & Gries, S. Th. (2003). Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics*, 8(2), 209–243.
- Thorgren, S. (2005). *Transaction verbs: A lexical and semantic analysis of rob and steal* (Reports from the Department of Language and Culture No. 3; Linguistics in the Midnight Sun). Luleå University of Technology. <http://urn.kb.se/resolve?urn=urn:nbn:se:ltu:diva-56412>
- Toury, G. (2000). The nature and role of norms in translation. In L. Venuti (Ed.), *The Translation studies reader* (pp. 198–211). Routledge.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., ... Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, 4(43), 1686. <https://doi.org/10.21105/joss.01686>
- Wickham, H., & Bryan, J. (2019). *readxl: Read Excel Files* [R package version 1.3.1]. <https://CRAN.R-project.org/package=readxl>
- Wittgenstein, L. (1953). *Philosophical investigations*. Macmillan.
- Zanettin, F. (2014). Corpora in Translation. In J. House (Ed.), *Translation: A Multidisciplinary Approach* (pp. 178–199). Palgrave Macmillan UK. [https://doi.org/10.1057/9781137025487\\_10](https://doi.org/10.1057/9781137025487_10)