

Nanyang Technological University

Crosslingual corpus-based analysis of English and Japanese Verbs

Final Year Project

Name: Sandra Soh Xue Ying

Supervisor: Associate Professor Francis Bond

Date of Submission: 17 November 2014

Acknowledgements

I would like to express my deepest gratitude to:

- Associate Professor Francis Bond, Division of Linguistics and Multilingual Studies, Nanyang Technological University (NTU) for his insightful and patient guidance, as well as his kind words
- My Japanese teachers who have passed on the knowledge of Japanese to me
- Hideka Wada, for her kind advice and discussion on translation issues
- My friends and family for their constant encouragements and support

TABLE OF CONTENTS

TABLE OF ABBREVIATIONS	3
ABSTRACT.....	4
CHAPTER ONE: INTRODUCTION	5
CHAPTER TWO: LITERATURE REVIEW	5
2.1 Contrastive Linguistics and Translation	5
2.2 Translation divergences	6
2.3 Verbs in English.....	7
2.4 Verbs in Japanese.....	9
2.5 Comparisons between Japanese compound verbs and English phrasal verbs	12
CHAPTER THREE: AIMS AND HYPOTHESIS	13
CHAPTER FOUR: METHODOLOGY.....	15
4.1 Resources	15
4.2 Tagging Multiword Expressions in WordNet.....	16
4.3 Annotating relationships shared between words.....	16
4.4 Tagging Issues	20
4.4.1 Polite forms	20
4.4.2 Poor Translations	20
4.4.3 Unexpected discoveries in corpus – concepts that were not linked	22
4.4.4 Tagging polysemous words	23
CHAPTER FIVE: RESULTS AND ANALYSIS.....	24
CHAPTER SIX: DISCUSSION	28
6.1 Analysis of zero matches in the corpora	28
6.2 Analysis of shifts in POS	28
6.3 Analysis of single English verbs and Japanese compound verbs	31
6.4 Analysis of English phrasal verbs and single Japanese verbs.....	32
6.5 Analysis of English phrasal verbs and Japanese compound verbs.....	33
6.5.1 Double Translations	34
6.5.2 Analysis of adverbs in PVs in relation to V ₂ s in VVs.....	36
6.5.3 Analysis of particles in PVs in relation to V ₂ s in VVs.....	36
CHAPTER SEVEN: CONCLUSION.....	40
REFERENCES	41

TABLE OF ABBREVIATIONS

Abbr.	Name	Ex.
NOM	nominative	-ga が
ACC	accusative	-o を
DAT	dative	-ni に/e へ
TOP	topic	-wa は
ADN	adnominal	-no の
FOC	focus	-mo も
QUO	quotative	-to と
FIL	filler(at the end of sentences)	-no の
COP	copula	-desu です
HON	honorifics	-go/o ご・お
ST	source text	
TT	translated text	
SL	source language	
TL	target language	
FSID	sentence ID in source text	
TSID	sentence ID in translated text	
PV	English phrasal verbs	
VV	Japanese compound verbs	
phrE	English idiomatic expressions	
phrJ	Japanese idiomatic expressions	
singleE	Single English verbs	
single	Single Japanese verbs	
POS	part-of-speech	
MWE	multiword expression	

ABSTRACT

This study looked into divergences and correspondences that occurred in the translation of English verbs to Japanese verbs. A quantitative and qualitative approach was adopted in this study to analyze the distribution of different types of verbs in English and Japanese, based on the parallel NTU Multilingual Corpus (NTU-MC). Verb types examined include English phrasal verbs, Japanese compound verbs, as well as idiomatic verbal expressions and single verbs in English and Japanese. English phrasal verbs and Japanese compound verbs were hypothesized to be correlated. Using a Chi-square test, they were proven to be translation correspondences of each other. Other patterns of translation divergences were also detected, such as lexical gaps and lexical deviations. These divergences were attributed to translation style, but more so on differences in the language systems, such as the system of honorifics in Japanese, of which a similar system is not found in English. Findings from this study can shed some light concerning translation issues of English phrasal verbs into Japanese, and also for learners of either language to understand the semantic behaviours of English phrasal verbs and Japanese compound verbs.

CHAPTER ONE

INTRODUCTION

Meaning and concepts are encoded differently in each language system. One method to discover such similarities and differences of two or more languages is through contrastive analysis or systematic comparison of the language systems. This comparison is useful in a theoretical sense, as well as in an applied perspective in lexicography and translation. Translation can be used as a means of establishing cross-linguistic relationships, and these relationships can be reinforced with the help of a multilingual corpus by comparing paired source and target language expressions. Comparisons of translated texts can also help to shed light on how meaning is marked similarly or differently in different language systems.

CHAPTER TWO

LITERATURE REVIEW

2.1 Contrastive Linguistics and Translation

Current contrastive linguistic studies have been found to use parallel corpora to compare between language systems and study corresponding linguistic phenomena in the languages compared (Johansson, 2007, Granger, 2003). A study conducted by Xu and Li (2014) examined English translations of Chinese splittable compounds in a Chinese-English parallel corpus, and studied differences in semantic and grammatical properties of their translation counterparts. In their discussion, they discovered variations in the rendering of argument structures and grammatical properties in the English translation, and there was a tendency for generalized meanings in Chinese verbs to be made specific in English, and for implicit arguments to be expressed overtly in the target language (abbr. TL) (Xu and Li, 2014: 99). This has given insights on how structurally divergent constructions in the TL can arise from the translational processes involved.

Castejon (2012) looked into translations of English gerunds into Spanish by incorporating a corpus-based approach in her study. She managed to establish that the Spanish infinitive is a formal correspondence, or in her paper, noted as “established equivalence” (Castejon, 2012: 128). Another equivalent translation was the Spanish that-clause, which was employed in different contexts. She also found the Spanish substantiative to be a non-corresponding translation counterpart of English gerunds, and she attributed this to another translation technique, known as TRANSPOSITION, which denotes a shift in word category

(Molina and Hurtado Albir, 2002). The English-Spanish parallel corpus was an example of a source with a collection of contexts in which these translation correspondences would occur.

Malá (2013) investigated the translations of English copula verbs into Czech. Due to a lack of functionally similar copula verbs in Czech, it was found that there were a variety of Czech divergent counterparts bearing meanings that were similarly conveyed in English. The semantic meanings conveyed in the counterparts were found to coincide with those of the English verbs, and distinctions between types of translation correspondences could be made with respect to particular copula verbs.

Denturck (2012) in particular focused on two translation techniques, EXPLICITATION and IMPLICITATION to examine how it applies to the translation causal connectives in French and Dutch. EXPLICITATION refers to the addition of information in the translated text (TT) where meaning in the source language (SL) is not explicitly stated, whereas in IMPLICITATION, meaning which is made explicit in the source text (ST) becomes implied in the target language (TL) (Molina and Hurtado Albir, 2002). In a bi-parallel comparison between these two languages, she found that translation shifts due to EXPLICITATION is more common in translations from French to Dutch, as she uncovered more causal conjunctions in Dutch, replacing non-causal constructions in French (Denturck, 2012: 224). Moreover, this trend was similarly reflected in Dutch-French translations, where French translations included causal constructions in cases where the ST did not have causal markers. Evaluating the types of translation techniques used in the translation process thus reveals how translation shifts come about. Translation techniques can thereby account for the appearance of varied translation counterparts of a particular construction in the ST.

The above studies have pointed out translation features such as EXPLICITATION commonly exhibited in TTs. We would like to identify translation features that exemplify translations shifts in our study.

Presently, no corpus-based studies have focused on contrastive analysis between English and Japanese verbs. This study will take a corpus-based approach to examine how verbs are translated between an Indo-European language and an Asian language, specifically English and Japanese.

2.2 Translation divergences

In translation, maintaining the main essence of the ST is necessary for an effective conveyance of meaning in the TT. Over the past decade or so, computational linguists have been revisiting this issue on differences in the representation of meaning in the TL and the SL (Laviosa, 2002). In fact, translation deviations occur more frequently than it is thought, and direct correspondences are exceptional cases to the norm (Cyrus, 2006). Bonnie Dorr was influential in her work of translation divergences. She saw

translation divergences in terms of differences in the structure of sentences of the ST and TT, in the conveyance of the same information (Dorr, 1993). She points out seven kinds of syntactic divergences commonly found in pairs of translated sentences: conflation, structural, thematic, categorical, demotional, promotional and lexical (Dorr, 1992/1993).

On the other hand, Vinay and Darbelnet (1958/1995), as well as Catford (1965) founded a new term “translation shift” to describe the notion of semantic deviation, in contrast to the concept of “translation equivalence”. A shift is said to be derived from a “departure(s) from formal correspondence” (Catford 1965: 73) in the ST and the TT. Formal correspondents refer to sentences in their respective ST and TTs belonging to the same categorical places. Translation equivalents are two texts which hold the relation that one is an exact translation of the other. Where a shift occurs, the translation equivalents have failed to be formal correspondents.

In view of translation discrepancies present in STs and TTs, many contrastive linguists have acknowledged that translation divergences can occur at a lexical, grammatical, or at a structural level (Catford, 1965; Baker, 1992). The following displays the circumstances in which translation shifts take place:

1. When a translation equivalent does not have the same synthetic structure as the ST, there is synthetic deviation (Dorr, 1993)
2. When the SL and TL lexicalize a similar concept with a different word or phrase, there is lexical deviation (Marello, 1989; Vinay & Darbelnet, 1958/1995)
3. When the translation equivalent is not able to reproduce all the connotations expressed by a word from the SL, there are differences in connotation (Brown, Mendes & Natali, 1995)
4. When the denotation of a word from the SL partially overlaps with the denotation of the translation equivalent, there are denotation differences (Lo Cascio, Boraschi, & Corda, 1995)

2.3 Verbs in English

2.3.1 Idiomatic expressions in English

Compositional phrases are said to be grammatically constructed, while idiomatic ones have idiosyncratic semantic and syntactic qualities which characterizes them to be part of the lexicon (Jackendoff, 2002). An idiomatic example would be *make up your mind* where its meaning of “to decide” is not derived from its components, *make up* and *mind*. In contrast, *walk down the aisle* is compositional in meaning, as it is predictable from its parts. In this view, there is a need for these two to be distinguished, as idiomatic phrases are unpredictable and thus need to be stored in the lexicon, while compositional ones need not stored (Snider and Arnon, 2012).

From a translational viewpoint, idiomatic expression or proverbs may be reworded into a different phrase, as they cannot be translated word for word. This is where OBLIQUE TRANSLATION occurs (Molina and Hurtado Albir, 2002). On the other hand, a correspondence of the idiomatic phrase that is structurally similar may exist in the TL. We would like to discover the possibilities of this occurring in Japanese translations of English texts as well.

In this study, we will look at multi-word expressions (word syntactically constructed linearly in a sequence) that are idiomatic, to help us identify phrases that can be potentially translated as a whole.

2.3.2 English phrasal verbs

A phrasal verb is generally defined as the combination of a verb and a preposition or a particle of an adverbial nature, (Vlad, 1998; Olteanu, 2012), and is taken to be a single lexical item.

Three possible combinations or syntactic categorizations of phrasal verbs are identified by Courtney, 1983:

1. Verb + adverb
2. Verb + preposition
3. Verb + adverb + preposition

Although it has been debated that phrasal verbs are to be viewed as idiomatic (Biber, Stig, Leech, Conrad, & Finegan, 1999: 403; Fraser, 1976), others have also given recognition to literal phrasal verbs as well (Claridge, 2000: 47; Bolinger, 1971: 16, 112–114). 1(i) demonstrates the literal meaning and 1(ii) shows the idiomatic meaning phrasal verbs can have.

(1) (i) He took up the coin from the floor.

(ii) He cleaned up his desk.

Example 1(i) demonstrates a literal phrasal verb that has a verb that specifies movement in space and a particle that expresses spatial information (Kilby, 1984). The example demonstrates the regular use of *up* as it denotes the directional upward movement, of the action indicated by the verb *took*. In 1(ii), the use of *up* has aspectual meaning, where *up* sets a “natural limit on the action” (Kilby, 1984: 105). In our paper, we will examine how this function of a particle contributes to the regular meaning of a phrasal verb.

2.4 Verbs in Japanese

Japanese is classified as an agglutinative language (Shibatani, 1990), where grammatical inflections are performed on the base form of verbs.

Politeness is conjugated on verbs, and depending on the type of affix used, they represent different levels of politeness. Polite affixes include *-masu* forms 食べます *tabemasu* “to eat”, while honorific conjugations consist of the form “お+ stem verb+ なる” (as in お待ちになる *omachi ni naru* “to wait” from the base form 待つ *matsu* “to wait”), and humble forms follow the pattern “お+ stem +する” (as in お待ちにする *omachi ni suru* “to wait”). Not limited to these structures, Japanese has a complex honorific system marking the hierarchical difference between referents (Paik, Ohtake, Bond and Yamamoto, 2004), and gathering pragmatic information to construe the social relationship between referents in annotation is a challenging task (Nariyama, Nakaiwa and Siegel, 2005). Since English does not have a similar honorific system, there is a possibility for translation mismatches to occur in the processing and translation of Japanese honorifics.

A large proportion of Japanese verbs is said to be from Chinese loanwords (Shibatani, 1990), and these verbs follow the periphrastic form noun + する *suru* “to do”. Termed as サ変 *sahen* verbs, these borrowed Chinese characters behave as verbal nouns which can be employed as standalone nouns in the lexicon. This grammatical behavior of verbal nouns has been observed to contribute to syntactic shifts in the translation of English verbs in the corpora (Gao, 2012). As *suru* “to do” works as a derivational suffix that changes word-class membership of the predicated it attaches to (Backhouse, 1993), *sahen* verbs are able to function equally and freely as nouns or verbs. Hence, we will observe if *sahen* verbs contributes to shifts in translation as well.

2.4.1 Idiomatic expressions in Japanese

Japanese verb phrases also vary on a semantic continuum. However, for those that are idiosyncratic, it is necessary that they are acknowledged, as they express a single concept. These opaque expressions are usually expressed in the syntactic forms *noun* + を+*verb* or *noun* + が+*verb*, where を *wo*, the accusative marker and が *ga*, the nominative marker are vital in our delineation of phrases in Japanese. An example of the former is 匙を投げる *saji wo nageru* “to throw a spoon, i.e. to give up (something)” (Backhouse, 1993: 86) and 気がする *ki ga suru* “to have a certain feeling”. We will examine idiosyncratic phrases marked with *wo* and *ga* in our paper.

2.4.2 Japanese Compound verbs

A Japanese compound verb (VV) is formed by the “concatenation of two or more verbs which function as a single multiword verb” (Breen & Baldwin, 2009: 35). As several types of VVs exist, we will narrow down our focus to examine VVs comprising of native Japanese verbs, which have a “non-finite verb stem” in first position (V₁) and a “verb with tense” in the second position (V₂) (Hokari, Kumagami and Akimoto, 2012: 156), in the manner following the combination of 打つ *utsu* “to hit” and 込む *komu* “to plunge into” to create 駆け込む *uchikom* “to drive in (a nail); to devote oneself to something”.

VVs can be differentiated into two types: syntactic and lexical VVs (Uchiyama, Baldwin & Ishizaki, 2005). Martin (1975), Tsujimura (1996), Hokari et al. (2012) identified two main differences between the syntactic and lexical types: their semantic behaviours and productivity. Syntactic types are semantically transparent and have meanings which can be deduced from its constituents, while lexical types can be semantically opaque. V₂s in syntactic VVs are able to compound freely with other verbs, while V₂s in lexical types are fairly restricted in use (Hokari et al., 2012).

Their productivity can be explained with their constituent structures as depicted below. Figure 1 and 2 show that the type of V₂ determines the type of VV it will be, and in turn affects their syntactic formation. The same V₁ is able to combine with a different V₂ to produce either a lexical or syntactic compound.

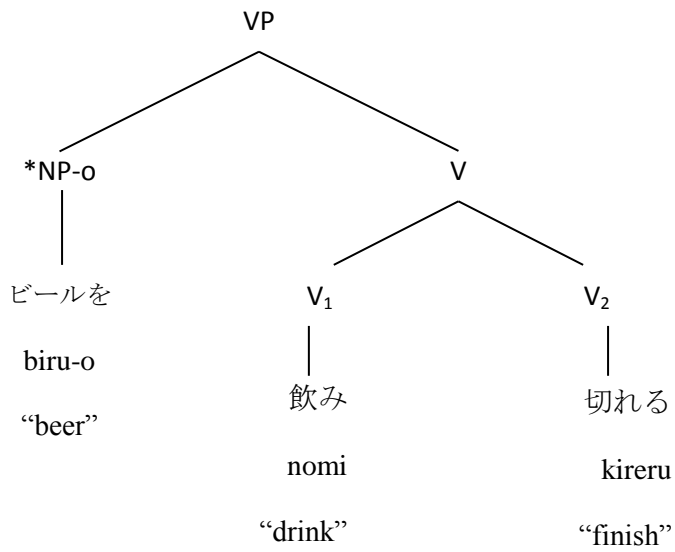


Figure 1: Lexical Compound Verb

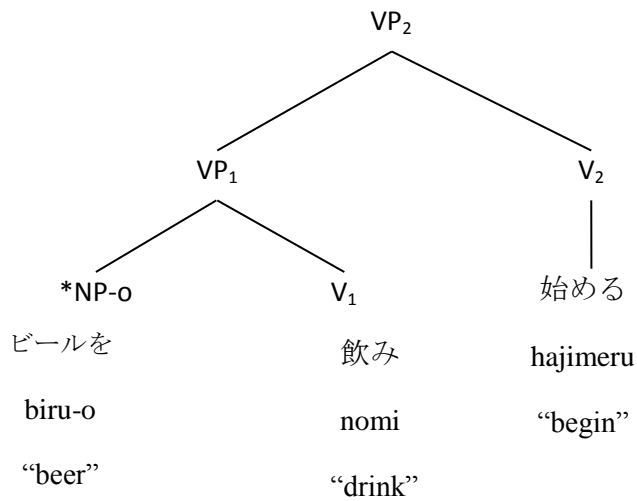


Figure 2: Syntactic Compound Verb

(adapted from Tamaoka, Lim, & Hiromu, 2004, pp.234, figure 1)

*NP-o refers to the noun phrase case-marked in the accusative

Hashimoto and Bond (2005) summarize the difference between syntactic and lexical in terms of productivity, compositionality, transitivity and semantics.

Properties	Syntactic	Lexical
Productivity	Highly productive	Less productive
Compositionality	Compositional	Varying degrees of compositionality, others are highly lexicalized
Transitivity	Transitivity of V1 is carried over to the transitivity of the entire VV	V1, V2, or both can contribute to the transitivity of the entire verb
Semantics	Embedding of V2's semantics on V1	Various kinds of semantic composition

Table 3: Comparison between Syntactic and Lexical VVs

2.5 Comparisons between Japanese compound verbs and English phrasal verbs

In Hokari, Kumagami, and Akimoto's study (2012), translation correspondences of Japanese VVs in English, Korean and Chinese were examined. Formal correspondences of VVs were said to be few in English in terms of their syntactic realizations, where combinations of two English verbs into a single compound such as *stir-fry* is uncommon (Hokari et al., 2012: 159). Although similarities in syntactic patterns and grammatical properties of English phrasal verbs (PV) and Japanese VVs were not analyzed, their study pointed out semantic commonalities between PVs and VVs, specifically verb-particle combinations.

Other studies have analyzed semantic features of Japanese verbs, and traced translation patterns that define semantic relationships of Japanese and English verbs in a variety of contexts. Uchiyama, Baldwin and Ishizaki, S. (2005) considered the similarities in semantic properties that PVs and VVs share based on the usage of the particle in the PV, and V₂ in the VV in a given context. Acknowledging that both English particles and Japanese V₂ verbs can potentially express the semantic component of path in a motion event, Uchiyama et al. (2005) captured 3 broad semantic categories to classify the types of V₂s: spatial, aspectual, and adverbial. This classification system effectively shows the polysemous behaviour that certain V₂s may have, and thus captures the variations of their verbal patterns manifested in various

syntactic combinations. Borrowing an illustration from Uchiyama et al. (2005), the V₂ 上がる *agaru* “go up” is highly polysemous, similar to “up” in English as well, as it serves all three functions, in ゆであがる *yudeagaru* “to finish boiling” (aspectual), 飛び上がる *tobiagaru* “to jump up” (spatial), and 振るえ上がる *furueagaru* “be terrified”. Such predictions of V₂ usage in context helps to disambiguate VVs in semantic processing.

Targeting Japanese motion verbs, Uchiyama and Ishizaki (2001) studied the combination of Japanese motion verbs that produce Japanese compound motion verbs. They looked into the semantic components and five “lexicalization patterns” of the constituents in the VV, which includes motion, path, figure, ground and manner/cause (Uchiyama and Ishizaki, 2001: 2). V1 and V2 combinations can be organized according to permutations of these five patterns as shown in the study¹. By delineating different types of verbal constituents, their productive patterns can be predicted. Identifying semantic relations between the constituents of the VV were found to aid translations of VVs into English.

Thus, it appears that studies on translations of English phrasal verbs and Japanese compound verbs have been limited to Japanese-English translations. Studies above were also based on word lexicons and word lists, which, in contrast to a translated text, have a smaller sample size and variety of VVs. As few studies have conducted an analysis based on English-Japanese translations, this study aims to examine translations of English phrasal verbs into Japanese compound verbs, and how semantic meaning of English PVs is rendered in Japanese equivalents.

CHAPTER THREE

3.1 Aims and hypotheses

We have seen the similarities that both English phrasal verbs and Japanese compound verbs share, and we would like to explore the correlation between them which can become a potential strategy in achieving translational equivalence. Not restricted to these types, we would like to see how translation shifts apply across various types of verbs in English and Japanese, by following patterns in English to Japanese translations. Our research questions are as follows:

1. What are the corresponding patterns of translation, as well as translation divergences, between English verbs and their Japanese counterparts?

¹ For more information on lexicalization patterns and combinations of verbs of varying lexicalization patterns, see Uchiyama, K., & Ishizaki, S. (2001). A study of translation rules for the Japanese compound motion verbs. *Proceedings of the AFNLP PACLING 2001*.

2. How can these patterns observed in the relation between English phrasal verbs and Japanese translations be accounted for from cross-linguistic and translational perspectives?

Presently, no English-Japanese parallel corpus based study has been found to conduct a cross-comparison study on verbs. Hence, this study will focus on four main types of verbs: English phrasal verbs, Japanese compound verbs, idiomatic expressions and single verbs in both Japanese and English.

Considering syntactic correspondence at the lexical level, the following hypotheses are proposed:

- i. English phrasal verbs would be most frequently matched to Japanese compound verbs, followed by Japanese idiomatic expressions and single verbs respectively.
- ii. English idiomatic expressions would be most frequently matched to Japanese idiomatic expressions, followed by Japanese compound verbs and single verbs respectively.
- iii. English single verbs would be most frequently matched to Japanese single verbs, followed by Japanese compound verbs and idiomatic expressions respectively.

CHAPTER FOUR

METHODOLOGY

4.1 Resources

WordNet

Founded and created by Fellbaum in 1998, the Princeton WordNet (PWN) of English is a large lexical database that classifies nouns, verbs, adjectives, and adverbs into sets of cognitive synonyms known as synsets. Each synset represents a distinct concept, and is linked to another by semantic and lexical associations (Fellbaum, 2005). Thus, their conceptual relationships can be built on synonymy, hyponymy, antonymy and so forth. Verbs in our study will only be considered for synonymy and antonymy, as we focus on the semantic closeness of the English verb and its translated counterpart.

Moreover, WordNet captures polysemous features of some verbs, and even specifies the type of sentence structure in which a particular verb sense can appear in (Clerk, Coleman, Willems, 2013). It enables us to see how units of meaning are manifested across various contexts in the languages compared. WordNet is also a rich resource to retrieve similar lexical features among synsets, which would be helpful in our comparison between verb types in English and Japanese, especially to trace a trend of a particular verb type falling under a specific semantic category, such as verbs of motion.

Published in 2008, the Japanese WordNet also follows a structure similar to that of PWN, where synsets are organized in the same hierarchical manner, based on lexical relations among words. Current work is done to ensure that JWN matches with the synsets and lexical structures of the PWN (Isahara, Bond, Uchimoto, Utiyama and Kanzaki, 2008), for better linkage of concepts across languages.

Corpus

For this study, we looked at a fiction short story, named “The Adventure of the Speckled Band” that is part of the Sherlock Holmes series, authored by Sir Arthur Conan Doyle (1892). The texts were paralleled and aligned in the NTU Multilingual Corpus. The texts had to be aligned sentence by sentence before they were automatically tokenized and given part-of-speech (POS) and lemma tags. Thereafter, verbs were tagged manually with senses obtained from the respective WordNets. As the two texts had been tagged by other annotators previously, the links between verbs from the source language and target language were checked thoroughly, and the mistakes in sense-tagging are further elaborated below.

4.2 Tagging Multiword Expressions in WordNet

New work was done in this study on MWEs. Multiword expressions (MWE) are defined to be “idiosyncratic interpretations that cross word boundaries (or spaces)” (Sag et al., 2002: 2). As one concept can be expressed by a group of words, it is necessary to form a single expression to represent a particular distinct concept. In this study, MWEs are examined in terms of compounds or idiomatic phrases. Verb compounds here refer to English phrasal verbs (PV) and Japanese compound verbs (VV). An example of a PV includes *jut out*, which was initially recognized in the system as two individual words. However, as the concept of *jut out* is lexicalized, and its meaning differs from its single components *jut* and *out*, a single MWE was created and tagged to one synset, in this case 02713372-v, denoting “extend out or project in space”, and finally concluding with the comment “PV”.

The same tagging procedure is performed for VVs. *突き出す tsukidasu* “to stick out” is in fact a derived form from two verbs *突く tsuku* “to stab” and *出す dasu* “to come out”, thus a MWE has to be formed for these two verbs. However, as found in many cases, the Japanese WordNet has yet to create distinct synsets for VVs, thus they are either tagged with the closest English equivalents in the English WordNet, or given meanings in the comment section. For *突き出す*, no suitable synset could be found, thus it was given the comment “to stick out”, followed by “VV”.

Phrases in English and Japanese which are commented with “phrE” and “phrJ” respectively, had lexemes grouped in the same manner. We ensured that their POS was tagged for “verb” by ensuring that a verb heads the idiomatic phrase, such as in *make up my mind*. Japanese phrases such as *音を立てる oto wo tateru* “to make a noise” or *気が済む kigasumu* “to feel good” have the ACC marker *を wo* and the NOM marker *が ga* to indicate the verb followed up behind, which determines the entire phrase to be a verb. Meanings were also manually keyed in, in the comment section if no suitable synset was found to fit the sense of the phrase.

4.3 Annotating relationships shared between words

For all tagged data, the program runs on an automatic linking system which links the SL and TL words in each sentence pair in the following circumstances:

1. Both words share the same synset e.g. think and *思う omou* “to think”, both tagged with the same synset, 00631737-v, defined as “expect, believe or suppose”

2. The two words share a hyponymy relationship e.g. *chat*, 01038666-v, which denotes “talk socially without exchanging too much information”, is a hyponym of *話す* *hanasu*, 00964694-v, defined as “to carry a conversation”

Following the tagging guidelines used in Mok (2012) (cited in Gao, 2012), we used the following symbols to denote the relationship between the source words and target words in the corpora.

- i. “=” for matched pairs with the same synsets

Each concept pair that shares the same synset is matched with the link “=” . In the ST, *wait* has the synset 02637938-v, defined as “stay in one place and anticipate or expect something”. Similarly, the equivalent verb *待つ* *matsu* “to wait” in the TT has the same synset 02637938-v. Thus, the pair was tagged with the link “=”.

This link type is also used in linking English verbs with the equivalent *サ変* *sahen* Japanese verbs (a noun treated as a verb in usage), which takes the form “NOUN + *する* *suru*”. For example, *marry* is tagged with the synset 02488834-v with the meaning “take in marriage”. In Japanese, the concept is translated into the same synset 02488834-v. A multiword expression (MWE) lemma is first created for the *sahen* verb *結婚する* *kekkon suru*, “to marry”, to join the noun *結婚* and the verb *する* together, before it is tagged as a whole to the synset. Thereafter, we linked both concepts with the symbol “=”.

- ii. “~” for those which share similar meanings or a non-hyponymy relationship and those with POS gaps

For pairs with concepts which are non-identical but similar, they are linked with “~”. Synonymous pairs are thus indicated by “~”. This symbol, however, does not define direct hyponymy in verb pairs, and similar concepts are linked regardless of the presence of a hyponymy relation. An example includes *prevent* tagged with synset 02452885-v, of the definition “keep from happening or arising”, but *妨げる* *samatageru* “to prevent” had the sense 02450505-v meaning “stop (someone or something) from doing something or being in a certain state”. These two words are not related by hyponymy in WordNet; they express similar concepts, as both refer to the act of prohibiting the occurrence of something. Therefore these two words were linked with “~”.

Additionally, word pairs with POS gaps were given the symbol “~”. This included *satisfy*, with the synset 01816431-v meaning “make happy or satisfied”, and its Japanese counterpart 満足 *manzoku* “satisfaction”, a noun tagged with 07531255-n, defined as “the contentment one feels when one has fulfilled a desire, need, or expectation”. Pairs like these showing a lexical shift were linked with “~”.

iii. “:” for similar not lexicalized meanings

We used this link “:” in tagging where the concept shared between words is not fully lexicalized in either language. This is especially so for ST and TT words which are not semantically close, yet one word is a translated correspondent of the other.

(2)...a face which spoke her joy

喜び	に	顔	を	<u>輝かせて</u>
yorokobi	ni	kao	wo	kagayakasete
joy	DAT	face	ACC	shining

“face beaming with joy”

FSID 10311 TSID 10375 NTU-MC

As above, we can observe the connection between the individual concepts. Although it is apparent that 輝く *kagayaku* “to shine” is the Japanese translational correspondent of *speak*, it cannot be said that they are semantically related or similar, because generally they do not express the same concept lexically. However, to show that they are related contextually to a certain degree, they are linked with “:”.

iv. “!” for direct antonyms

For words with opposite meanings, their relationship is represented with the “!” symbol. A pair that demonstrates antonymy would be the lemma come and the Japanese translation 行く *iku* “to go”.

v. “#” for weak antonyms

More commonly found than direct antonyms, weak antonyms in the corpora were represented with the symbol “#”. The example below shows an English phrasal verb that was interestingly translated into an equivalent Japanese compound verb.

(3)...taking in every detail of the apartment

この 部屋 の どんな 細部 も 見逃さなかつた

kono heya no donna saibu mo minogasa-nakatta

this room ADN what detail FOC did not overlook

“did not miss any detail of the room”

FSID 10189 TSID 10236 NTU-MC

In the TT, the Japanese compound verb *見逃す minogasu* “to overlook” was tagged with the synset 02166361-v, defined as “look past; fail to notice”, which had a sense somewhat opposite from *take in*, tagged with 01539063-v defined as “take in; also metaphorically”. Synset 01539063-v did not explain *take in* well, but it holds the meaning of absorbing visual information and processing it. *見逃す* however does not encompass the sense of cognitive processing at a deeper level, unlike *take in*, so we deemed the pair as weak antonyms and linked as “#”.

A 95% level of confidence was given to every lemma that was tagged manually. Table 4 briefly summarizes the types of link symbols used to represent the relationship between word pairs. In our study, translational shifts are represented by the symbols “~”, “:”, “!”, and “#”.

Symbol	Relationship
=	Same synset
~	Similar or related meaning, including different POS
:	Translation equivalent but not clearly linked lexically
!	Direct antonym
#	Weak antonym

Table 4: Symbols used to indicate relationships between words

4.4 Tagging Issues

4.4.1 Polite forms

Some Japanese honorific verbs were identified in the TT and those that were purely polite forms in themselves (i.e. not inflected for polite forms) were tagged manually and given the comment, “polite” in the comment section.

(4) You know me, then?

では、私 の こと を ご存知なの で？

dewa, watashi no koto wo gozonji nano de

so, 1SG POSS matter ACC HON-know FIL COP

“so, you know me then?”

FSID 10073 TSID 10047 NTU-MC

As opposed to 知る *shiru* “to know” which is the non-polite form in Japanese, ご存知 *gozonji* “to know” is employed by the speaker in his/her inquiry to an addressee of a higher status, on whether the addressee is informed on something. As the core semantic meaning still remains, but additional information on social status is indexed in the word itself, the words “know” and ご存知 *gozonji* “to know” were linked with “~”.

4.4.2 Poor Translations

Special cases of loose translations were detected in the TT, and these were omitted from the overall data set. This could be a misinterpretation on the translator’s part, and this is also a manifestation of translation problems based on the ability of the translator. One of the cases that was observed is shown in (5) below where there is a literal translation of a metaphorical English phrase into Japanese. It exemplifies the wrong use of EXPLICITATION, a translation technique used to convey original information in an explicit manner:

(5)... *he shut himself up in his house and seldom came out save to indulge in ferocious quarrels with whoever might cross his path.*

(a)	義父	は	家	の	中	に	閉じこもって、
	gifu	wa	ie	no	naka	ni	tojikomotte,
	stepfather	TOP	house	ADN		within	LOC shut in
	ほとんど						出てきませんでした。’
	hardly						come out-NEG-COP- PST
	hotondo						detekimasendeshita.

“stepfather keeps himself in the house, and rarely comes out”

(b) []	私道を	渡ろう	とする人と	激しい	喧嘩を
	shidou wo	watarou	to suru hito to	hageshii	kenka wo
	personal path ACC	cross	to do person	strongly	argues ACC

するのです。
suru no desu.

to do FIL COP

“he will argue with anyone who crosses his personal path”

FSID 10081 TSID 10107 NTU-MC

As above, “cross his path” would be an idiomatic phrase that would mean to meet or encounter someone unexpectedly in this situation. However, the phrase was literally interpreted to mean the physical action of crossing into the lane of the other person, as in 私道を渡ろう *shidou wo watarou* lit. “cross over (his)private road”. Thus, since the original meaning of the phrase was not reflected in the translated Japanese sentence, we decided to leave the two phrases unlinked and excluded them from our analysis.

4.4.3 Unexpected discoveries in corpus – concepts that were not linked

Another common occurrence includes the appearance of a corresponding verb in the TT where the word is not present in the original text. This means that an additional Japanese verb is added into the translation although the verb equivalent does not exist in the ST. Often times, they are a repeated verb (verb in ST is translated twice) or an extra verb inserted to convey additional information. The following example illustrates the first scenario.

(6)...went out to Calcutta, where, by his professional skill and his force of character, he established a large practice.

カルカッタへ 行って 開業し、 技量と 押しの強さで

karukatta e itte kaigyō shi, giryou to oshi no tsuyosa de

Calcutta DAT go establish skill and force of character COP

大いに繁盛しました。

ooi ni hanjouu shimashita

greatly flourished

“He went to Calcutta to set up his practice, and with his ability and force of character, his business flourished”

FSID 10071 TSID 10093 NTU-MC

As above, the final expression 大いに繁盛しました *ooi ni hanjoushimashita* “greatly flourished” was additionally inserted to support the argument in the ST of establishing “a large practice”. However, “establish... a practice” was translated into one verb 開業する *kaigyō suru* “open business”, and the adjective “large” by 大い later on in the sentence. The sense of a flourishing business was implied from “a large practice” and was made explicit in the additional Japanese verb 繁盛する at the end. As we did not want to overlap words in annotation, we did not tag such extra translated verbs with the verb/phrase where the implied meaning was interpreted from.

Another type of Japanese verbs that were not tagged includes those that are employed as adjectives that modify clauses. They behave as participles in the TT, and they always precede a modified noun in the clause.

(7)...like those of some hunted animal.

狩られる 動物の ようだった。

karareru doubutsu no you data.

being hunted animal ADN seems COP

“like a hunted animal”

FSID 10031 TSID 10040 NTU-MC

Although its form remains a verb, 狩られる *karareru* “to be hunted” modifies the following noun 動物 *doubutsu* “animal”, thus making it an adjective. Thus, participles in the TT will be ignored and left unlinked.

4.4.4 Tagging polysemous words

Certain ST words with vague meanings were found to be translated with another ambiguous word. We encountered a few with such ambiguity, such as the example below.

(8) But I am now sleeping in the middle one (room).

でも、私は、今は 真中 の 部屋で 寝起きしています。

demo, watashiwa, ima wa mannaka no heya de neoki shiteimasu.

but I TOP now TOP middle AND room COP sleep and wake up

“but, as for me, I sleep and wake up in the middle room”

FSID 10332 TSID 10400 NTU-MC

“sleep” and 寝起き *neoki* “awake” are fundamentally two antonyms in concepts. However, both are contextualized to mean “to stay/use the room to sleep and wake up in” in both languages. Therefore, we decided upon the implied meaning, and tagged both concepts with the same synset 02637202-v 'dwell' as they do convey the same implications in the respective SL and TL. We then linked the two with “~” as their basic meanings are different.

CHAPTER FIVE

RESULTS AND ANALYSIS

With the words tagged to their concepts and linked, we proceeded to count the number of verbs that were linked in the corpus. Out of 1665 verbs in total in the ST, 869 were tagged with synsets. This means that 52.2% of the English verbs were linked, and this number also represented the linking of verbs to words of a different POS in the TT. The unlinked verbs could be accounted for by the light verbs and auxiliary verbs in English which hold little semantic content and were excluded from the analysis. On the other hand, out of a total of 1211 Japanese verbs in the TT, 248 verbs were not linked. This could be attributed to the translation of English words to Japanese verbs of a different POS, where these English words that were not verbs were not linked since a unidirectional analysis of parallel corpora is performed here, and our focus is on shifts in POS in the translation of English verbs to their Japanese counterparts.

After tabulating the number of linked concepts in the TT and their POS tags, links that represented a shift in word class in the translation process were presented in Table 5 below. The figures demonstrate all linked English verbs except for those linked with “=”. Those that had the link “=” represented pairs of exact correspondences, and were thus excluded.

POS Gap	Total Number of Matches	%
Verb-Verb	367	93.10
Verb-Noun	19	4.82
Verb-Adjective	6	1.52
Verb-Adverb	2	0.51
Total	394	100.00

Table 5: Cross lingual matches of POS relations across words

From the table above, the highest percentage of matches came from linked concepts sharing the same POS, which is to say that 93.1% of translation shifts in our corpora had occurred between SL and TL words of the same word class. It is apparent that English verbs are mostly translated into Japanese verbs, although they are not exact correspondences of each other. The next highest percentage of syntactic mismatches occurred between English verbs and Japanese nouns (4.82%). However, mismatches between verb-adjective and verb-adverb do not occur frequently.

As for verb-verb matches in the corpora, Table 6 depicts the association by link types between the matched verb pairs of various verb types, namely, PV,VV, phrE,phrJ, singleE,singleJ. Of the total number of verb-verb matches, more than half of the pairs (55.87%) were tagged with “=”, which represents that they are exact correspondents of each other. The next significant shared relationship between the pairs would be represented by “~” (25.50%), where there exists a semantic dissimilarity between them. Following up closely is the link type “:” (18.15%), where the pairs do not share a similar lexicalized concept. Direct and weak antonyms were however few in the corpora. In total, there were 843 pairs of linked verbs in the corpora.

Matched verb-pairs/Link type	=	~	:	#	!	Total
PV-VV	18	9	3	1	-	31
PV-phrJ	1	4	2	-	-	7
PV-singleJ	13	23	13	-	-	49
phrE-VV	7	-	1	-	-	8
phrE-phrJ	4	2	1	-	-	7
phrE-singleJ	1	1	5	-	-	7
singleE-VV	28	26	14	-	-	68
singleE-phrJ	10	9	-	-	-	19
singleE-singleJ	388	141	114	1	3	647
Total	471	215	153	2	3	843
%	55.87	25.50	18.15	0.24	0.36	100

Table 6: Cross-matches of various verb types in the corpora and their link types

As predicted for PVs, the PV-VV pair were linked with “=” the most (18), which manifests that they are formal correspondences of each other, and there was a lesser number of PV-singleJ pairs similarly linked with “=” (13), showing that they are less similar conceptually. On the contrary, there were very few PV-phrJ pairs uncovered in total.

PV-singleJ were interestingly found to be mostly tagged with “~” (23), which was against our predictions that the most number of “~” linked pairs would be from PV-phrJ pairs (4). Hence, with respect to cases of non-correspondences in translation (as represented by “~” and “:”), PV-singleJ dominated in nominal

figures. This disproved the prediction that PV-phrJ pair to be the second most commonly tagged pair with these links types.

Similarly for single verbs, the number of singleE-singleJ pairs were found to be linked with “=” the most (388), followed by singleE-VV (28) pairs and lastly singleE-phrJ pairs (10), which means there is a relatively high number of formal correspondences for singleE-singleJ pairs. However, even for conceptually dissimilar verbs linked with either “~” or “:”, the singleE-singleJ matches (141 and 114) were the highest compared to singleE-VV (26 and 14) matches and singleE-phrJ (9 and 0) matches.

An interesting trend was spotted for singleE-VV (28 and 26) pairs and PV-singleJ (13 and 23) pairs, that they have a relatively high number of “=” and “~” link types for each of these pairs. These figures exceeded expectations, and they were tested for random occurrence in the Chi square test below.

Unfortunately, we were unable to find a trend for matches with English phrases, as there were too few data collected from the corpora. We thus rejected the second hypothesis (refer to section 3 part ii) on translations from English idiomatic expressions to Japanese idiomatic expressions, and proceeded to focus our discussion on the other hypotheses.

With the remaining two hypotheses, we thus tested whether the remaining SL verb types, PV and singleE, had an effect on the distribution of the corresponding TL verb types, VV and singleJ, by applying the Chi square test. The Chi square test is a statistical hypothesis test that depicts the situation where the null hypothesis is true, and the sampling distribution of a test statistic follows a Chi-square distribution (Jenset, 2008). By showing that the distribution of our observed data does not follow a Chi square distribution, we can then determine that variations in our data are not due to random variation, and thus reject null hypothesis in favour of the alternative hypothesis.

In order to determine that the patterns of verb translations are not due to chance, we used a Chi-square test on the data we collected. Earlier in the literature review, we predicted that particular verb types in the SL would be related to the type of verb constructions found in the TL (i.e. a specific verb type contained in the ST would influence the type of verb to be constructed in the TT). Verbs in English are divided into three groups: phrasal verbs, idiomatic phrases and single verbs, while verbs in Japanese are categorized according to the following three types: compound verbs, idiomatic phrases and single verbs. In this study, the dependent variable is the *verb correspondences found in the TT*, and the independent variable is the *verb type employed in the ST*. Therefore, the Chi-square test is suitable as a correlation test since the data is nominal and variables measured are organized into non-overlapping categories.

English/Japanese	VV	SingleJ	Total
PV	31 (9.96)	49 (70.04)	80
SingleE	68 (89.04)	647 (625.96)	715
Total	99	696	795

Table 7: Summary of expected frequencies of each verb type

The Chi-square test shows that there is an influence of the verb types in English on the translations of the corresponding verb types in Japanese ($\chi^2 = 56.423, p < 0.001$)². Thus, the null hypothesis is rejected.

As shown above, the observed frequencies for PV-singleJ matches (49) and singleE-VV matches (68) did not reach the figures of expected frequencies calculated (70.04 and 89.04 respectively). This means that despite the relatively high nominal numbers tabulated for these matches, they point to chance occurrence, i.e. matches were contingent on random relations.

In contrast, PV-VV matches (31) and singleE-singleJ matches (647) have both exceeded the calculated expected figures (9.96 and 625.96 respectively), thus the data seem to indicate that they are parallel realizations of each other in the ST and TT. Specifically for English PVs and Japanese VVs, further investigations would be conducted in this paper to examine their common underlying semantic structures that has established the pair to correspondences of each other.

² *p-value is the value of significance at which the null hypothesis should be rejected, and in this case, $p < 0.05$ is the standardized p-value in the humanities and social sciences (Jenset, 2008, Litosseliti, 2010). Here, $p < 0.001$ shows that the results are highly significant.

CHAPTER SIX

DISCUSSION

6.1 Analysis of zero matches in the corpora

It was observed that a fair number of ST verbs (796) did not have a lexical equivalent in the TT, i.e. the concept was not represented at all in the TT. One possibility to account for this is the difference in realization of grammatical functions, where a phrase can be condensed into a less dense construction using verb conjugations in Japanese. The following shows synthetic correspondence (Malá, 2013), where the meaning of the ST verb is represented with a grammatical inflection in Japanese.

(9) I can only say, madam, that I shall be happy to devote the same care to your case as I did to that of your friend.

マダム、お友達に してあげた くらい のことは、喜んでさせていただきますよ。

madamu, otomodachi ni shite ageta kurai no koto wa, yorokonde saseteitadakisimasu yo.

madam, friend DAT to give like ADN thing TOP, happy to have the privilege to do

“Madam, as I have offered the same to your friend, I am happy to have the privilege (to devote the same care to you).

FSID 10057 TSID 10073 NTU-MC

The meaning of *devoted* is incorporated into the Japanese verb 喜ぶ *yorokobu* “to be pleased” which is inflected with the polite causative suffix させていただく *sasete itadaku* “to have the privilege of doing”. This inflection is commonly used in Japanese to refer to an act of offering a service to another person. Although *sasete itadaku* is actually a derived form of する *suru* “to do” (a light verb) but functions as an honorific suffix, it is not a lexical equivalent and was left unlinked in the corpora. Politeness in Japanese is one problem significant in the area of translation divergences.

6.2 Analysis of shifts in POS

Translation shifts observed in this study highlighted typological differences in the distribution of linguistic features in the representation of concepts. Findings on translation shifts support those of Gao (2012), who also found the highest number of POS shifts between verbs and nouns, followed by verb-adjective and verb-adverb pairs (Gao, 2012: 27, table 9). Language influences on categorical shifts (as mentioned in Dorr (1992/1993)), or TRANSPOSITION (as termed by Vinay and Darbelnet, 1958) in English-Japanese translations can account for such shifts.

The POS category that English verbs were found to be commonly translated to was the noun category (19, 4.82%). The example below shows how the same concept can be realized differently in the two languages.

(10) Why, what do you mean?

まあ、それは いったい どういう 意味ですか？

maa, sore wa ittai douiu imi desu ka

well, that TOP on earth what meaning QUOT

‘Goodness, what is the meaning of that?’

FSID 10195 TSID 10243 NTU-MC

The above example shows a difference in lexical construction in expression of the concept “mean”. Constructions between languages are said to differ when they do not share either the same formal or functional features in the expression of a particular concept (Serbina, 2013). The concept of *mean* in the SL here is expressed as a verb, given the synset 00955148-v, defined as “mean or intend to express or convey”, but was translated into a different form, a noun, 意味 *imi* “meaning”, linked with synset 06601327-n denoting “the message that is intended or expressed or signified”. The difference in formal construction here demonstrates a lack of a formal correspondent in the TL.

Another case of nominalization of the SL verb is exemplified below, and this phenomenon could be attributed to the lack of an equivalent verb of the same transitivity in the TL.

(11) Dr. Grimesby Roylott's chamber was larger than that of his step-daughter, but was as plainly furnished.

グリムズビー・ロイロット博士 の部屋は、 義理の娘の部屋より

gurimuzubi roirotto professor ADN room TOP, stepsister's daughter's room

大きかったが、家具は 同様に 質素なものだった。

bigger than furniture TOP same DAT simple thing FIL COP

“Grimesby roylott's room was bigger than that of his step-daughter, but (his) furniture was similarly simple.”

FSID 10386 TSID 10457 NTU-MC

Here, the translator employed a noun, 家具 *kagu* “furniture” in Japanese, to express the same concept of “equipped with furniture” as conveyed in the original verb *furnished*. However, in the English context, the verb *furnished* refers to the state of the room being equipped with furniture. In contrast, in the Japanese translation, the focus of description falls on “furniture” itself, and not the room. In the TT, the use of the topic marker は *wa*, which always follows after a noun, has placed this focus on “furniture”.

Moreover, the positioning of the topic marker twice -- first after the noun 部屋 *heya* “room” and then after “furniture” -- sets up a semantic contrast between the state of the room and the state of the furniture. This grammatical construction for contrast between two topics is unique to Japanese, and は *wa* creates an obligation for the topic in the clause-initial position to be a noun, not a verb, as verbs in Japanese cannot be interpreted as topics (Vermeulen, 2009). Thus, the syntactic construction sets up a need for the use of a noun. Moreover, although a verb equivalent of *furnished*, which is 備えつける *sonaetsukeru* “to install/furnish”, is available in Japanese, it is not used as it is transitive and is not congruous with the intransitive *furnished*. Therefore, in adapting the Japanese translation to TL grammatical norms, it led to the change in lexical realization in the TT.

Besides verb-noun matches, there were a few ST verbs found to be translated into adjectives (6, 1.52%) in the TT. Interesting cases included an English verb that was translated into an adjective part of a fixed expression in the TL.

(12) ...you are at liberty to defray whatever expenses I may be put to, at the time which suits you best.

とはいえ、 <u>ご都合のよろしい</u>	ときに 実費を	ご負担して頂いても	結構
towaie,	gotsugou no yoroshii	tokini jippi wo	gofutan shite itadaitemo kekko
nonetheless, POL convenient	time DAT expenses ACC	POL bear	fine

ですよ。

desuyo.

COP

“Nonetheless, you can defray the costs at any time convenient”

Perhaps in favour of conventionalized TL features in translation, the verb *suit* was observed to be adapted for a frequently used expression in Japanese, 都合のよろしい *tsugou no yoroshii* “convenient”. Although the two concepts are not related lexically, taking the entire proposition into account “time which suits you” would make for a better interpretation that supports the translation into the Japanese adjectival phrase of “convenient”. Both the English verb and the Japanese adjective よろしい *yoroshii* in this case would then share the same concept of “be satisfactory to (somebody)”. Therefore, the POS shift here is undertaken by the translator to cater to the TL audience, and NORMALIZATION, which refers to the approach of conforming to the language patterns of the TL (Baker, 1996: 176 - 177), has facilitated a more natural translation in the TT.

Due to limited data gathered for verb-adverb shifts (2, 0.51%), further studies is recommended to explain the relationship between the English verbs and Japanese adjectives and adverbs respectively.

6.3 Analysis of single English verbs and Japanese compound verbs

Many single English verbs were found to be matched with VVs in the corpora (68 in total). For verbs that are semantically ambiguous, translations into VVs are found to explicate the implied meaning of the ST verb. EXPLICITATION is a technique used by the translator where implicit information in the ST is made explicit in the TT (Baker, 1996: 180).

(13)... you can see deeply into the manifold wickedness of the human heart.

あなたは	人間の心の、	様々な邪悪さを	<u>見通せる</u> 方と
anata wa	ningen no kokoro no,	samazama na jaakusa wo	mitooseru kata to
you TOP	human ADN heart ADN	many wickedness ACC	see through person QUO

伺っています。

ukagatteimasu

heard

“(I heard that) You can see the manifold wickedness of the human heart.”

FSID 10062 TSID 10082 NTU-MC

The meaning of see here is made literal with 見通す *mitoosu* “to see through”, where the implied meaning in see is “predict through perception” and that is done with the addition of 通す *toosu* “to pass through” to explicate the overall meaning of examining below the surface. The verb *toosu* here thus helps to amplify the sense of 見る *miru* “to see” as well.

Yet, EXPLICITATION is known to produce redundant information in the TT. EXPLICITATION can be done by the translator unconsciously (Olohan, 2004: 94), but here we would like to suggest that it can be due in part to language effects, or lexical expressions in the specific language. It is typical in Japanese to conflate the semantic factors of motion and path in motion verbs (Uchiyama and Ishizaki, 2001), and this can thus account for the frequent translation of single motion verbs in English into VVs in the corpus here. An alternative example include *jump (up from the floor)* and 飛び上がる *tobiagaru*, formed from 飛ぶ *tobu* “to fly” and 上がる *agaru* “up”. The V₂ here reinforces the degree to which the action denoted by V₁ is carried out.

6.4 Analysis of English phrasal verbs and single Japanese verbs

SIMPLIFICATION has occurred in cases where the phrasal verbs are conflated into a single verb in Japanese. SIMPLIFICATION can mean the reduction in the length of constructions and/or condensing complex information (Olohan, 2004).

This is illustrated with the PV come out, which has the motion verb “come” and the particle projecting the direction of “come” which is outwards. Japanese, however, is said to be a verb-framing language which incorporates the semantic component of path into verbs (Talmy, 1991, 2000). Thus, the corresponding verb of come out in Japanese is 出る *deru*, which expresses the same meaning of “come out”.

Translations of other PVs simply show the redundancy of the particle which contributes little semantic meaning to the whole, and thus is not reflected in the corresponding translated word. Such pairs were found to be linked with “=” and were thus semantically equivalent, since the meaning of the PV was fully conveyed in the single Japanese verb.

(14) The windows of the three rooms open out upon the lawn.

三つの 部屋の 窓は、 芝生へと 開きます。

mitsu no heya no mado wa, shibafu he to akimasu.

Three ADN rooms ADN window TOP lawn DAT open

“the windows of the three rooms open out to the lawn”

FSID 10105 TSID 10140 NTU-MC

The directional particle へ (pronounced /he/) which specifies the motion towards a specified target is parallel to the English particle *out* here in the PV open out. Therefore, in Japanese, there is no need for additional marking of direction in the verb itself.

However, in parallel with English phrasal verbs with metaphorical meanings which can be substituted for single verbs, these PVs can also be replaced with a single Japanese verb in translation. This is another illustration of explication in the TT as well. In the ST, *go on* was used to mean “to continue”, and 続く *tsudzuku* “to continue” was the corresponding word used in the translation to explicate this literal sense of *go on*.

6.5 Analysis of English phrasal verbs and Japanese compound verbs

In this section, combinations of matched PVs and VVs in the corpora, as well as their semantic commonalities will be discussed.

Following past classifications of ambiguous V_2 s into three semantic classes, namely: aspectual, spatial and adverbial (Uchiyama, Baldwin, Ishizaki, 2005: 6; Uchiyama and Ishizaki, 2001, para. 2.2), our study will examine, where possible, the manifestation of such semantic properties in V_2 s. As most PVs found in the data belonged mostly to the semantic class of verbs of motion, we will also deduce generalizations of patterns based on this group of PVs and their corresponding VVs.

Congruence between the main verb of the PV and V_1 , as well as the particle and V_2 was found in cases where the PV and its corresponding VV share common semantic components. In particular, a PV containing a motion verb with a particle expressing a spatial factor is found to be paralleled with a congruent motion verb in V_1 position, and another motion verb in V_2 position depicting the same spatial feature expressed by the particle. This is further supported from the findings that these verbs were

commonly marked with the “=” symbol which denotes that they are correspondences of each other, and are thus semantically similar.

A case in point would be the phrasal verb *take out* that was directly translated into 取り出す *toridasu* “to take out” in the TT. Both the PV and the VV in this case demonstrate semantic transparency, i.e. the meaning is apparent and is predictable from its parts. Their structures can thereby be broken down as such: The motion indicated by the verb *take* is modified by the particle *out* which conveys the change in location from an interior place to an exterior place (Neagu, 2007). Similarly in the translation equivalent, 出す *dasu* “to put out” modifies the manner in which the action of 取る *toru* “to take” is done. Here, in both languages, the second component in the PV or VV has marked spatial properties on the motion expressed by the main verb. Where there is semantic embedding of the second component on the first component of the PV or VV, the phrasal verb is considered to be semantically compositional (Jackendoff, 2002), and the VV syntactically compounded (Hashimoto and Bond, 2005: 5; Tamaoka, Lim, Hiromu, 2004: 234).

6.5.1 Double Translations

Interestingly, it was discovered that PVs had translations into both VVs and single verbs.

A case in point would be the phrasal verb *caved in* that was translated into 崩れ落ちる *kuzureochiru* “to fall in” in the TT.

(15) ... the roof was partly caved in, a picture of ruin.

屋根は	部分的に	崩れ落ち、	さながら	廃墟を	絵に
yane wa	bubun teki ni	kuzureochi,	sanagara	haikyo wo	e ni
roof TOP	partly DAT	caved in,	just like	ruins ACC	picture DAT

描いたようだった。

kaita you datta.

drew seems COP

“The roof was partly caved in; like a picture of ruin”

As above, the phrasal verb *cave* has the meaning of “collapse”. Its construction is metaphorical in a sense that it can be understood as a reference to the situation where something above falls inwards and downwards into a hollow beneath it, such as a cave, and the significance of this meaning has been transferred into the idiomatic meaning of “collapse”. The meaning of collapse is then exemplified in V₁ of the VV equivalent: 崩れる in itself means “collapse”, whereas in V₂ of the Japanese equivalent, 落ちる *ochiru* “to fall” supports the argument of V₁ in portraying the direction of movement to be downwards towards the ground. Additionally, as they are semantically equivalent, we linked them with “=” to denote this relationship of correspondence.

Yet, *kuzureochiru* can be used in a different way, and its meaning can be extended to denote the physical collapse of the person, as depicted below.

(16) ... but at that moment her knees seemed to give way and she fell to the ground.

...そのとき、 姉の膝の力 が抜けて、 床に 崩れ落ちました。

sonotiki, one no hiza no chikara ga nukete, toko ni kuzureochimashita.

that time sister's knees ADN strength NOM gone floor DAT collapsed

“At that time, her knees gave way and she collapsed on the ground”

FSID 10145 TSID 10184 NTU-MC

For the paired sentence above, the translator gave *kuzureochiru* as the translation equivalent for the single English verb *fall*. Other Japanese lexical equivalents could have been chosen, such as another single verb like 倒れる *taoreru* “to collapse” to describe the physical fall, but this lexical choice in the translation ultimately reflects stylistic preferences of the translator. This translation can be considered natural in a sense that *kuzureochiru* sounds natural in this context and can be understood in the same manner by Japanese readers. Since the sense of “fall” in *kuzureochiru* is also stronger, than in *fell*, this slight difference in semantic quality was denoted by the link “~”.

6.5.2 Analysis of adverbs in PVs in relation to V₂s in VVs

Apart from particles, we discovered that PVs with adverbs were translated into VVs as well.

Example	PV	VV	Corresponding translation from adverb
i	step forward	進み出る “to step forward”	forward → 出る <i>deru</i> “to come out”
ii	drive away “to get rid of”	取り除く “to remove”	away → 除く <i>nozoku</i> “to remove”
iii	cast round “to look around”	見回す “to look around”	round → 回す <i>mawasu</i> “to turn round”

Table 8: Summary of PVs with adverbs translated into VVs

Example (i) demonstrates a PV that has the main verb conveying the path of movement, and the adverb encoding directionality. The directional adverb is then translated into a Japanese motion verb encoded for path, 出る *deru* “to come out” (Tsujimura, 2002). For such compositional PVs whose meanings can be broken down into meanings of its constituents, literal translations of the adverb can be carried out to produce a corresponding V₂ in Japanese.

Example (ii) and (iii) depicts collocations whose metaphorical meanings cannot be deduced from its constituents. The adverbs away and round has its extended meanings translated literally in 除く *nozoku* “to remove” and 回す *mawasu* “to turn round” respectively. Here we see the V₂ of the VV holding the respective metaphorical meanings of the PV on the whole.

6.5.3 Analysis of particles in PVs in relation to V₂s in VVs

An interesting pattern that we discovered was the translation of the particle *up* into four different verbs in Japanese.

Example	PV	VV	<i>up</i> translated into
a	throw <u>up</u> “kick up”	巻き上げる	<i>ageru</i> “up”
b	creep <u>up</u> “to get near to”	歩み寄る	<i>yoru</i> “near”
c	knock <u>up</u> “to wake someone up”	たたき起こす	<i>okosu</i> “to cause to wake”
d	spring <u>up</u> “to stand up”	飛び起きる	<i>okiru</i> “to wake /get up”

Table 9: Variations in translations of *up*

Based on the context in the story, example (a) represents a regular use of *up*, where *up* indicates spatial movement upward. Interestingly, where the literal meaning of upward movement is denoted by the particle *up*, the same spatial quality is expressed in the Japanese translation equivalent *あげる* *ageru* “up”. We have seen this in the examination of *take out* previously, where the particle expressing spatial meaning can be paralleled by a directional verb in Japanese. In this case, in spite of the presence of a semantically corresponding counterpart in Japanese, *up* and *ageru* are not syntactically congruent since the forms belong to different word categories in their respective languages. Thus they can be defined as divergent correspondences (Johansson, 2007: 23).

Examples (b) and (c) show the alternative uses of *up*. In b, *up* defines the “closeness of the target position to some object” (Kilby, 1984: 105). In terms of cognitive schemas, *up* here represents a “goal oriented” field, where the direction of movement is towards a targeted goal; it does not signify a prototypical direction of movement upwards (Cook and Stevenson, 2006: 48). In example (c), *up* demonstrates the “path into the state of activity”, here signifying the upright vertical posture of a person (Cook and Stevenson, 2006: 48). In such cases which involves the extended use of the particle metaphorically, we noticed that if in the translation equivalent does not bear the sense denoted by the English particle, there is a tendency for explication to take place, that is, to “introduce information from the ST that is implicit from the context or situation” (Molina and Alibir, 2002). Here, the translation equivalent of “up”, *あげる* *ageru* “to go up”, does not include the senses of *up* in example (b) and (c). A new V_2 is then introduced in the VV to manifest overtly the implicated meanings of *up*. In b, the sense of close proximity in space is expressed by the Japanese verb *寄る* *yoru* “to be close to”, while in (c) the meaning of “get up” is transferred into the verb *起こす* *okosu* “to wake someone”.

Similar to extended uses of *up* in different contexts, *起きる* *okiru* lit. “to wake up” has extended meanings as well. The literal meaning of *okiru* which means “to be aroused from sleep” is not implied in *飛び起きる* *tobiokiru* (example (d)). Instead, the meaning of *okiru* is extended to mean “to move upright from a position of rest”, which is congruent to the meaning of *up* here: “to get on one’s feet”. In Japanese, *tobiokiru* is lexical VV since it is idiomatic and its meaning is not directly derived from its parts: *飛ぶ* *tobu* means “to fly” which is unrelated to the meaning of the whole. Example (d) is an instance of a match of two idiomatic correspondences in English and Japanese.

Examples (c) and (d) illustrate two V_2 verbs which form a causative-inchoative pair: *起きる* *okiru* is the inchoative verb which denotes a change in state in an entity; while *起こす* *okosu* is the causative-inchoative counterpart denoting something causing the change in state of an entity. As noted earlier in

Hashimoto and Bond (2005) for lexical VVs, the transitivity of VVs can be inherited from V₂ verbs. According to the theory of TRANSITIVITY HARMONY (Kageyama, 1993), the transitivity of the verb components limits the type of VV that can be compounded. Hence, this explains why *okiru* and *okosu* are not interchangeable in these two VVs, as it would create an incompatibility of transitivity of the V1 and V2, thereby affecting the transitivity of the entire compound as well. Thus, the pair 叩く *tataku* “to knock” and *okosu* are compatible as they are both transitive, to form an overall transitive VV, while 飛ぶ *tobu* “to fly” and *okiru* logically form an intransitive VV.

One noteworthy point here is that the transitivity of the PV is maintained in V₂ of the VV itself: *creep up* and *yoru* is intransitive, while *knock up* and *okosu* is transitive. In example (d) as well, 起きる *okiru* follows the intransitivity of the verb in the SL *spring up*. We thus carried out a further investigation on how transitivity of the PV influences the transitivity of V₂ in VVs, in the translation various particles in PVs, *out* and *in*.

Example	PV	VV	<i>out</i> translated into
e	took <u>out</u> “to remove”	取り出す	<i>dasu</i> “to cause to come out”
f	drag <u>out</u> “to lengthen”	生き延びる	<i>nobiru</i> “to extend”
g	shut <u>in</u> “to seclude oneself”	閉じこもる	<i>komoru</i> “to be confined in”
h	cave <u>in</u> “to collapse”	崩れ落ちる	<i>ochiru</i> “to drop”

Table 10: Variations in translations of *out*

Examples (e) shows the regular use of *out*, which depicts the change in location or the spatial movement from one place to another, and because 出す *dasu* lit. “to take out” manifests the same spatial meaning, *out* and *dasu* is another divergent correspondence. However, in (f), we see *out* translated into another verb 延びる *nobiru* “to extend”, and this is another case of EXPLICATION, as the extended meaning of *out* used to refer to the metaphorical extension of an entity is represented with *nobiru*. In (g) and (h) both V₂s represent metaphorical extended meanings of *in*, and as shown in the above translations, the meaning of *in* is disambiguated in the V₂ as above.

Moreover, all of the above examples also demonstrate the transference of its transitivity to V₂ in the VV translation equivalents. The transitive PV *took out* corresponds with the transitive verb *dasu*, and the rest of the intransitive PVs are also translated into intransitive V₂s, *nobiru*, *komoru* and *ochiru*. Overall we

see a high probability for a PV expressing movement to have its transitivity encoded in V_2 of the VV counterpart.

CHAPTER SEVEN

CONCLUSION

In this paper, we have examined translational shifts in a parallel English-Japanese corpus, which can occur across word classes as well as within the verb class. Translation divergences of verbs can occur due to the differences in language systems, as well as the translation techniques employed in the translation process. Comparisons of semantic categories in English and Japanese have allowed us to see how these two language systems relate to each other.

An interesting phenomenon uncovered in this study was the correlation of English phrasal verbs with Japanese compound verbs. We found that where the counterparts share the same meaning, they are similar in formal construct, especially for motion verbs. Transitivity of phrasal verbs expressing motion was discovered to be enacted on V_2 of the corresponding VVs. Further research could be performed to investigate the relationship of phrasal verbs expressing other semantic categories with the transitivity of the corresponding VVs.

Our study focused mainly on spatial V_2 verbs in VVs, thus future work could be done on aspectual or adverbial V_2 types for a better overview of the relationship between particles in PVs and V_2 in VVs. Nevertheless, our analysis of spatial particles has worked to disambiguate semantic ambiguities in phrasal verbs as well.

Through the cross-lingual analysis done on English and Japanese verbs, we hope that this can help second language learners of Japanese to better understand how concepts in English can be represented in different ways based on the context. Materials from this study can also be used by other scholars for future work on translations of English and Japanese verbs.

REFERENCES

- Backhouse, A.E (1993). *The Japanese Language: An Introduction*. London: Oxford University Press.
- Baker, Mona (1992) *In Other Words: a Coursebook on Translation*. London: Routledge.
- Baker, M. (1996). Corpus-based translation studies: The challenges that lie ahead. In H. Somers (Ed.), *Terminology, LSP and translation. Studies in language engineering in honour of Juan C. Sager* (pp. 175–186). Amsterdam: Benjamins.
- Biber, D., Stig J., Leech G., Conrad S., & Finegan E. (1999). *Longman Grammar of Spoken and Written English*. London: Longman.
- Bolinger, D. (1971) *The Phrasal Verb in English*. Cambridge: Harvard University Press.
- Breen, J., & Baldwin, T. (2009). Corpus-based Extraction of Japanese Compound Verbs. In *Proceedings of the 2009 Australasian Language Technology Workshop*, (ALTW 2009): 35-43.
- Brown, V., Mendes, E., & Natali, G. (1995). *False friends and bugs and bugbears*. Zanichelli, Bologna.
- Castejón, M. A. G. (2012). Contrastive analysis and translation study from a corpus linguistics perspective. *IJES, International Journal of English Studies*, 12(2): 111-132.
- Catford, J. C. (1965). *A linguistic theory of translation* (Vol. 74). London: Oxford University Press.
- Claridge, C. (2000). *Multi-word Verbs in Early Modern English: A Corpus-based Study*. Amsterdam: Rodopi.
- Clerck, D.B., Coleman, T., & Willems, D. (2013). Introduction: A multifaceted approach to verb classes. *Linguistics*, 51(4), 663-680.
- Cook and Stevenson, S. (2006). Classifying particle semantics in English verb-particle constructions. In *Proceedings of the Workshop on Multiword Expressions: Identifying and Exploiting Underlying Properties*, (pp 45–53). Sydney, Australia.
- Courtney, R. (1983). *Longman dictionary of phrasal verbs*. England: Longman.
- Cyrus, L. (2006). Building a resource for studying translation shifts. In *Proceedings of LREC 2006-Second International Conference on Language Resources and Evaluation, Genoa, Italy*, (pp 1240-1245).
- Denturck, K. (2012). Explicitation vs. Implication: A Bidirectional Corpus-Based Analysis of Causal Connectives in French and Dutch Translations. *Across Languages and Cultures*, 13 (2): 211–227.
- Dorr, B.J. (1992/1993). The Use of Lexical Semantics in Interlingual Machine Translation. *Machine Translation*, 7: 135-193.
- Dorr, B.J. (1993). *Machine Translation: A View from the Lexicon*. Cambridge, Massachusetts: MIT Press.
- Doyle, A.C. (1892). *The Adventures of Sherlock Holmes*. London, England: George Newnes Ltd.

- Fellbaum, C. (2005). WordNet and wordnets. In Brown, K. et al. (eds), *Encyclopedia of Language and Linguistics* (2nd edition.), pp. 665-670. Oxford: Elsevier
- Frazer, B. (1976). *The verb-particle combination in English*. New York: Academic Press.
- Gao, E.H. (2012). *Crosslingual comparison of linguistic phenomenon in English, Japanese and Chinese*. (Final Year Project). Retrieved from DR-NTU: <http://hdl.handle.net/10220/9431>
- Granger, S. (2003). The Corpus Approach: A Common Way Forward for Contrastive Linguistics and Translation Studies? In Granger S., Lerot J. and Petch-Tyson, S. (eds.). *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*. (pp. 17-29). Amsterdam: Rodopi.
- Hokari, T., Kumagami, M. & Akimoto, T.(2012).The Production of Japanese Verb-Verb Compounds by English, Chinese, and Korean Speakers: A Corpus Study. *九州大学言語学論集*, 33: 155-175.
- Hashimoto, C. and Bond, F. (2005). A Computational Treatment of V-V Compounds in Japanese. *Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar*, 143-156.
- Isahara, H., Bond, F., Uchimoto, K., Utiyama, M. & Kanzaki, K. (2008). Development of the Japanese WordNet. In *Proceedings of the 6th Conference on Language Resources and Evaluation (LREC 2008)*, Marrakesh, Morocco.
- Jackendoff, R. (2002). English Particle Constructions, the Lexicon, and the Autonomy of Syntax. In Nicole Dehé, Ray Jackendoff, Andrew McIntyre and Silke Urban (eds.). *Verb-Particle Explorations*. (pp. 67-94). Berlin / New York: Mouton de Gruyter.
- Johansson, S. (2007). *Seeing through Multilingual Corpora: On the use of corpora in contrastive studies* (Vol. 26). Amsterdam/Philadelphia: John Benjamins.
- Jenset, G.B. (2008). Basic Statistics for Corpus Linguistics. Retrieved from: http://www.researchgate.net/publication/265965333_Basic_statistics_for_corpus_linguistics
- Kageyama, T. (1993). *Bunpô-to Gokêsê*. (Grammar and Word Formation) Tokyo: Hituzi.
- Kilby, D. (1984). *Descriptive Syntax and the English Verb*. Great Britain: Croom Helm.
- Laviosa, S. (2002). *Corpus-based translation studies: theory, findings, applications* (Vol. 17). Amsterdam: Rodopi.
- Litosseliti, L. (2010). *Research Methods in Linguistics*. London: Continuum
- Lo Cascio, V., Boraschi, P., & Corda, A. (1995). Correspondence between senses and translation equivalents: automatic reversal of a bilingual dictionary. In *Translation and Meaning. Part 3 of the Proceedings of the 2nd International Maastricht-Lodz Duo Colloquium on 'Translation and Meaning'*, (pp. 221-231). Maastricht.
- Martin, S. E. (1975). *A Reference Grammar of Japanese*. Japan: Tuttle.
- Marello, C. (1989). *Dizionari bilingui: con schede sui dizionari italiani per francese, inglese, spagnolo, tedesco* (Volume 6). Zanichelli, Bologna.

- Malá, M. (2013). Translation counterparts as markers of meaning: The case of copular verbs in a parallel English–Czech corpus. *Languages in Contrast*, 13(2): 170-192.
- Molina, L. and Hurtado Albir, A. (2002). Translation techniques revisited: A dynamic and functionalist approach. *Meta: Journal des traducteurs / Meta: Translators' Journal*, 47(4): 498-512.
- Nariyama, S., Nakaiwa, H., & Siegel, M. (2005). Annotating honorifics denoting social ranking of referents. In *Proceedings of LINC the 6th international workshop on linguistically interpreted corpora*, (LINC-05): 91-100.
- Neagu, M. (2007). English verb particles and their acquisition: A cognitive approach. *Revista española de lingüística aplicada*, (20), 121-138.
- Olteanu, A.R. (2012). *A Holistic Approach to Phrasal Verbs*. Retrieved 23 September, 2014, from: http://lectura.biblioteca.digital.ro/Olteanu_Adriana/A_holistic_approach_to_phrasal_verbs.pdf
- Olohan, M. (2004). *Introducing corpora in translation studies*. Routledge.
- Paik, K., Ohtake, K., Bond, F., and Yamamoto, K. (2004). Source Language Effect on Translating Korean Honorifics. *Proceedings of the 5th International Conference on Intelligent Text Processing and Computational Linguistics*, (CICLing-2004), LNCS 2945: 334-337.
- Sag, I.A., Baldwin T., Bond F., Copestake A. and Flickinger D. (2002). *Multiword Expressions: A Pain in the Neck for NLP*. In *Lecture Notes In Computer Science*, 2276: 1-15.
- Serbina T. (2013) Construction Shifts in Translations. *Constructions and Frames*, 5(2): 168–191.
- Shibatani, M. (1990). *The languages of Japan*. Cambridge University Press.
- Snider, N., & Arnon, I. (2012). A unified lexicon and grammar? Compositional and non-compositional phrases in the lexicon. In S. Gries & D. Divjak (Eds.). *Frequency effects in language representation*. (pp. 127-163). Berlin: Mouton de Gruyter
- Talmy L. (1991). Path to Realization: A Typology of Event Conflation. *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society*, 480-519.
- Talmy, L. (2000). *Toward a Cognitive Semantics: Typology and Process in Concept Structuring (Vol 2)*. Cambridge: MIT Press.
- Tamaoka, K., Lim, H., Hiromu, S. (2004). Entropy and Redundancy of Japanese Lexical and Syntactic Compound Verbs. *Journal of Quantitative Linguistics*, 11(3): 233–250
- Tsujimura, N. (1996). *An Introduction to Japanese Linguistics*. Cambridge, USA: Blackwell.
- Tsujimura, N. (2002). Motion and path in the acquisition of Japanese intransitive verbs. In *Proceedings of Stanford University Child Language Forum*, 21-30.

Uchiyama, K., & Ishizaki, S. (2001). A study of translation rules for the Japanese compound motion verbs. *Proceedings of the AFNLP PACLING 2001*.

Uchiyama, K., Baldwin, T., & Ishizaki, S. (2005). Disambiguating Japanese compound verbs. *Computer Speech & Language*, 19(4): 497-512.

Vermeulen, R. (2009). Topics, contrast and contrastive topics in Japanese. In *MIT Working Papers in Linguistics 58: Proceedings of the 5th Workshop on Altaic Formal Linguistics* (pp. 361-372).

Vinay, J., & Darbelnet, J. (1995). *Comparative Stylistics of French and English: A Methodology for Translation*. (J.C. Sager & M. Hamel, Trans.). Amsterdam and Philadelphia: John Benjamins. (Original work publish 1958).

Vlad, Eduard. (1998) *English Group Grammar*. Constanta: Tipografia Universitatii Ovidius.

Xu, J., and Li, X. (2014). Structural and semantic non-correspondences between Chinese splittable compounds and their English translations: A Chinese-English parallel corpus-based study. *Corpus Linguistics and Linguistic Theory*, 10(1): 79-101.