

**Analysis of Translation Shifts
Using Systemic Functional Linguistics**

**Textuality of news translation
between Japanese and English**

Mika Takewa

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
School of Languages, Cultures and Societies
Centre for Translation Studies

November, 2018

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

© 2018 The University of Leeds and Mika Takewa

Acknowledgements

This work would have been difficult without the partial fee scholarship the School of Languages, Cultures and Societies, University of Leeds awarded me. I am particularly grateful to Professor Claire Horness, former Head of School who was supportive of my wish to undertake a PhD and consulted a School committee which agreed to provide me with this opportunity.

My sincere thanks go to my supervisors, Dr Serge Sharoff and Dr Martin Thomas, whose guidance and patience enabled me to complete this thesis. Sergei constantly reminded me of the big picture of my project while Martin always gave me detailed, constructive feedback based on his thorough reading of my manuscripts. I am also fortunate to have a wonderful exam panel: Professor Erich Steiner from the Universität des Saarlandes, Dr Akiko Sakamoto from the University of Portsmouth and Professor Jeremy Munday from the University of Leeds. I thank them all for their time to read my thesis as well as their enthusiastic questions during the viva.

I would like to thank my colleagues in East Asian Studies, Mr Kazuki Morimoto and Ms Yuka Oeda for their understanding and support, and Dr Martin Ward for proofreading my thesis. Professor Ruru Li and Dr Ning Yi have been an inspiration to me. Professor El Mustapha Lahlali and Dr Ruba Khaman in Arabic and Middle Eastern Studies also gave me encouragement and advice whenever they saw me. Mrs Karen Priestley, the School's Education Service Officer also offered reassurance and administrative help throughout my study.

Special thanks go to Professor Tony Hartley, former Director of Centre for Translation Studies, University of Leeds, who encouraged me to embark on a doctorate and provided me with continuous inspiration. I also appreciate his wife Masumi's hospitality and friendship. Two relatively recent students of Tony's, Dr Alina Secala and Dr Dragoş Ciobanu cheered me up whenever I had a chance to talk to them.

Friends have been source of energy. My gratitude and affection go to Vivian and Marilena, with whom I shared happy laughs at the beginning of our PhD, Tomoko Zama, Jessica and Peng Chiu, who helped me towards the end, and Agnieszka, Catherine, Kathryn, Kaeko & David Ward and Miwako Kashiwagi, whose friendship sustained me throughout.

Finally, I thank my mother for accepting my choices and supporting me whatever my choices are, my late father for exemplifying it is never too late to go to university, and my aunt Hiroko and her late husband Moriya for telling me not to miss an opportunity for education.

Abstract

Translated news are target-oriented and therefore translation shifts are general practice in news translation. However, little research has been conducted to reveal specific shifts involving English and Japanese. With a view to contribute to translator training, the main objective of this project is to establish patterns and motivations of shifts in news translation between English and Japanese. Two types of shifts are investigated. First, information content; additions, deletions or moves, and second, information flow, focusing on Theme.

This project is corpus-based and a specialised corpus is compiled, consisting four sub-corpora; original news articles in English and in Japanese, their translations into Japanese and into English. They form parallel corpora as well as comparable. All the data are manually annotated for additions, deletions and moves and also for Theme using the UAM CorpusTool, adopting Systemic Functional Linguistics as framework for analysis.

Literature claims most news articles undergo additions, deletions or moves during translation. In my study, additions are more common in English translations while in Japanese translations it is deletions. Rank-wise, below-the-clause shifts occur throughout the text while above-the-clause shifts tend to occur towards the beginning or end of the text.

Regarding information flow, general trend of Theme realisation in translated news are similar to that of the TL in general. In addition Theme choice at the beginning of text units moves towards the TL norm, while still showing the ST influence. Particular types of Theme appear exclusively at specific locations in text in Japanese translation, in line with the non-translation Japanese texts. This suggests genre conventions of the TL are at work in Theme choice.

News translation into English and Japanese indicates processes of standardisation and ST interference. Shifts concerning information content and flow are motivated by the genre conventions and the target reader's relevance to and expectations for news articles.

Table of Contents

Acknowledgements	iii
Abstract	iv
Table of Contents	v
List of Tables	ix
List of Figures	x
Chapter 1 Introduction	1
1.1 Statement of the problem	1
1.2 Research aims, objectives and questions	5
1.3 Thesis outline	8
Chapter 2 Systemic Functional Linguistics as framework of the study	10
2.1 Language seen from the point of view of SFL	10
2.1.1 Language as a resource for making meaning	10
2.1.2 Meaning making by choice	13
2.1.3 Meaning making and metafunctions.....	15
2.2 SFL and translation studies.....	20
2.3 Text and textuality	23
2.3.1 The nature of text	23
2.3.2 Textuality and Theme.....	24
Chapter 3 Translation shifts and news articles	30
3.1 Investigation of translation shifts	30
3.1.1 Translation shifts	30
3.1.2 Research on translation shifts	33
3.2 News translation.....	36
3.2.1 Genre and register	36
3.2.2 News as a genre	37
3.2.3 News structure and discourse	39
3.2.4 News translation and shifts	43
Chapter 4 Research design and analytical approaches	46
4.1 Corpus and tool.....	46
4.2 Unit of analysis	51
4.2.1 T-unit to analyse Theme realisation	51
4.2.2 T-unit in English	52
4.2.3 T-unit in Japanese.....	55

4.2.3.1 Unit of analysis, Theme and particle 'wa'	55
4.2.3.2 T-unit and the suspensive form	62
4.2.3.3 The suspensive form and taxis.....	69
4.2.3.4 Identifying T-unit in Japanese.....	74
4.2.4 Discourse unit to analyse thematic development	74
4.2.4.1 Identifying the nucleus and satellites.....	75
4.2.4.2 Annotating the nucleus and satellites	82
4.3 Identifying Theme.....	84
4.3.1 Boundary of Theme and Rheme	84
4.3.2 Theme in English.....	86
4.3.3 Theme in Japanese.....	89
4.3.3.1 Implicit or explicit Theme.....	91
4.3.3.2 Simple and multiple Theme	96
4.3.3.3 Experiential Theme	100
4.3.3.4 Absolute Theme	103
4.3.3.5 Logical Theme.....	104
4.3.3.6 Projection and Theme	105
4.3.4 Annotating Theme	109
4.3.4.1 Systems of Theme and general principles for annotation	109
4.3.4.2 Guidelines specific to English.....	112
4.3.4.3 Guidelines specific to Japanese	113
4.4 Identifying additions, deletions and moves.....	113
4.4.1 Rank of additions, deletions and moves.....	113
4.4.2 Annotating additions, deletions and moves	117
Chapter 5 Theme selection and translation shift.....	120
5.1 Theme in En-orig and Ja-orig.....	120
5.1.1 Simple and multiple Theme	120
5.1.2 Ideational Theme.....	123
5.1.3 Summary of comparison	125
5.2 Theme in En-trans (compared with En-orig & Ja-orig)	126
5.2.1 Simple and multiple Theme	127
5.2.2 Textual and Interpersonal Theme.....	128
5.2.3 Ideational Theme.....	130
5.2.3.1 Unmarked Theme.....	131
5.2.3.2 Marked Theme	132

5.2.3.3 Enhanced Theme	136
5.2.4 Patterns of Theme selection in En-trans	137
5-3 Theme in Ja-trans (compared with Ja-orig & En-orig).....	140
5.3.1 Simple and multiple Theme	140
5.3.2 Interpersonal and Textual Theme	142
5.3.3 Ideational Theme	143
5.3.3.1 Participant Theme	145
5.3.3.2 Circumstance Theme	145
5.3.3.3 Logical Theme.....	147
5.3.4 Patterns of Theme selection in Ja-trans	150
5.4 Shifts in En-trans and Ja-trans	152
Chapter 6 Textual coherence of news and translation shift.....	154
6.1. Text structure and field of discourse (information content).....	154
6.1.1. Frequency of additions, deletions and moves in translation.....	155
6.1.2. Rank and type of additions, deletions and moves in translation.....	156
6.1.2.1. Additions	156
6.1.2.2. Deletions	160
6.1.2.3. Moves.....	163
6.1.3. Patterns of additions, deletions and moves	166
6.2 Text structure and mode of discourse (thematic progression).....	168
6.2.1 Thematic progression in En-orig and Ja-orig.....	169
6.2.1.1 Theme at the beginning of the text nucleus.....	169
6.2.1.2 Theme at the beginning of satellite	186
6.2.1.2 Comparison of En-orig and Ja-orig	192
6.2.2 Thematic progression in translation.....	193
6.2.2.1 Theme at the beginning of text nucleus.....	194
6.2.2.2 Theme at the beginning of satellite	202
6.2.2.3 Comparison of En-trans and Ja-trans	209
6.2.3 Translation shift in thematic progression	210
6.2.3.1 Comparison of translation and target language original	210
6.2.3.2 Comparison of translation and source text	217
6.2.4 Patterns and motivations of shifts in thematic progression	224

Chapter 7 Conclusions and further research	226
7.1 Main findings and contributions	226
7.1.1 Translation shifts in Theme selection (general frequency)	226
7.1.2 Translation shifts in Theme selection (at generic stages)	230
7.1.3 Translation shifts in additions, deletions and moves	232
7.1.4 Theme in Japanese news	234
7.1.5 Contributions	236
7.2 Limitations	238
7.2.1 Corpus.....	238
7.2.2 Methodology.....	239
7.2.2.1 Implicit Theme	239
7.2.2.2 Consistency of analysis	239
7.2.2.3 Scope	240
7.3 Further research.....	240
List of References	242
List of Notations and symbols	257

List of Tables

Table 5-1: Summary of contrast between En-orig and Ja-orig	126
Table 5-2: Adjunct Theme types: significance (En-trans)	135
Table 5-3: Orientation of Theme realisation (En-trans)	138
Table 5-4: Circumstance Theme types: significance (Ja-trans).....	146
Table 5-5: Orientation of Theme realisation (Ja-trans)	151
Table 6-1: Number of texts involving additions, deletions or moves.....	155
Table 6-2: Number of texts in which additions, deletions or moves co- occur	155
Table 6-3: Number of texts involving additions according to the semantic rank.....	157
Table 6-4: Number of texts involving deletions according to the semantic rank.....	161
Table 6-5: Number of texts involving moves according to the semantic rank	163
Table 6-6: Number of texts involving moves according to the semantic rank across the boundary of the text nucleus and satellites.....	164
Table 6-7: Number of texts involving moves from the text nucleus towards satellites according to the semantic rank.....	164
Table 6-8: Number of texts involving moves towards the text nucleus from satellites according to the semantic rank	165

List of Figures

Figure 2-1: SFL's stratified model of language (from Martin 2014).....	10
Figure 2-2: System of POLARITY in English	13
Figure 2-3: SFL's stratified model of language (from Bartlett & O'Grady 2018)	16
Figure 3-1: Nucleus and Satellites model (Thomson <i>et al.</i> 2008: 219)	40
Figure 4-1: Data source and size	48
Figure 4-2: Relationship of sub-corpora	49
Figure 4-3: Suspensive form	65
Figure 4-4: Possible postpositional phrases in the corpus	66
Figure 4-5: Annotation example of the nucleus on the UAM CorpusTool....	83
Figure 4-6: The most general system of THEME (Teruya 2006: 133)	90
Figure 4-7: System network of THEME (Thomson 2001: 119)	91
Figure 4-8: System of THEME in English (adapted from Thompson 2004)	110
Figure 4-9: System of THEME in Japanese	111
Figure 4-10: Semantic rank of translation shift	115
Figure 5-1: Simple & multiple Theme (En-orig & Ja-orig)	121
Figure 5-2: Multiple Theme type (En-orig & Ja-orig).....	121
Figure 5-3: Textual Theme type (En-orig & Ja-orig)	122
Figure 5-4: Unmarked or other Theme (En-orig & Ja-orig).....	123
Figure 5-5: Marked Theme type (En-orig & Ja-orig)	124
Figure 5-6: Adjunct/Circumstance Theme type (En-orig & Ja-orig)	125
Figure 5-7: Simple & Multiple Theme (En-trans)	127
Figure 5-8: Multiple Theme type (En-trans).....	128
Figure 5-9: Textual Theme type (En-trans)	129
Figure 5-10: Ideational Theme (En-trans)	131
Figure 5-11: Marked Theme type (En-trans)	132
Figure 5-12: Adjunct/Circumstance Theme type (En-trans)	134
Figure 5-13: Enhanced Theme type (En-trans)	136
Figure 5-14: Theme realisation gaps between En-trans and En-orig	137
Figure 5-15: Simple & Multiple Theme (Ja-trans).....	140
Figure 5-16: Multiple Theme type (Ja-trans).....	141
Figure 5-17: Textual Theme type (Ja-trans)	143
Figure 5-18: Relative Theme type (Ja-trans).....	144

Figure 5-19: Experiential Theme type (Ja-trans)	144
Figure 5-20: Circumstance Theme type (Ja-trans).....	146
Figure 5-21: Logical Theme type (Ja-trans)	148
Figure 5-22: Projection Theme type (Ja-trans).....	149
Figure 5-23: Theme realisation gaps between Ja-trans and Ja-orig	150
Figure 6-1: Subject Theme or non-Subject Theme at the beginning of the text nucleus (En-orig & Ja-orig).....	171
Figure 6-2: Non-Subject Theme type at the beginning of the text nucleus (Ja-orig)	178
Figure 6-3: Circumstance Theme type at the beginning of the text nucleus (Ja-orig)	183
Figure 6-4: Subject Theme or non-Subject Theme at the beginning of satellite (En-orig & Ja-orig).....	186
Figure 6-5: Non-Subject Theme type at the beginning of satellite (En- orig & Ja-orig).....	190
Figure 6-6: Circumstance Theme type at the beginning of satellite (Ja- orig).....	191
Figure 6-7: Subject Theme or non-Subject Theme at the beginning of the text nucleus (En-trans)	194
Figure 6-8: Subject Theme or non-Subject Theme at the beginning of the text nucleus (Ja-trans).....	195
Figure 6-9: Non-Subject Theme type at the beginning of the text nucleus (Ja-trans)	199
Figure 6-10: Circumstance Theme type at the beginning of the text nucleus (Ja-trans)	201
Figure 6-11: Subject Theme or non-Subject Theme at the beginning of satellite (En-trans).....	203
Figure 6-12: Subject theme or non-Subject Theme at the beginning of satellite (Ja-trans).....	203
Figure 6-13: Non-Subject Theme type at the beginning of satellite (En- trans).....	206
Figure 6-14: Non-Subject Theme type at the beginning of satellite (Ja- trans).....	207
Figure 6-15: Circumstance Theme type at the beginning of satellite (Ja- trans).....	208
Figure 6-16: Subject Theme or non-Subject Theme at the beginning of the text nucleus (En-trans & En-orig)	211
Figure 6-17: Subject or non-Subject Theme at the beginning of the text nucleus (Ja-trans & Ja-orig)	212
Figure 6-18: Non-Subject Theme type at the beginning of text nucleus (Ja-trans & Ja-orig)	212

Figure 6-19: Circumstance Theme type at the beginning of text nucleus (Ja-trans & Ja-orig)	213
Figure 6-20: Subject or non-Subject Theme at the beginning of satellite (En-trans & En-orig)	214
Figure 6-21: Subject or non-Subject Themes at the beginning of satellite (Ja-trans & Ja-orig)	215
Figure 6-22: Non-Subject Theme type at the beginning of satellite (En- trans & En-orig)	215
Figure 6-23: Non-Subject Theme type at the beginning of satellite (Ja- trans & Ja-orig).....	216
Figure 6-24: Circumstance Theme type at the beginning of satellite (Ja- trans & Ja-orig).....	217
Figure 6-25: Subject Theme or non-Subject Theme at the beginning of the text nucleus (En-trans & Ja-orig).....	218
Figure 6-26: Subject Theme or non-Subject Theme at the beginning of the text nucleus (Ja-trans & En-orig).....	221
Figure 6-27: Subject Theme or non-Subject Theme at the beginning of satellite (En-trans & Ja-orig).....	222
Figure 6-28: Subject Theme or non-Subject Theme at the beginning of satellite (Ja-trans & En-orig).....	223
Figure 6-29: Non-Subject Theme type at the beginning of satellite (En- trans & Ja-orig).....	223
Figure 6-30: Non-Subject Theme type at the beginning of satellite (Ja- trans & En-orig)	224

Chapter 1 Introduction

1.1 Statement of the problem

Translating news articles entails much more than a linear transfer of the content from source text (ST) to target text (TT), regardless of the language pair (Bielsa & Bassnett 2009). The present study focuses on a specific language pair of Japanese and English, and investigates what diversions occur in translated news. Example 1-1 provides such examples in the opening paragraphs of 'Famed writer's English lessons published', which is translated from Japanese into English.

1-1 A copy of lecture notes taken during a series of English lessons taught by writer Lafcadio Hearn (1850-1904) at a school in Kumamoto has been discovered at a library in Toyama.

The notes have now been converted into a book recently published by *Gen Shobo* called "Lafcadio Hearn's English Lessons." The text includes the transcribed lecture notes alongside their Japanese translation.

Experts say the notes are important because they shed light on the activities of Hearn, who is also known as *Yakumo Koizumi*, during his days in Kumamoto, which had been largely a mystery due to a lack of documentation.

4 April 2013, the Japan news *by Yomiuri*

Example 1-2 below is a glossing translation of the corresponding part of 1-1 in the original text written in Japanese (my translation). It preserves the content and the text structure, paragraph boundaries and order. Clause structure is preserved as much as possible too but the basic word order of SVO in English is respected over the SOV order of Japanese.

1-2 A copy of lecture notes that a student took during the English lessons novelist Lafcadio Hearn (*Yakumo Koizumi*, 1850-1904) who is known for *Kwaidan* and the like conducted at the Kumamoto 5th Higher Secondary School (Kumamoto city) under the old system has been found at the library of the University of Toyama (Toyama city) and published by *Gen Shobō* as *Lafcadio Hearn's English Lessons*, including the Japanese translation along with the original English.

Researchers point out [they are] important data that fill the gap of [his] activities during [his] days in Kumamoto that lack documentation and tell an aspect [of his] as an excellent English teacher.

3 April 2013, *Yomiuri Shimbun*

This short excerpt reveals a number of interesting differences in the TT (1-1) in comparison to the ST (1-2). The first two paragraphs in 1-1, which consist of three sentences, are originally written in one paragraph, in fact in one sentence. In terms of the news focus, the translation gives an impression that it is the discovery of the lecture notes but the original article places essentially an equal weight on discovery of the notes and its publication, if not more on the latter. The detailed information such as which library the notes were found in or which school Hearn taught at disappears, the book title of *Kwaidan* [Horror stories] is also omitted, Hearn's Japanese name is introduced very late and the fact that he was an English teacher as well as a novelist is omitted.

These are informative as examples of translation not exactly matching the ST. They are called 'translation shifts' (Catford 1965), although Catford's definition is rigidly linguistic and nowadays the term is used more flexibly (Chapter 3 Section 1). Bearing in mind that a lot more than purely linguistic issues take place in news translation (Bielsa & Bassnett 2009), the present study observes and analyses translation shifts as widespread professional practice in news translation. This topic has been researched little in general, and in particular with respect to Japanese and English.

A particular focus is placed on the textuality of translated news, i.e. readability and acceptability. News translation is target-oriented and the translated news articles are tailored to suit the target reader's needs and expectations (*ibid.* 10). An examination of professional behaviour both quantitatively and qualitatively as it appears in the product of news translation will contribute to understanding what kind of shifts are taking place in professional news translation, if there are any patterns of them as well as the possible reasons behind them. These shifts are the results of translator's choice of the strategies as Vinay and Darbelnet (1958) explain, although it is also possible that they are errors and/or involving a process of editorial interventions.

Translators are said to rely on 'their experiential, creative and intuitional (tacit knowledge) resources' (Wilss 1996: 3) in deciding how best to deliver the ST message in the target language (TL). They are also said to acquire 'techniques' of translation through their experience (Anzai *et al.* 2005: 71). An empirical study of translation shifts using a specialised corpus will shed light

on the current situation of shifts in professional news translation from Japanese into English, and from English into Japanese, which is an understudied area in translation studies (Chapter 3).

The investigation of professional translation shifts will also contribute to addressing the gap that exists between the professional translation and trainees'. Students I teach at the Centre for Translation Studies, University of Leeds, often recognise the necessity to add explanations to assist the Japanese readers' understanding. However, they are neither sure if it is acceptable to do so, nor do they know how much information to add if it can be done.

Omitting information or changing the narrative flow seems hardly ever to occur to them as possible translation strategies. They say they don't know to what extent '意訳 [meaning translation]' is allowed and that they ought to keep close to the ST. When they say 'be loyal to the ST', their concern seems to be only with the syntactic structure. Their translations and comments during in-class discussions show their reluctance to employ a TL word whose word class is different from the original in the ST, or change word order when it is possible to keep the same order as the ST, i.e. systemically equivalent option is available. However, Teich (2001) points out that in reality a systemically equivalent option is often rejected during translation and shifts occur for register reasons.

Structural differences between English and Japanese are one of the obvious causes for translation shifts but they are not the only reasons. With respect to news translation, other issues such as differences of the level of familiarity of the content or interest in it between the ST and the TT readers, relevance of the news to the source language (SL) and TL societies, different rhetorical conventions of writing news articles in both languages are also likely to prompt translation shifts. In addition, it must be noted that the text function of the translated news in the target society is different from that of the ST in the society the original news is aimed at.

In order for the translated news to make sense or communication in the TT to be properly carried out, an intervention from the translator is often required. Causes for shifts vary depending on the genre of the ST, topic, target readership, and so on. It can be explained that one of the problems of the Japanese trainee translators' described above are the lack of the notion of audience. Anzai *et al.* (2005: 49-65) relate this to the translation exercise they do in English lessons in Japan. Translation in English class is a supplementary means for learners to understand the language; therefore it is important for

them to depict every word in a sentence even if that results in unnatural Japanese sentences. Trainee translators must realise that the function of translation is different in the professional setting, in which there is a communicative function to fulfil.

In order to analyse translation shifts Systemic Functional Linguistics (SFL) is employed. SFL views language as resource for making meaning. Meanings are realised by lexicogrammar; the combination of grammar and lexis. In discussing translation shifts, this realisational concept is useful to analyse linguistic differences between the ST and TT as different realisations of the same meaning in different languages, and therefore enables a comparison.

In addition to the realisational relationship between the meaning and lexicogrammar, SFL situates language in context. This means contexts influence language use, i.e. we use language differently in different situations (Thompson 2004: 41). If the context is the same, then there is a characteristic language use, i.e. register. SFL's approach to see language as systemic choice supports the analysis of translation shifts taking into account of the context of translation; topic, participant roles and how the communication is organised. Focusing on a particular register of news articles allows the present study to observe typical language use in that specific context of situation. Also focusing on the translation between English and Japanese will help recognise translation shifts specific to this language pair.

In addition to the register perspective, which is concerned with typical and frequent language use in a particular context and thus defined text-internally (Sharoff 2017, 2018), the notion of genre is important in the present study. Thompson explains genre accessibly as 'register + purpose' (2004: 42). According to Sharoff, genre is defined text-externally (2017, 2018) and a purpose of a text is external to text. A text's purpose influences how a text is structured (Thompson 2004: 42) and a genre perspective offers a basis for analysing shifts in relation to the text structure of news, which follows a culturally expected way of a specific genre of news articles in each culture.

A genre perspective means the text as a whole is the target of analysis. There are many research in contrastive analysis of English and Japanese that offer valuable insight in terms of linguistic differences. However, most of them is conducted at the sentence level and out of context, thus discussions above the sentence level are scarce.

In SFL the clause is the basic unit of analysis but how clauses are connected as a whole text can also be discussed. Focusing on the textual metafunction,

in particular Theme enables an analysis of translation shifts in terms of textuality at different levels such as the T-unit, structure components and the whole text. It will not only contribute to translation studies but also to contrastive analysis between English and Japanese.

Another aspect of contribution the current study aims to offer is for a better understanding of linguistic features of the register of news articles in Japanese. Findings of contrastive analysis in many cases are based on literary texts and qualitatively examined (Kawahara 2011c). Quantitative research using a specialised corpus of the news articles will be useful to understand a typical language use in the register of news articles as well as a future comparison with the available findings in the register of literary texts.

Investigation of the Japanese language using SFL has been developing (Thomson 1998, 2001, 2005a; Tatsuki 2000, 2004; Teruya 2006) but research on translation within the framework of SFL with respect to Japanese has been few and limited to small scale case studies. A systematic analysis of translation with the language pair of English and Japanese using SFL addresses a gap in translation studies as well as SFL of Japanese.

Apart from translation studies and SFL of Japanese, this project seeks to make a contribution to a better understanding of news discourse in Japanese, where little research has been conducted on linguistic characteristics of news articles. Observing linguistic features in relation to the conventional structure of news articles will shed light on how a particular structure component is likely to begin, and what type of information is expected in a particular type of structure component.

1.2 Research aims, objectives and questions

The overall aim of the present study is to find patterns of translation shifts taking place in the professional translation of news articles from Japanese into English and from English into Japanese. This study is corpus-based and focuses on textual meaning. Through critical analysis of translations by professional translators, I hope to gain understanding of translation shifts that take place in news translation and establish a pattern of them. The findings will provide trainee translators with evidences to guide them for or against shift as well as to support the shifts they attempt.

The aims of the study are to:

- Investigate what kinds of translation shifts take place in professional news translation and to discover whether they have any patterns;

- Investigate factors for translation shifts;
- Explore Japanese news articles in terms of linguistic features and textual structures in comparison to those written in English.

The objectives of the study are to:

- Identify a pattern of shifts in news translation;
- Provide a contrastive model of news discourse in terms of linguistic features in Japanese;
- Provide trainee translators with the evidences to guide them for or against shifts in translating news;
- Construct specialised corpora of hard news articles; written in Japanese, written in English, Japanese translation of English articles and English translation of Japanese articles.

Three sets of research questions are posed in order to identify patterns of shifts. Each set introduced below has a different focus; Theme choice in terms of frequency, Theme choice at specific locations in the text, and added, deleted and moved information in translation. In addition, a further set of questions is prepared to discover the thematic features of news articles written in Japanese.

The first set, questions 1-2, concerns Theme selection in terms of frequency, which is then discussed in Chapter 5. The questions are as follows:

- 1) What are the patterns of Theme realisation in translation?
 - What elements are frequently realised as Theme in translation into English and into Japanese?
- 2) Is the Theme realisation in translation ST oriented or TL oriented?
 - How does Theme realisation in translation compare with that of the comparable corpus in either English or Japanese as the TL?
 - How does Theme realisation in translation compare with that of the parallel?

The questions above are intended for an investigation of Theme from the point of view of 'register' (Halliday & Hasan 1989), in that they are about the frequency of different types of Theme in the specific register of news articles. The second and third sets of questions on the other hand aim to examine

Theme from the viewpoint of 'genre' (Halliday 1977/2003), in that they involve the notion of generic structure of news as part of texture creating resources.

Thus, the second and the third sets of questions concern textual coherence including additions, deletions or moves of information, and thematic development. The second set, questions 3-5, concerns translation shifts in terms of information content, and the related findings are discussed in Chapter 6 Section 1. The questions are as follows:

- 3) How frequently do additions, deletions or moves occur in news translation?
- 4) Are there any patterns of additions, deletions or moves?
 - At what semantic rank do they occur?
 - What kinds of information are added, deleted or moved?
 - Is there a link between the ranks and types of information that are added, deleted or moved?
- 5) Do English translations and Japanese translations have similar patterns for additions, deletions or moves?

In order to answer these questions I examine the relationship between the text structure and information that are added or deleted in news translation. I will also examine moves of information from one place in the source text to another in the translation, bearing in mind that they are also relevant to information flow.

The third set, questions 6-8, aims to discover patterns of translation shifts in terms of thematic progression and the findings are discussed in Chapter 6 Section 2. The questions are as follows:

- 6) Are there any patterns of Theme realisation in news translation at specific locations in text; at the beginning of different structure components?
 - Are Themes realised at the beginning of a specific structure component distinguishable from those at the beginning of different type of structure component?
 - What are the factors influencing translation shifts, if they occur?
- 7) How does Theme realisation at the beginning of different structure components in translation compare with the comparable TL corpus in English or Japanese?

- 8) How does Theme realisation at the beginning of different structure components in translation compare with the parallel ST corpus in English or Japanese?

In order to answer these questions, I examine if Theme functions as an orienter of the message more globally than the local clause rank. These questions concern the internal thematic structure of the text, and they are set in an attempt to establish a path towards understanding the relationship between the Theme and functional segments of text in translation.

In addition to understanding and establishing patterns of shifts in translation, a further aim is to gain a contrastive understanding of the discourse structure of Japanese news articles and how Theme relates to it in comparison to English. To this end a further set of questions are presented below, adapted from those earlier in the present section. They are addressed in Chapter 6 Section 2.

- 9) Are there any patterns of Theme realisation at specific locations in Japanese news articles as non-translation; at the beginning of different structure components?
- Are Themes realised at the beginning of a particular structure component distinguishable from those at the beginning of a different type of structure component?
 - How does Theme realisation at different structure components compare with that in the whole corpus of the original Japanese texts with no reference to realisation locations in the text?
- 10) How does Theme realisation at different structure components compare with the original ST corpus in English?

These research questions are investigated through the empirical analysis of a sample of news articles and their translations. Research design is explained in Chapter 4.

1.3 Thesis outline

This chapter has introduced the problem the present study aims to investigate and identified the gap in knowledge it aims to address. It has also outlined aims and objectives as well as research questions.

Following this, **Chapter 2** introduces Systemic Functional Linguistics as the framework of study. It describes how SFL sees language and why it is useful to investigate translation shifts.

Chapter 3 provides the definition of the translation shift this study employs and describes previous research conducted on them. It also explains news articles as a genre and their conventional structure.

Chapter 4 sets out the design of research, which explains how the analysis is conducted; what data is used, what the analytical tool is, how the analysis is carried out in order to answer the research questions. This chapter also discusses several methodological issues for consistent analysis both in English and in Japanese.

Chapter 5 presents the findings of the quantitative analysis of Theme choice from the viewpoint of register, i.e. frequency. Theme choice in translation will be discussed in comparison to the comparable corpus of non-translated texts in the TL as well as the parallel corpus of the STs.

Chapter 6 concerns textuality of translated news and translation shifts are observed in relation to the text structure. The chapter presents the findings of the analysis in terms of information content; what is added, deleted and moved in translation, and information flow; how Theme develops.

Chapter 7 summarises the study and presents main findings and contributions. It also recognises limitations of the study and indicates directions for further study.

Chapter 2

Systemic Functional Linguistics as framework of the study

This chapter explains the reasons why Systemic Functional Linguistics (hereafter SFL) is suitable as a framework to investigate translation shifts. I will first describe the way SFL sees language, continuing to an account of SFL's applicability to translation studies, and, finally, explore Theme, an aspect of SFL that is a major focus of this study.

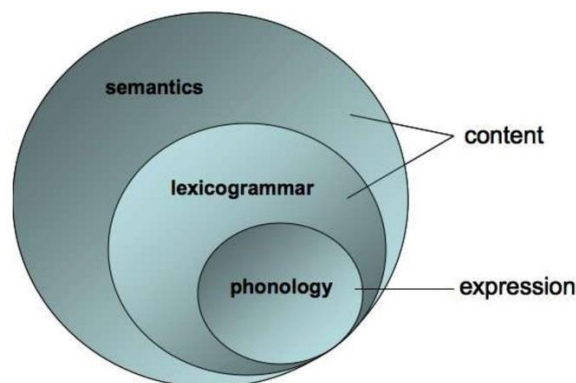
2.1 Language seen from the point of view of SFL

2.1.1 Language as a resource for making meaning

SFL was developed in the 1960s by M.A.K. Halliday, who had been influenced by the Prague linguistic circle, British linguist J.R. Firth, as well as British anthropologist B. Malinowski. It is a functional theory of language concerning how language is used to fulfil its functions in society.

According to SFL, language is a resource for making meaning, having a coding system on three levels, or strata; semantics, lexicogrammar which is a combination of grammar and lexis (vocabulary), and phonology/graphology.

Figure 2-1: SFL's stratified model of language (from Martin 2014)



As translation is essentially an activity of interpreting the meaning created in the source text (ST) and representing it in the target language (TL), semantics and lexicogrammar are the focus of this study.

The significant aspect of this model in the present project, which investigates translation shifts, is the relationship between semantics and lexicogrammar. A source text (ST) and its translation are related with each other by sharing the same meaning (semantics), which is realised in each language using different resources (lexicogrammar).

Steiner (2001: 186) models translation on the level of semantics, arguing that this is 'the most plausible level at which to hold features constant'. Translation is constrained by the ST unlike multilingual text generation (*ibid.* 185) and, thus, features of the ST must be preserved. If the features of the ST to be preserved are placed on the level of semantics, a translation will avoid 'typologically-caused weaknesses'. While if it is on the level of lexicogrammar, the TT will likely be marked or unnatural, in particular with pairs of typologically different languages. Halliday (2001: 17) affirms that if equivalence is sought at the semantics stratum, grammar may be allowed to vary.

Modelling translation on the semantics level has several important implications. Firstly, as Steiner (2001: 186) emphasises, semantics in SFL is a multifunctional one, so the investigation of translation covers ideational, interpersonal and textual meanings, instead of focusing on experiential meaning. Multifunctional meaning will be discussed further in Chapter 2 Section 1.3. The second point, which is implied rather than explicitly stated in Steiner, is that a semantics-oriented translation model indicates that the target of investigation is the text rather than sentences. A further discussion on this point is in Chapter 2 Section 3.1.

Returning to the stratified model of language, in addition to the realisational relationship between the semantics and lexicogrammar, another significant aspect is that the semantics of language interface with the context in which language is used. In SFL, meanings are construed by language on the one hand, and on the other it derives from function in context. In other words, language is socially situated; i.e. embedded in context within a semiotic system in the relationship between language and the social context.

We use language to make sense of our experience, and to carry out our interactions with other people. This means that the grammar has to interface with what goes on outside language: with the happenings and conditions of the world, and with the social processes we engage in.

(Halliday & Mattiessen 2004: 24)

Considering language in context means that the circumstances of a particular communicative event, such as what is being communicated (subject matter), what the social relationship of the participants is (formal or informal) and in what mode the communication is carried out (written or spoken), are taken into account in analysing language.

The importance of context in relation to translation is well illustrated by Malinowski regarding his difficulties in translating into English the language of

the Trobriand Islands where he conducted research (cited in Halliday & Hasan 1989: 5-8). He had to provide the context of situation in various ways in order for his translation to be intelligible for the English-speaking reader. He recognised the importance of both contexts of situation and culture.

Although the social and cultural distance between Japan and English speaking countries in the present day is most likely smaller compared to Malinowski's case, some shifts in news translation between Japanese and English are predicted to occur, and indeed they do.

For example, in a news article reporting the government's plan of reforming English education in Japan, the ST (Ja-orig-Y30) informs four points. They are 1) English at primary school will start from the third year instead of the current fifth, 2) English will be made a subject, i.e. will have assessment, for fifth and sixth year pupils, 3) English will be taught in English at middle school, and 4) high school students will learn higher level English by means of discussions or presentations.

In the TT (En-trans-Y30) not only are these points covered, but elaborated explanations are also provided. They include 1) English-only classes are already introduced at high school, 2) primary school children are expected to acquire basic reading and writing skills and will also be introduced to some content of middle school English classes, 3) therefore, they should be ready to be taught in English at middle school, 4) means of supporting instructors to improve their English and instruction skills, and 5) the level of English middle school pupils are expected to reach will be raised to *Eiken* (英検) Grade Pre-2 from the current Grade 3. *Eiken* is an English proficiency test widely recognised in Japan, but not outside, as an indication of the holder's level of English. Apparently, it is presented without any explanation in the TT.

The reader of the original ST is likely to be familiar with most of the extra information added to the TT since the reform of English education is an ongoing issue which is reported regularly. In contrast, it may not be common knowledge for the reader of the TT, given that not all the articles published in Japanese are translated into English. Therefore, the translator must have considered it necessary or useful to add information so that they understand the news better.

What the reader of the TT may lack is the broader background in order to comprehend the text: its 'context of culture' (Halliday & Hasan 1976: 47). As Halliday claims 'linguistic structure is the realization of social structure' (1978:

186), any linguistic structure appearing in the ST represents the society or the culture of the SL.

The power of SFL is that the language use is linked to all the linguistic strata and context outside language; from lexicogrammar to semantics, and to context. This means that when we understand the context in which language is used, we can assume to a certain extent what language we are likely to encounter in that context. At the same time, by a specific language use, we can predict the situation in which it is used.

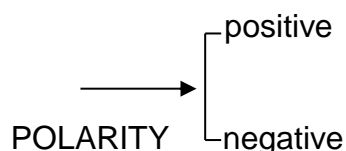
Since two languages are involved in translation, I recognise the significance of the context in the two relationships of text and context; the ST and its context and the TT and its context. I should also note in discussing translation that these two relationships are not independent from each other, but that the translation is constrained by the ST (Steiner 2001: 185).

2.1.2 Meaning making by choice

In this section three concepts that are important in SFL in discussing how meanings are created are introduced and related to translation. They are system, instance and register, and they can be explained through the notion of choice.

Choice in SFL means two things; first, the system of options and second, choosing among the options of a system. The former is what Halliday means when he states that language is a 'system of meaning potential' (Halliday 1978: 39). System is an essential category in SFL as it aims to explain the grammar of a language by making use of system networks. They represent an overall potential of a language, i.e. what options are available from which to choose. System networks are "sets of interrelated choices having the form 'if *a*, then either *b* or *c*'" (*ibid.* 128). For example, the system of POLARITY in English is presented as below in Halliday & Matthiessen (2004: 149).

Figure 2-2: System of POLARITY in English



The system of a language is 'instantiated' in the form of text as an instance of language use. This is the choosing among the options in a system, which indicates the process of instantiation of the system that results in a particular

structure or text, that is, what is actually chosen. Thus, system and text are related through instantiation on the cline of instantiation (*ibid.* 26-27).

On the cline of instantiation between the system and instance, there are intermediate patterns, which are called registers. They are functional varieties of language (Halliday & Hasan 1989: 38), 'patterns of instantiation of the overall system associated with a given type of context' (*ibid.* 27). In terms of choice, a register indicates what is typically chosen in a particular context of situation.

In addition to register, there is another notion that is important to this study, the notion of genre. These terms are at times used interchangeably (Sharoff 2017, 2018) but this study distinguishes them, as will be discussed in Chapter 3.

Situated between the system and instance, registers are seen in different lights depending on from which pole of the cline of instantiation they are looked at. From the system's pole, they are sub-systems under particular contexts of situation. Seen from the instance pole, they are groups of texts that display certain similar linguistic features.

Halliday (2001) points out that a translator tends to see a register from the instance pole, i.e. as a text type rather than a sub-system. He assumes it is because the translator needs to engage with language as text/instance in order to improve translation. However, by characterising a 'good' translation as a text 'in respect of those linguistic features which are most valued in the given translation context' (*ibid.* 17), he indicates the importance of viewing registers as sub-systems, from the pole of the system. Relating it to translator training, students are often unable to see beyond a particular text instance and not easily convinced by a correction or suggestion. The concept of register offers principles to judge suitability of language use in a particular context of situation.

Teich (2001: 198) explicitly asserts the importance of the notion of register as sub-system. She emphasises, in addition to contrastive knowledge about the systems, that it is indispensable for translators to have contrastive knowledge about the systems in use, i.e. instantiations of language systems according to certain situational and text-type specific conditions. Choices in translation cannot always be explained only by the differences in the systems. Moreover, even when there is a systemically equivalent option in the two languages, it is often not chosen in translation, due to culture specific conventions of text types (*ibid.* 198).

Mattiessen characterises translation as ‘the recreation of meaning in context through choice’ (2014: 272), referring to Steiner, who claims the register/context remains relatively constant across STs and TTs in translation (2004: 166), which is where translation differs from multilingual text production.

In summary, the notion of register on the cline of instantiation with system and instance on both poles is an important reference point to examine translation shifts. With the register notion, the translator’s choices can be readily analysed or justified against the features of a particular register. If a type of choice occurs frequently in translation, it can be said to be as a result of a strategy employed with a motivation behind it. By looking at news translations in significant quantities and observing them against the register in the SL and TL I aim to discover the current state of translation shifts in news translation between English and Japanese.

2.1.3 Meaning making and metafunctions

We have seen in Chapter 2 Sections 1.1 and 1.2 that language is embedded in context. We have also seen that within the stratified model of language semantics is an interface between context on the one hand and lexicogrammar on the other. In this section I observe how SFL explains the relationship between the context and language through the notion of metafunction, which organises different types of meaning.

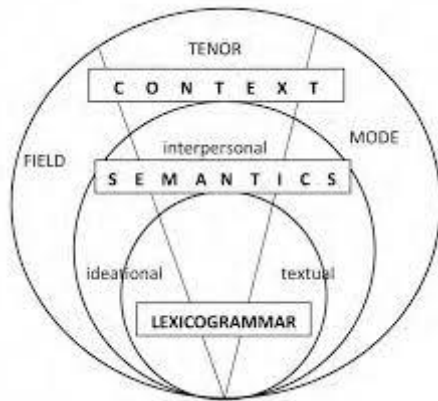
Language is said to fulfil three major functions. The metafunctions are explained as functional components of language at the level of semantics. They suggest that there are three types of meaning-making that are inherent in all texts.

1. Ideational metafunction: expressing the speaker’s experience of the external world, and of his own internal world, that of his own consciousness. This can be divided into experiential and logical functions.
2. Interpersonal metafunction: expressing relations among participants in the situation, including the speaker’s own intrusion into it.
3. Textual metafunction: expressing the language’s operational relevance to the total communication process.

(Halliday 1978: 45-46)

These components of meanings, i.e. metafunctions, systematically relate to the components of the social context (*ibid.* 189).

Figure 2-3: SFL's stratified model of language (from Bartlett & O'Grady 2018)



This diagram shows the relationship between the three modes of meaning in the context plane and the three metafunctions at the semantics level. Semantics, in Halliday's words, has 'an internal organization in which the social functions of language are clearly reflected' (Halliday 1978: 187).

- 1' Field of social process (what is going on): related to ideational metafunction
- 2' Tenor of social relationships (who are those taking part): related to interpersonal metafunction
- 3' Mode of discourse (how meanings are exchanged): related to textual metafunction

These contextual meanings are construed as linguistic meanings by the set of strategies at the semantics level (Matthiessen 1990: 324), as described as 1-3 earlier, and thus moving into the linguistic system (*ibid.* 324). Each metafunction at the semantic level is realised by different structural configurations at the level of lexicogrammar. Major ones are;

- 1'' Transitivity: related to experiential metafunction
- 2'' Mood: related to interpersonal metafunction
- 3'' Theme: related to textual metafunction

The present study focuses on the textual metafunction because that is the metafunction most closely related to the text's readability and acceptability. However, because of the nature of its 'text-forming' function, which will be discussed further in Chapter 2 Section 3, both experiential and interpersonal metafunctions are relevant to the textual. Therefore, I will briefly describe each metafunction as well as the textual.

The system of TRANSITIVITY provides a model for construing a domain of experience, making use of elements such as agent, process and goal to configure a structure. It is part of the experiential metafunction which concerns construing human experience of the world.

In order to explain additions, deletions and moves that occur in translation (Chapter 6 Section 1), it is necessary to understand how the experiential meaning is realised in English and in Japanese. There are three main types of process in the system of TRANSITIVITY, namely material, mental and relational. Material Processes refer to actions and events of the external world, Mental Processes refer to inner experience of the external world, and Relational Processes are those which identify and classify relationships.

A finding of the contrastive study of English and Japanese is that Japanese is a 'to become (*naru*)' language, while English is a 'to do (*suru*)' language (Ikegami 1981), which can potentially affect the way a process is translated. Consequently, it may influence what is realised as Theme because Theme realisation is related to word order.

Structural differences such as the above are one of the obvious causes for the surface differences between the ST and the TT with respect to translating news. However, they do not constitute the only reasons. Other issues, such as differences of the level of understanding/familiarity of the content between the ST and the TT readers, or relevance of the news to the source language (SL) and TL, are also likely to be reasons for adjustments in translation, namely additions or deletions of information (e.g. Blum-Kulka 1986; House 1997; Olohan & Baker 2000). For example, Baker (1992: 34) suggests a strategy of using a loan word plus explanation to translate culture-specific items. In terms of the experiential metafunction, such adjustments, i.e. shifts are anticipated to take place with notions or issues that are culturally and socially bound. For example, an explanation may be added which does not exist in the ST, as is seen in the Lafcadio Hearn article in Chapter 1.

These shifts are the target of investigation here, and not the formal changes that necessarily take place due to the syntactic difference between English and Japanese. In other word, the translation shifts of additions or deletions with which this study is concerned are the adjustments the translator can choose to apply instead of the obligatory ones that the syntax forces him/her to (Jakobson 1959).

Therefore, attention will be given mainly to the information in the text that can be considered extra, i.e. not that required by the grammar as constituents. In

addition to additions and deletions, moves will also be analysed. Within experiential meaning, which relates to the field of discourse, types and rank of additions, as well as deletions and moves, will be analysed. Methods of identifying them will be discussed in Chapter 4.

The second metafunction is the interpersonal meaning, which is principally realised through the system of MOOD. The components are Subject and Finite (underlined) and in English these appear at the beginning of the clause.

2-1 Should a captive chimpanzee have the same rights as a 'legal person'?

En-orig-C28

In comparison, interpersonal significance in Japanese is placed at the end of the clause, and Mood types are indicated by the Predicator and the Negotiator that follows (Teruya 2006: 135) as in example 2-2.

2-2

どうぞ	キャッシュ家のために	祈ってください。
<i>Dōzo</i>	<i>Kyasshu-ke NO TAME NI</i>	<i>inottekudasai.</i>
Please	Cash family for	pray-ASP-IMP
Please pray for Cash family.		
Modal Adjunct	Circumstance	Process

Ja-trans-C52

Example 2-2 shows the interpersonal significance at the end of clause '祈ってください [pray (please)]'. The example also includes an additional interpersonal element 'どうぞ [please]' at the beginning of the clause, which is realised as an interpersonal Theme. What can be seen here is that the interpersonal significance in Japanese is typically realised in the clause's final position, but initial position is not irrelevant either.

Interpersonal meanings fundamentally realise speech functions such as offer, command, statement and question. Given that the main function of news articles is to report newsworthy events, the relevance of the interpersonal meaning to news translation is rather limited, although it does not mean that these meanings are not realised in news (e.g. White 1998). In the present study, interpersonal metafunction will be investigated in relation to the textual metafunction, in particular Theme.

The third metafunction is the textual, a main focus of this study in the investigation of readability and acceptability of translations. This is the 'form

of organisation, whereby it fits in with, and contributes to, the flow of discourse' (Halliday & Matthiessen 2004: 64). This aspect of meaning is carried out by resources such as reference, information structure and thematic structure.

Theme is 'the element which serves as the point of departure of the message; it is that which locates and orients the clause within its context' (*ibid.* 2004: 64). Hence, how the text is organised can be perceived by looking at the Themes. Each Theme contextualises its message in three possible dimensions; discourse, tenor and the field of activity, by the textual, interpersonal and experiential Themes respectively (Halliday & Matthiessen 2004; Rose 2001: 111).

Theme is realised positionally in English. The most natural, or unmarked, Theme realisation in declarative clauses in English is the Subject, an experiential element which is obligatory in Theme. An experiential Theme on its own forms a simple Theme.

Textual Themes such as 'and' or 'but', and other interpersonal Themes such as 'probably' or 'honestly' can appear as part of a multiple Theme. (Halliday & Matthiessen 2004: 71-87). There are also marked experiential Themes, in contrast to the neutral unmarked Theme. Marked Themes occur only when 'contextual reasons overrule the unmarked choice of Subject as Theme' (Thompson 2004: 145). Examples of marked Theme are in 2-3 and 2-4 (underlined).

2-3 According to the statement, a 2012 annual report from the National Park Service counted 3.7 million visitors to Liberty Island in 2011 ...

En-orig-C10

2-4 Hours later, authorities pronounced Sandidge dead. En-orig-C14

Themes can be realised differently in different languages (Halliday & Matthiessen 2004; Rose 2001) and as an example Halliday & Matthiessen offer the particle *wa* in Japanese, calling it the Theme marker (2004: 64). However, research conducted on Japanese suggest that Theme realisation in Japanese does not exclusively rely on the particle *wa* and clause-initial position plays a more prominent role (Thomson 1998, 2001, 2005a; Tsukada 2001; Teruya 2006). Theme in Japanese is discussed in Chapter 2 Section 3.2.

These differences in how discourse operates in English and Japanese have the implication that some form of shifts is inevitable in translation, as simply following the textual structure of the ST will, for example, misplace markedness/unmarkedness, which can result in delivering different nuances

(Naganuma 2000). Specific issues in relation to Theme realisation and identification will be discussed in more detail in Chapter 4.

So far I have discussed major concepts of SFL which are important to the current study. To reiterate, the current study will benefit from an SFL approach, which situates language in context. The realisational relationship between context and language enables us to investigate a characteristic or typical language use in a given context.

With regard to translation, the SFL model is particularly powerful in that the semantic stratum relates to the lexicogrammar of the two languages involved; meanings are provided for the translator in the form of ST in one language which s/he recreates in another language using its lexicogrammar.

Both theoretically and in practice it is possible that several forms are available to realise one meaning (Halliday 1978: 44) and the translator's choice of one form over others is motivated. SFL's approach of seeing language as systemic choice supports the analysis of the translation shifts taking into account the context of translation; what is being talked about, what the social relationships between the speaker/writer and the audience are, and how the communication is organised. It aims to 'explain the nature of language in such a way as to relate it to its external environment' (*ibid.* 48).

Moreover, SFL's metafunctional approach can be applied to the analysis of any language instances, which means that the functional framework is independent of a particular language. Therefore, SFL provides one common set of analytical tools for analysing texts in different languages, although grammar varies.

2.2 SFL and translation studies

From the beginning of SFL studies, as we have seen before, SFL has had a link to translation studies (Catford 1965). SFL's description of language has been developing mainly via its analysis of English, but the framework is applied to many other languages including French, Spanish, Chinese and Korean, to name a few. Descriptions of grammar of Japanese have also been developing (Thomson 1998, 2001, 2005a; Tsukada 2001; Tatsuki 2004; Teruya 2004, 2006, 2015). SFL's applicability to multilingual analysis has already been established, as Halliday declares that the 'general linguistic theory should be a theory that can be applied to any language' (1961/2002).

There is also a growing number of studies in translation studies using SFL. Earlier work includes Baker (1992), House (1997), and Hatim & Mason (1990, 1996), which were influential for translation studies in the 1990s, according to Munday (2008).

Indeed, research on translation using the framework of SFL exists with many languages paired with English. Kim & Matthiessen (2015: 340) list selected examples of research on the translation of Theme with the language combination of English and other languages such as German, Norwegian, Spanish, Portuguese, Chinese and Korean. In addition to these, research involving Italian (Manfredi 2008), German (Steiner 2001; Teich 2003), Arabic, Hungarian and Persian are found, though their focus is not limited to translation of Theme unlike those listed in Kim & Matthiessen (2015). With regard to Japanese, and in particular with the translation of Theme, Naganuma (2000, 2008) continues to contribute to the field.

Among the wealth of previous research on translation, Kim (2007, 2011) and Teich (2001, 2003) are most relevant to the current study. The current study also draws on Teruya (2006) and Thomson (2001) with respect to SFL of Japanese, which are referred to later in this chapter as well as in Chapter 4.

Research by Kim (2007, 2011) focuses on translator training, and centres on analysing and explaining errors in translation. Based on the observation that accurate lexical choices do not necessarily ensure the TT's readability, she investigates textual meanings (flow of information) in English and Korean translations by analysing Theme choices in the STs and their translations.

Her research is relevant to the present study in a number of ways. Firstly, her work is pedagogically motivated. Secondly, she focuses on the textual meaning, in particular Theme. Thirdly, she works with the language pair of English and Korean, the latter is typologically similar to Japanese.

Korean is characterised as a 'topic prominent' language (Li & Thomson 1976; Kim *et al.* 2014), which has an SOV word order, makes use of a case marking system, and often uses ellipsis to track identifiable referents. These features of Korean are shared by Japanese and, therefore, she faces similar methodological issues that are relevant to the present study.

The present study also differs from Kim's in a number of aspects. One of the main foci of her research is the error analysis of trainee translators and, hence, she analyses translations produced by trainees. In contrast, my focus is to find translation shifts that occur in professional translation to provide trainee translators with evidence to support their practice.

Next, research by Teich (2001, 2003) is of particular importance to the current study. She presents a model for contrastive linguistic investigation to discuss specific properties of translated texts. Within the model, she firmly situates the notion of register which, according to her, defines the notion of comparability (2003: 222) for texts written in different languages.

...while an SL and a TL text are comparable because they stand in translation relation, another basis of comparison must be found for multilingual original texts. Such a basis can be suitably provided by the notion of register: if two texts in a language *I1* and language *I2* are used in similar situational contexts in which they fulfil comparable functions, then these two texts can be compared across the two languages.

(Teich 2003: 4)

The concept of register is also used by Teich to make 'more precise the hypotheses of SL shining through and TL normalization' (*ibid.* 222). Situating register in a central position of the analytical model enables the researcher an analysis of lexicogrammatical features of translation with reference to the contexts of situation in both TL's and SL's.

The current study is similar to Teich's in that I aim to discover a pattern of shifts in translation using parallel and comparable corpus. At the same time, it is different from hers in that I focus on the textual metafunction in the register of news, while the practical aspect in Teich (2003) is to test the methodology and analyse language in the field of science fiction in terms of different parameters covering three metafunctions.

Overall, SFL has been widely and successfully applied in a variety of language-related research, which in itself is an evidence of its applicability and analytical power. Nevertheless, there has also been criticism.

Firstly, there is a danger of imposing analytical tools or systems that are based on English on to other languages (Baker 1992; Munday 1998a, 2000). This is a valid point, in particular for languages that have been analysed little using the SFL's framework. With regard to the present study, however, recent progress in the SFL of Japanese provides us with useful resources.

Secondly, SFL is sometimes criticised for its theoretical complexity. Halliday himself acknowledges it but argues '(i)t is complicated because language is complicated, and there is no point in pretending that it is simple' (Halliday 2009: 61). He also states that 'the aim has been to construct a grammar for purposes of text analysis: one that would make it possible to say sensible and

useful things about any text, spoken or written, in modern English' (Halliday 1985: xv).

SFL is indeed rather a complex model. Kim explains that the nature of complexity results from its diversity as 'it engages in meaning at various ranks including phonology, lexicogrammar, meaning and context with authentic texts used in real life' (Kim 2016: 8). However, she insists that 'there are ways of applying SFL in a manageable and focused way' (*ibid.* 8), and these are exemplified in a number of studies on Theme cited in Kim & Matthiessen (2015).

2.3 Text and textuality

2.3.1 The nature of text

SFL is primarily concerned with analysis of texts as the unit of meaning. In order to investigate readability and acceptability of translated news articles, the current study takes the whole text as the unit of analysis and comparison. A text is not merely a collection of sentences but has characteristics as a unit.

Halliday presents a passage below to demonstrate that some form of structure is needed for grammatically correct sentences to make appropriate text, which thus ends up as 'non-text'.

Now comes the President here. It's the window he's stepping through to wave to the crowd. On his victory his opponent congratulates him. What they are shaking now is hands. A speech is going to be made by him. 'Gentlemen and ladies...

(Halliday 1978: 134)

Citing this, Baker (1992: 124) claims that we do encounter translation of this type of 'non-text' despite Halliday's argument that we normally do not.

What makes the above passage 'non-text', or why it is not considered as a text, can be explained from the way a text is defined in SFL. A Text is 'a unified whole' that is coherent and has structure (Halliday & Hasan 1976). The unity of a text is called 'texture' that is characterised by having 'a generic structure', being 'internally cohesive' and 'function(ing) as a whole as the relevant environment for the operation of the theme and information systems' (Halliday 1978: 136). Here, Halliday emphasises the importance of the context again, claiming that a text is 'the product of its environment, and it functions in that environment' (*ibid.* 136). Thus, a text is language 'that is functional... that is

doing some job in some context' and not isolated words or sentences out of context (Halliday & Hasan 1989: 10).

As we have seen earlier, translated news articles are read in an environment which is different from that in which the original news articles are written. In order for the translated news to be readable and acceptable in a new environment, i.e. to make sense to the new audience, adjustments are often required during translation (Bielsa & Bassnett 2009).

Among the components of the texture above, the textual system of THEME is one of the main foci of this study in investigating readability and acceptability of translated news. The problem of the 'non-text' above results from the way information is organised in each sentence and across text, which textual metafunction, in particular the system of THEME can usefully help explain. This will be verified in the following section.

2.3.2 Textuality and Theme

One of the resources of the textual metafunction, the system of Theme, is a grammatical concept which organises information in the clause. However, its contribution is not limited to the clause but it also functions above the clause with regards to texture creation in text. Halliday explains in the first edition of *An Introduction to Functional Grammar (IFG thereafter)* that the Theme is 'the starting point for the message; it is what the clause is going to be about' (Halliday 1985: 39). This explanation contains two separate notions about Theme and this 'double-sided definition' (Downing 1991: 122) has attracted researchers to investigate what exactly the functions of Theme are.

Subsequently, scholars find that the Theme is not necessarily what the clause is about (Downing 1991) but is 'more than what the message is about' (Martin 1995: 254). The function of the Theme is explained using a variety of expressions such as an 'orienter' (Fries 1994: 244), 'logogenetic growth-point' (Matthiessen 1995b: 27), 'an orientation', 'a perspective' and 'a point of view' (Martin 1992: 489).

By 'orienter' Fries means the Theme 'orients the listener/reader to the message that is about to be perceived and provides a framework for the interpretation of that message' (1995a: 318). By 'logogenetic growth-point' Matthiessen means the Theme is 'the point of integration of the clause as a message or logogenetic growth-point' (Matthiessen 1995b: 27). By introducing different expressions for Theme, Martin attempts to explain that information in a clause is distributed systematically as Theme or Rheme to contribute to the development of discourse. He states that 'significant patterns

of information flow through Theme in ways that are critical to an interpretation of the meaning of a text' (Martin 1995: 254).

These discourse-based studies demonstrate explicitly and clearly what 'the point of departure' means via their analyses of real texts of different register. Their findings are reflected in the definition of the Theme in the third edition of *IFG*, in which the aboutness of the Theme is notably backgrounded.

The Theme is the element which serves as the point of departure of the message; it is that which locates and orients the clause within its context.

(Halliday & Matthiessen 2004: 64)

This is a suitable definition of Theme for the current study where the focus is the flow of information, and aboutness is not considered as a defining concept for Theme. Given that the Theme is part of the textual meaning of which the function is to create discourse, adopting an 'external' view of the Theme in relation to the surrounding discourse (Thompson 2007: 678) is more suitable for investigating the textuality of translation than the 'internal' view of the Theme in relation to Rheme (*ibid.* 678). In addition, Themes can be multiple, involving textual and interpersonal Themes combined with an obligatory experiential Theme too, and, thus, the general function of Theme as the point of departure is appropriate.

However, aboutness is not entirely ignored since the concept relates to the notion of Topic, which is a major grammatical category in Japanese. According to Li & Thompson (1976) Japanese is a language that is both subject-prominent and topic-prominent, and it makes use of both the Subject-Predicate structure as well as the Topic-Comment. When the Theme conflates with the Topic (marked by *wa*), it is indeed what the clause is about.

At this point I will discuss how Theme is realised in Japanese. The discussion here is limited to an introduction of two slightly different views, based on which I will examine in Chapter 4 how this study identifies Theme in Japanese and how it deals with the Topic marker *wa*, which is often mentioned as the Theme marker.

There is general consensus amongst researchers working in SFL of Japanese that Japanese does not exclusively rely on the particle *wa* to identify Theme and that clause initial position plays a major role (Thomson 1998, 2001, 2005a; Tsukada 2001; Teruya 2006). Nevertheless, there are two different views on the relationship between Theme and the particle *wa*. The difference lies in the level of *wa*'s involvement one allows in the Theme structure.

On one hand there are researchers who claim *wa* plays an important role in the system of THEME and Teruya is representative thereof. He claims that 'Japanese deploys segmental marking as well to assign thematic status' (Teruya 2006: 71) to elements in configuring the textual metafunction, and also states that Theme in Japanese is 'realised positionally by being slotted in at the clause-initial position' (*ibid.* 51). He explains that *wa* gives the thematic prominence by highlighting an element in the clause as Theme.

However, only the nominal group can be segmentally marked this way and the other classes of group would have to be realised as Theme by an alternative way, i.e. by being placed in initial position. Since elements marked by *wa* are generally slotted in initial position, for Teruya, initial position and *wa* marking do not contradict each other in recognising Theme. The recognition criteria he offers for Theme at the clause rank are:

- a) Whatever precedes the element marked by *wa* or other theme markers is an unmarked Theme;
- b) If there is a marked participant or circumstance, the Theme always extends up to and includes that element marked by *wa* or other theme markers;
- c) If, however, the element described in (a) and (b) above is absent, the Theme likewise extends up to and includes the topical theme which may be marked by *ga*.

(Teruya 2006: 96)

Teruya's description includes any element marked by *wa* or by other Theme markers as Theme, irrespective of a preceding marked Participant or Circumstances (1998 cited in Thomson 2001). Hence, *wa* plays a significant role in the Thematic structure.

In contrast, Thomson has a view different from Teruya, and argues that the particle *wa* does not play a substantial role in Thematic organisation (1998, 2001, 2005a). She claims that the Theme in Japanese is realised in clause-initial position in the same way as in English (Thomson 1998). Elements marked by *wa* do appear in initial position, but not all those appearing there are marked by *wa*. Hence, she concludes that *wa* does not function as the Theme marker although it has a textual function because it is associated with information status (New/Given).

Teruya also acknowledges a link between *wa* and the information structure, describing that Theme is marked by *wa* if Given, and if New by *ga*, other things

being equal (2006: 53-54). When analysing texts, he distinguishes Given or New using the terms highlighted for the Theme marked by *wa* and non-highlighted for the Theme marked by *ga* or with no marking respectively.

The particle *wa* is the Topic marker and Topics in the Topic-Comment structure are inherently Given (Li & Thompson 1976). However, Themes do not always have to be Given, although in unmarked cases they are indeed Given (Halliday & Matthiessen 2004: 89).

The Given is information presented by the speaker as recoverable from the context or cotext, while the New is presented as not-recoverable (*ibid.* 91). They are elements in a structure of information unit and contribute to creating discourse. Within the information unit, the sequence of Given followed by New is unmarked, although it can be reversed if there is a reason for doing so. It is also possible that the Given information is omitted and the information unit consists of the obligatory New only (*ibid.* 89).

The system of INFORMATION is related to but different from the system of THEME. Both systems are structural resources for creating discourse. However, while the system of THEME construes the clause with Theme and Rheme, the system of INFORMATION construes the information unit with Given and News (*ibid.* 87-94). Although these systems function in separate units, in an unmarked case the information unit and the clause are co-extensive, and the Theme conflates with the Given while the Rheme with the New. With respect to how to interpret the relationship between the system of INFORMATION and the system of THEME, Fries (1983) offers terms for two different approaches; the 'combining approach', equating Theme with Given, or the 'separating approach', disassociating the two.

Going back to the discussion of Theme of Japanese, Teruya and Thomson share the opinion that the particle *wa* relates to information structure. However, they are different in that Thomson clearly separates Theme and New/Given, while Teruya considers that they are 'interrelated' (2006: 107). The practical difference between them in identifying Theme is apparent when there is more than one element marked by *wa* in initial position.

2-5¹

福井県警の調べでは	犯人は	身長 175 センチくらい
<i>Fukui-kenkei NO Shirabe DE WA</i>	<i>han'nin WA</i>	<i>shin'cho 175 sen'chi kurai</i>
Fukui prefectural police NO investigation DE WA	the suspect WA	height 175 cm about
According to the Fukui prefectural police, the suspect is roughly 175cm tall		
Circumstance: Angle	Token	Value
Theme	Rheme	

Thomson (1998: 233)

Thomson takes the element in initial position in 2-5, a Circumstance, as Theme, leaving the second which is also marked by *wa* in Rheme. Teruya will include both elements marked by *wa* in Theme. Thomson states that an element+*wa* is 'part of the textual metafunction', but 'does not operate in the Theme/Rheme structure' (1998: 235) and therefore, does not contribute to the thematic development of text. Teruya in contrast treats element+*wa* as profoundly integrated with Theme. I will discuss how the present study identifies Theme in Japanese in Chapter 4.

Having discussed the functions of Theme within clause and outlined the two views as to how Theme is identified in Japanese, I will move on to the role of Theme in text. Although Theme is a grammatical construct at the rank of clause, it 'contributes to the flow of discourse' (Matthiessen & Halliday 2004: 64). With respect to this, Fries introduces the notion of the 'method of development of the text', which is 'the way a text develops its ideas' (1995b: 9). He explains that the Themes of the clauses or clause complexes of the text will express meanings that relate to the method of development.

This will be a useful concept to adopt, if one is concerned with realisation of Theme in successive clauses. Although that is an interesting aspect of discourse, the present study is more interested in whether Theme functions as an orienter at specific locations in text, not necessarily sequentially. It is claimed that Theme signals the boundaries of sections in the text (Thompson 2004: 165). The current study is concerned with the relationship between the Theme and text structure in relation to the text's readability.

¹ Gloss is conducted following Teruya's methods generally with some modifications to simplify them to suit the purpose of this study.

According to Halliday, text structure is a 'configuration of semantic functions' (1981/2002: 235), thus regarded as consisting of text segments, each serving a particular communicative function to develop the text's meanings. Whether or not the text segments of news articles can be identified by Theme merits an investigation. It will also be beneficial to uncover patterns, if the text segments are indeed indicated by Theme.

Chapter 3

Translation shifts and news articles

This chapter provides the definition of ‘translation shift’ the present study employs and describes previous research conducted on them (3.1). It also explains news as a genre and discusses the conventional structure of news articles, which is organised by one text nucleus followed by a set of satellites. Finally, how a genre based analysis helps examine shifts in news translation in terms of discourse is presented (3.2)

3.1 Investigation of translation shifts

3.1.1 Translation shifts

The investigation of translation shifts started in the late 1950s. Vinay and Darbelnet (1958/2000), working with English and French, developed a system of translation procedures. They offer seven procedures, three of which are classified as ‘direct translations’ (borrowing, calque and literal translation), and the remaining four of which are referred to as ‘oblique translations’. These oblique procedures result in differences between the ST and the TT, i.e. shifts, although the authors themselves do not use this term.

The oblique procedures are ‘transposition’ (change in word class), ‘modulation’ (change in semantics), ‘equivalence’ (completely different translation) and ‘adaptation’ (change of situations). Vinay and Darbelnet write that oblique procedures should be used when a direct one would lead to a unsuitable or awkward translation, which, Cyrus observes, has ‘a slight prescriptive undertone’ (2006: 1240). Thus, Vinay and Darbelnet view translation shifts as undesirable but inevitable.

The term ‘translation shift’ was first introduced by Catford, which is defined as ‘departures from formal correspondence in the process of going from the SL to the TL’ (1965: 73). ‘Formal correspondence’ is a type of equivalence, which is explained as;

any TL category (unit, class, element of structure, etc.) which can be said to occupy, as nearly as possible, the “same” place in the “economy” of the TL as the given SL category occupies in the SL ...

(ibid. 27)

There is another type of equivalence called 'textual equivalence' which is described as;

any TL text or portion of text which is observed on a particular occasion
... to be the equivalent of a given SL text or portion of text.

(*ibid.* 27)

According to Munday, '(t)extual equivalence is tied to a particular ST-TT pair, while formal equivalence is a more general system-based concept between a pair of languages' (2008: 60). This means shifts occur if textual equivalents are not formally correspondent with their source.

According to Catford (1965: 73), there are two major types of shifts: 'level shifts' and 'category shifts'. Level shifts occur between grammar and lexis, e.g. a Russian perfective aspect *сделат* is translated using the English lexical verb 'to achieve' (*ibid.* 75). Category shifts are divided into 'structure shifts' (changes of clause structure), 'class shifts' (changes of word class), 'unit shifts' (changes of rank), 'intra-system shifts', e.g. changes of term in systems such as number even though the languages have the same number system) (*ibid.* 75-82).

Catford's definition of shift is 'strictly linguistic' (Munday 1998: 544) and problematic as it 'relies heavily on the structuralist notion of system and thus presupposes that it is ... possible to determine and compare the *valeures* of any two given linguistic items' (Cyrus 2006: 1240). His approach to shifts seems to presume a certain level of symmetry between the languages, which makes it difficult to apply to language pairs such as English and Japanese that are structurally distant. However, the concept of shift has become wider than the original definition of Catford. In *Routledge Encyclopaedia of Translation Studies* published in 2009, Bakker, Koster & van Leuven-Zwart describe that:

- shifts commonly refers to changes which occur or may occur in the process of translating;
- shifts of translation can be distinguished from the systemic differences which exist between source and target languages and cultures;
- shifts ... result from attempts to deal with systemic differences.

(Bakker, Koster & van Leuven-Zwart 2009: 269)

It has been more than 50 years since Catford first introduced the term 'shift'. The definition he employs for it reflects his equivalence-based description of translation, i.e. the *replacement* of textual material in one language (Source

Language, SL) by *equivalent material* in another language (Target Language, TL) (Catford 1965: 20).

The notion of translation equivalence has been ‘for many years considered essential in any definition of translation’ (Snell-Hornby 1988: 15), as is seen in Catford’s definition. Nowadays however, there are few contemporary theorists ‘who subscribe to the “same meaning” approach to translation’ and when they ‘posit relationships of equivalence between texts, they do so without appeal to language-neutral, objective meanings and they acknowledge the role of translating subjects, translators, in creating and maintaining these relationships’ (Kenny 2009: 96).

For example, Pym (2007) asserts ‘one directional’ equivalence; target side equivalence, as opposed to assuming pre-existing ‘natural’ equivalence. Similarly, with a particular relevance to news translation, Bielsa & Bassnett discuss the skopos theory which Reiss and Vermeer developed in the 1980s, which ‘postulated that the objective of the target text would determine how it was translated’ (2009: 8). They continue to say;

This meant that a translation could deviate enormously from the source and yet fulfil the original purpose.

Equivalent effect (underpinning skopos theory) is probably the best one can hope for with most translations, when we reflect that what happens in translation is that a text is read, decoded and then reshaped in the target language to accommodate differences of structure, style, context and audience expectation.

(Bielsa & Bassnett 2009: 8)

Target-oriented translation means taking into account differences such as structure, style, context, and audience expectation between the languages involved in translation. It signifies an inevitability of shifts and allows them to be seen neutrally rather than negatively. The present study views shifts in this manner and tries to describe the state of them in professional news translation between English and Japanese.

In addition to viewing shifts neutrally, target-oriented translation means accepting ‘approved behaviour’, i.e. norms, within the target community (Toury 1995: 24). According to Toury, translations are ‘facts of the culture which hosts them’. He believes translation is a norm-governed behaviour, and norms are pervasive in translation practice (*ibid.* 24). Thus, target-oriented translation means that translations are expected to move closer to TL norms, as a result of shifts.

Hence, shifts are the changes that are observed between the ST and the TT, which are the result of the choice taken by the translator. Claiming that translators require 'decision-making skills', Pym (2016) offers seven 'translation solutions' as options available to them for addressing translation problems. His solution types are pedagogically motivated, and depending on the training needs, they can be grouped to three simple categories; Copying, Expression change and Material change. Or, for more focused work, the seven types can be extended to unlimited number of specific sub-types.

Pym states seven categories of translation solutions are of the most versatile level for teaching purpose as they are neither too abstract nor too confusing (*ibid.* 221). The names for the solution types are transparent for clarity; Copying Words, Copying Structure, Perspective Change, Density Change, Compensation, Cultural Correspondence, and Text Tailoring. Solutions likely to be relevant to the current study are Perspective Change, Density Change and Text Tailoring. This is because employing these will result in shifts which are the target of investigation here; Theme realisation and additions, deletions and moves of information content.

3.1.2 Research on translation shifts

Van Leuven-Zwart (1989) is an early example of approaching translation shifts neutrally, without describing what translators should do (Cyrus 2009). She observes and describes 'the translator's interpretation of the original text and the strategy adopted during the process of translation' (1989: 154). To describe 'shifts in integral translations of narrative texts' (*ibid.* 151) she develops a two-part method consisting of a detailed analysis of shifts on the 'microstructural level', i.e. within sentences, clauses and phrases, and an investigation of the effects of them on the 'macrostructural level', i.e. 'the characters, events, time, place and other meaningful components of the text' (*ibid.* 155).

The microstructural shift analysis interests the current study. It consists of four steps: first, the units to be compared – 'transemes' are established. These are either a predicate and its arguments, or adverbials without a predicate. Then each pair of source and target transemes are compared to determine the 'architranseme', i.e. the 'common denominator', which is then compared with each separate transeme to establish the relationship between them; either synonymic or hyponimic. Finally, the pairs of transemes are classified according to this relationship; 'modulation' (hyponimic with respect to the architranseme), 'modification' (two transemes are contrasting with each other) and 'mutation' (architranseme cannot be established). Each of them has

several subcategories that total 37 and this is a reason why the model has sometimes been criticised for its complexity (e.g. Munday 1998a: 3).

Van Leuven-Zwart analyses shifts manually, but ten years later Munday (1998a) attempts a computer-assisted approach to the analysis of translation shifts. He uses basic corpus linguistics tools and lexicography 'to enable accurate and rapid access to surface features over a whole text, reducing the arduousness and tedium of what has previously been a manual task' (*ibid.* 2). In addition, he makes use of comparable corpora to ensure a systematic analysis 'to break out of the confines of a single pair of texts to enable preliminary consideration of the influence of typical target-language patterns and of the translator's specific idiolect in the creation of shifts' (*ibid.* 2-3).

Munday (1998a) recognises that not every shift can be thoroughly explained by simply adopting a computerised approach, for example, whether or not the Spanish word 'casa' is consistently translated into English cannot be observed quantitatively as it can mean both English 'home' and 'house'. Nonetheless, it facilitates the feasibility to go beyond the individual shift and 'see if it becomes a trend over the whole text' (*ibid.* 10). He also emphasises the importance of taking into account systemic differences between the languages involved in translation in interpreting the results.

Cyrus (2006) is another research which combines corpus linguistics and translation studies. She created a corpus resource in which translation shifts are explicitly annotated and categorised. Her annotation is based on predicate-argument structures and the shifts are categorised as grammar shifts, semantic shifts or structural shifts. Cyrus' categories for shifts are simpler than van Leuven-Zwart's, although there are some similarities. Cyrus claims her project distinguishes itself from other similar annotation work in that it deals 'not with one language, but with two (and potentially more) languages, and in particular with parallel data' (2006: 1244).

With respect to research on translation shifts involving Theme, Kim & Matthiessen (2015: 340) provide selective examples as already mentioned in Chapter 2. However, according to Alekseyenko (2013), corpus-based comparative research involving Theme is not extensive. Indeed, with regard to Japanese, little corpus-based research has been conducted on shifts. Kawahara (2009a) is a rare example of study of translation shifts but it is in fact on the subject of interpreting and it is not corpus-based. He investigates translation shifts in the act of interpreting between English and Japanese from the perspective of cognitive linguistic typology. His focus is on grammar and he aims to find out how often shifts occur in order for the naturalness of

languages to be maintained. Focusing on features such as 'reference' and 'predication', he compares English and Japanese to examine what makes natural English or Japanese.

In an attempt to uncover patterns of translation shifts, the current study explores Theme realisation in translation both in English and in Japanese. As is seen in Chapter 2, there have been studies carried out into translation shifts between English and various languages, but little with Japanese. Examining shifts in hard news with respect to the specific language pair of English and Japanese, in a systematic manner and using a computerised approach, I will address a gap that exists in terms of shifts involving Japanese in translation studies.

Another gap of knowledge the study hopes to address relates to the type of texts investigated, i.e. hard news articles. According to Kawahara, Non-literary translation has rarely been researched with respect to Japanese (2011b), and investigation into hard news is also new (Thomson, *et al.* 2008: 66). By means of an empirical, systematic analysis of translation shifts in hard news on a large scale, the study hopes to address the gap of knowledge as described above.

In this study, one of the foci is Theme, because it is a major resource for creating discourse. News articles are generally expected to read naturally and, thus, readability and acceptability of the translation are important notions in this study. In relation to readability and acceptability of translated news, the content of information is also relevant in addition to how the information is organised. For this reason, this study investigates additions, deletions and moves of information too.

Additions and deletions are also called explicitation and implicitation, which are widely researched in translation studies (e.g. Blum-Kulka 1986/2000, Olohan & Baker 2000, Károly 2012). Among the research on additions and deletions in particular relevance to news translation are Bielsa & Bassnett (2009: 115) and Aktan & Nohl (2010). The latter is insightful in a practical sense as their empirical analysis identifies various forms of moderation during 'trans-editing' news reports on BBC World Service to Turkish readers. They report that factors 'such as cultural discrepancies and local relevancies force the trans-editors to generate explicit and implicit changes, which shall reinforce the comprehension of the texts' (Aktan & Nohl 2010).

3.2 News translation

3.2.1 Genre and register

Register and genre are important notions in the present study as they offer different perspectives to investigate translation shifts. The two terms are often used ‘in various incompatible ways’ (Sharoff 2017, 2018). However, the current study distinguishes them and intends to use ‘register’ when referring to typical language use, and ‘genre’ when referring to conventional structures that texts in the same group or category share.

As discussed in Chapter 2, Halliday defines register as a functional variety of language (Halliday & Matthiessen 2004: 5). A register is a characteristic way of using language according to situation type. According to Biber & Conrad, the underlying assumption of the register perspective is that ‘core linguistic features like pronouns and verbs are functional, and, as a result, particular features are commonly used in association with the communicative purposes and situational context of texts’ (2009: 2).

Since the register perspective is concerned with typical and frequent language use in a particular context, Sharoff employs text-internal criteria to define register. They are based on the ‘parameters related to lexicogrammatical choices made within the text’ (Sharoff 2018: 68). The text-internal view is compatible with Halliday’s definition of register.

In contrast to registers, which are defined text-internally, genres are defined text-externally (Sharoff 2017, 2018). Text-external criteria are based on ‘parameters related to the context of its (the text’s) production’ such as ‘the author, communicative aims or perception by the audience’ (Sharoff 2017: 538). By applying text-external criteria, genres are defined through communicative aims without references to their lexicogrammatical realisations (Sharoff 2018: 68). Defining genres without referring to linguistic phenomena enables multilingual studies, including the present one, to compare texts written in different languages because they are related by the same communicative functions. Sharoff mentions the needs for genre analysis across languages too.

Once a reliable text-external framework has been established, we will be able to investigate its correlation with any appropriate text-internal features in a given language,

(Sharoff 2018: 68)

A genre perspective of linguistic analysis focuses on complete texts and the conventional internal structure or rhetorical organisation of texts from a variety (Biber & Conrad 2009). Texts that belong to the same variety conform to the culturally expected way of constructing texts belonging to that variety (*ibid.* 16). It is because they share similar communicative functions to fulfil and a similar context of situation in which language works.

However, where translation is concerned, even if the communicative functions and the context of situation remain more or less the same, translation and the original are likely to display differences. Translated text belongs to a culture that is different from the original. These differences are recognised generally in news translation as a form of strategy.

In the present study, taking the conventional structure as one of the important aspect of meaning making (Halliday 1977/2003), part of the investigation is Theme realisation at specific locations in the text. The aim of this is to observe whether or not Themes realised at these particular locations are recognised as a conventional device in constructing texts.

Within functional linguistics, where language in context is in focus, it is said that '(m)embership of some text in a genre allows firm predictions to be made concerning the linguistic details of that text and, conversely, the occurrence of particular combinations of linguistic features can be (...) a strong indicator of genre membership' (Bateman *et al.* 2007: 147). Thus, a genre-based approach to translation in general as well as translation shifts is meaningful. It firstly allows us to describe the linguistic features and the way discourse operates in a specific genre, in this study hard news reports, both in the SL and the TL, and then to reflect them in translation.

In summary, the register perspective describes the typical linguistic features of text varieties in relation to the particular context of situation of the variety. Analysing Theme from the viewpoint of register identifies the types of Theme that are frequently used in news articles (Chapter 5). In contrast to this, the genre perspective describes the conventional way to organise information within text. Analysing Theme from the viewpoint of genre identifies the types of Theme that are conventionally used at certain locations in the text (Chapter 5). In addition to Theme, the current study investigates additions, deletions and moves as an aspect of readability and acceptability (Chapter 6).

3.2.2 News as a genre

Journalistic genres are considered well-established genres that 'exhibit various patterns of similarity in terms of structure, style, content and intended

audience' (Swales 1990: 58). There are two major types of news reports; hard news and soft news (Tuchman 1973; Bell 1991; Iredema *et al* 1994b). Hard news concerns 'events potentially available to analysis or interpretation and consists of "factual presentations" of events deemed newsworthy' (Tuchman 1973: 113). Soft news are also called 'feature articles' and 'human interest stories' as well, and it concerns the lives of human beings, which are interesting (*ibid.* 114).

Bell acknowledges that the boundaries between hard and soft news are unclear (1991: 14). Tuchman also explains that it is possible for the same event to be treated as either a hard news or a soft news story (1973: 114). However, she offers a useful categorisation for distinguishing hard news from soft news in terms of scheduling. Reflecting the question of scheduling, hard news is either about unscheduled, unexpected events such as a fire or prescheduled event such as a debate on a legislative bill, and soft news is about non-scheduled events whose date of distribution as news is decided by the newsmen (*ibid.* 117-118). She notes however, that there are also non-scheduled hard news stories, which tend to involve investigative reporting (*ibid.* 118).

Hard news, which this study examines, must be timely, where soft news does not need to be.

Before moving on to the following section to discuss the issues of textual organisation of news, there are a few clarifications needed that relate to hard news as a genre. One is the distinction between the tabloids and the broadsheets. A news report in the *Guardian* is written differently from one in the *Sun* and many of the differences can be explained by reference to factors such as 'the purpose and nature of the newspaper and the nature of the readership' (Bhatia 1993: 20). In Japan, so-called sports newspapers may be considered in a way similar to British tabloids, and the national daily newspapers are in a way regarded similar to the British quality papers in terms of what Bhatia describes above. In order to keep the variables comparable, as well as the practical issue of Japanese sports newspapers not having English sites, tabloids and sports newspapers are excluded from this study.

Another clarification required here is the difference between the traditional print editions of newspapers and their web editions. Web editions are analysed in this study due to their accessibility. Though the presentation of news reports may be different on the print and the web editions due to technological differences (Knox 2009; Caple 2009), the textual structure of hard news is believed to be unaffected since the electronic media only came

into being after the primacy of the genre of hard news had been established (Iedema *et al.* 1994a: 9), although this is not incontrovertible.

Changes in translation occur for various reasons, such as the genre of the ST, topic and target readership. Different languages have their own rhetorical conventions in general as well as in specific genres. Exemplars of a genre 'exhibit various patterns of similarity in terms of structure, style, content and intended audience' (Swales 1990: 58).

Iedema *et al.* state that different text types that appear in the newspapers have different social functions and thus different patterns of textual organisation (1994b: 1). It must be also noted that different cultures have their own way of organising texts.

For White (1998: 243) news reports are 'grounded in communicative events such as speeches, interviews and press releases', which 'act primarily to represent, not activity sequences, but the points of view of various external sources'. They are classified as communicatively-based rather than event-based. A news report should strive to remain objective and use neutral language, while presenting a diversity of opinions, voices, and perspectives of the event, incident, or issue under discussion.

3.2.3 News structure and discourse

Different text types that appear in newspapers have different social functions and thus different patterns of textual organisation, and the hard news discourse in English is said to have a specific kind of generic structure (Iedema *et al.* 1994a; White 1997; Thomson *et al.* 2008). It is often described by the term 'inverted pyramid' (Bell 1991, 1998; Bhatia 1993; Pöttker 2003). In this structure the most important information in the first or 'lead' paragraph of the story is placed at the tip, and the remaining details are arranged in the following paragraphs in order of importance.

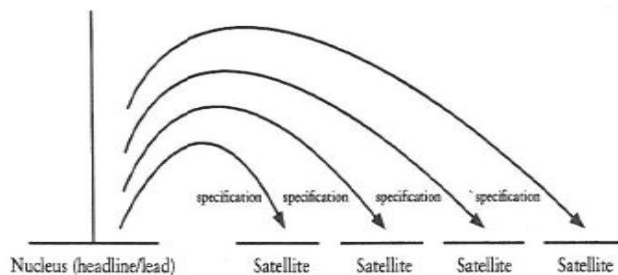
Iedema *et al.* (1994a) and White (1997) clarify the structure of the body part of the news report, which follows the headline and the lead. They explain that it acts to specify the meaning presented in the headline and the lead by elaborating, providing causes or consequences, contextualising or assessing attitudes (Iedema *et al.* 1994a; White 1997). White claims that the role of the body of the news report is 'not to develop new meanings nor to introduce new information but to refer back to the headline/lead through a series of specification' (1997: 112). The structure of the body part of the news report is described as a set of 'satellites' to the headline/lead, which is called the

'nucleus' to which satellites are dependant and whose meaning they specify (Iedema *et al.* 1994a; White 1997).

These terms are used differently in Rhetorical Structure Theory (RST), in which a nucleus and a satellite constitute a pair of successive spans in a text. The nucleus is more essential and central than the satellites and is always placed before the satellites. The satellites gain their significance through the nucleus, without which they are incomprehensible (Mann & Thompson 1987: 31-38).

The terms nucleus and satellite are used in this project in the way that Iedema *et al.* (1994a) and White (1997, 1998, 2000) use them. RST describes hierarchically how parts of text are related to each other and make up a coherent whole. While RST dynamically facilitates an illustration of the relationships between parts of text to the elemental level of clause, that is beyond the scope of the current project and thus, a deep level of RST analysis is not conducted.

Figure 3-1: Nucleus and Satellites model (Thomson *et al.* 2008: 219)



The relationship between individual satellites and the nucleus is important but those between each satellite are not, and these are not well linked with each other. Texts with this structure are non-linear and orbital (Iedema *et al.* 1994a; White 1997).

The function of the nucleus is 'to provide a selective synopsis and/or abstract of the event at issue, to single out points of social salience or significance' (White 1997). The nucleus is followed by satellites which act to specify the meaning presented in the headline and the lead (Iedema *et al.* 1994a; White 1997).

The role of the body of the news report is 'not to develop new meanings nor to introduce new information but to refer back to the headline/lead through a series of specification's (White 1997: 112). Five major modes or relationships of specification between the nucleus and satellites are defined by White (1997,

1998, 2000). These relationships may appear the same or similar to the ones offered in RST. White states that they are based on logico-semantic relationships between clauses or clause complexes, as proposed by Halliday. For example, between the nucleus and a satellite of Elaboration lies a relationship that can be expressed as 'that is' or 'for example'. Or a relationship of 'because' can be found between the nucleus and the satellite of Cause-&-Effect. White defines these relationships as follows.

- Elaboration: one sentence or a group of sentences provides more detailed description or exemplification of information presented in the headline/lead, or acts to restate it or describe the material in the headline/lead in different terms.
- Cause and Effect: one or more sentences describe the causes, the reasons for, the consequences or the purpose of the 'crisis point' presented in the headline/lead.
- Justification (issues reports): one or more sentences provide the evidence reasoning which supports the newsworthy claim presented in the headline/lead nucleus. This justification could be seen as a text internal cause-and-effect, in that it explains why a particular claim has been made and hence could be included within a single broader 'Cause-and-Effect' category.
- Contextualisation: one or more sentences place the events or statements of the headline/lead in a temporal, spatial or social context. The geographical setting will be described in some detail or the 'crisis point' will be located in the context of preceding, simultaneous or subsequent events. Prior events of a similar nature may be described for the purpose of comparison.
- Appraisal: elements of the headline/lead nucleus are appraised, typically by some expert external source, in terms of their significance, their emotional impact, or by reference to some system of value judgement.

(White 1997: 115)

An illustrative example is presented in 3-1 to show the structure of news and how the nucleus and satellites are realised in text. Satellite 1 is an example of specification that elaborates how 'tough' NYC Council will become on tobacco, satellites 2 and 4 describe the consequences of the 'approval of an anti-tobacco law', satellite 3 provides the reasoning to support the 'anti-tobacco law' and satellite 5 places the 'approval of an anti-tobacco law' into social and geographical contexts.

3-1

Nucleus	NYC Council gets tough on tobacco, approves raising purchase age to 21 New York (CNN) -- The New York City Council voted on Wednesday night to approve an anti-tobacco law that will raise the tobacco-purchasing age from 18 to 21.
Satellite 1 Elaboration	In addition to the "Tobacco 21" bill, which includes electronic cigarettes, the council also approved a second bill, "Sensible Tobacco Enforcement." It will prohibit discounts on tobacco products and increase enforcement on vendors who attempt to evade taxes.
Satellite 2 Cause & Effect	Mayor Michael Bloomberg has 30 days to sign the bills into law. Given his previous support, that is likely to happen soon.
Satellite 3 Justification	"By increasing the smoking age to 21, we will help prevent another generation from the ill health and shorter life expectancy that comes with smoking," Bloomberg said in a statement on Wednesday.
Satellite 4 Cause & Effect	"Tobacco 21" will take effect 180 days after it is enacted, according to the council's news release.
Satellite 5 Contextualisation	New York City has now become the largest city to have an age limit as high as 21. Needham, Massachusetts, raised the sale age to 21 in 2005, according to the New York City Department of Health.

En-orig-C24

The relationship between individual satellites and the nucleus is important but that between each satellite is not. Since satellites are not well connected with each other, the order thereof can be changed without breaking the coherence of the text (ledema *et al.* 1994a; White 1997). Since the reordering of discourse elements is one of the aspects of the investigations, this orbital structure is employed in this study. The Generic Structure Potential (GSP) that Hasan develops (Halliday & Hasan 1989) is also a potentially useful model for this study. However, she stresses the importance of the order of discourse elements that appear in text, although allowing some flexibility.

Since the generic or conventional structures are established within culture, different cultures may make use of different structures. For example, Katsuta

(2001), who compares *Asahi Shimbun* and *the Guardian* articles, focusing on the functions of the first paragraph, finds that the leads from *Asahi Shimbun* summarise the whole article while those from *the Guardian* indicate the main topic.

In addition, Thomson *et al.* (2008: 66) cites the term ‘three readings’ from the Japan Newspaper and Editors Association (1994). This practice offers the reader three options in digesting news reports. A level 1 reading involves reading the headlines and subheadlines. A level 2 reading involves reading the headlines and the lead. A level 3 reading covers the entire article (Thomson *et al.* 2008: 67). This implies that the headlines/subheadlines, the lead, and the body, contain the same information but with a different degree of detail.

According to Thomson *et al.* (2008: 66) research on hard news in Japanese is new. Quantitative analysis of the hard news discourse in Japanese as part of this project aims to shed light on the generic structure of them to contribute not only to translation studies but also media studies.

3.2.4 News translation and shifts

Bhatia claims that ‘a successful achievement of the specific communicative purpose that the genre in question serves, depends on the use of specific conventionalised knowledge of linguistic and discursal resources, depending upon the subject-specific, socio-cultural and psycholinguistic factors typically associated with the setting with which the genre is associated’ (1993: 9). When translating an English news articles into Japanese, or vice versa, the differences between the two languages and societies in terms of the conventions of news writing as well as the subject-specific, socio-cultural factors, must be taken into account in order for the translated text to make sense for the target readership. Bielsa & Bassnett argue;

In news translation, the dominant strategy is absolute domestication, as material is shaped in order to be consumed by the target audience, so has to be tailored to suit their needs and expectations.

(Bielsa & Bassnett 2009: 10)

They continue by stating that the news translator ‘does not owe respect and faithfulness to the source text’ (*ibid.* 65), unlike the literary translator, but s/he transforms the source text ‘in a context in which faithfulness is due more to the narrated events than to the source text’ (*ibid.* 73). They list the most frequent modifications to which the source text can be subjected in the process of translation as below:

- Change of title and lead: titles and leads (informative subtitles) are often substituted for new ones so as to better suit the needs of the target reader or the requirements of the target publication;
- Elimination of unnecessary information: information can become redundant either because it is already known by the target readers or because it becomes too detailed and specific for a reader who is geographically and culturally removed from the reality described;
- Addition of important background information: when the target readers change it becomes necessary to add background information that will not necessarily be known in the new context;
- Change in the order of paragraphs: the relevance of the information in a new context and the style of the publication might make it necessary to alter the order of paragraphs;
- Summarising information: this method is often used to fit the source text into the space available and to reduce lengthy paragraphs which are no longer fully relevant to the target readers.

(Bielsa & Bassnett 2009: 64)

Negishi (1999) and Sprung & Vourvoulias-Bush (2000) confirm from their own experiences of working for newspaper companies that these modifications take place in news translation. The Lafcadio Hearn article in 1.1 is also an example of the second modification example above.

Translation is not regarded as separate from other journalistic tasks of writing up and editing, although it is an important part of journalistic work, which is subject to the same requirements of genre and style that govern journalistic production in general (Bielsa & Bassnett 2009: 57). Kawahara (2009 cited in 2011a) demonstrates that this is also the case in the Japanese media for producing international news articles. He presents three processes; direct translation, complex process and direct process where journalists, not translators, are involved in translation. The direct translation process involves the journalists translating news or news articles distributed by news agencies or broadcasters such as Reuters or the BBC. The complex process is a practice in which local staff either translate or summarise news from the local media for correspondents from Japan, who then, with the help of local assistants, investigate the news and interview people, and after that write up a news report in Japanese to send to the Tokyo head office. The direct coverage is a process whereby correspondents from Japan interview people in the local language and write a report in Japanese to send to Tokyo.

Kawahara (2011a) clarifies that the act of translation in news writing by journalists is indeed a part of editing and it naturally includes the manipulation of impartial information. Bielsa adds;

A precondition for the successful transmission of news texts across geographical and linguistic boundaries is their thoroughgoing transformation so that, even if the news source is the same, the text can fully respond to the particularities of a new context in which the narrated events will impact in different ways. This explains why radical changes such as the full rewriting of titles and leads (first paragraphs), the restructuring the order of paragraphs, and the *addition and/or elimination of information are the norm rather than the exception* in news translation.

(Bielsa 2010: 164, my emphasis)

There is broad consensus in research to acknowledge shifts in news translation as widespread journalistic practice (Sprung & Vourvoulias-Bush 2000; Lee 2006; Bielsa & Bassnett 2009; Kawahara 2009 cited in 2011a; Bielsa 2010). However, not much is known about shifts in news translation, in particular with respect to the language pair of English and Japanese. Therefore, the main aim of the current study is to examine what kind of shifts take place between the two languages, on which subject very few studies has been conducted.

As mentioned above, news articles may be translated significantly differently from other text types. For example, the data used here include examples that show multiple STs are used to produce one TT, which is unlikely to happen such as in literary translation. Additionally, news articles may be 'trans-edit(ed)' (Aktan & Nohl 2010) instead of simply translated, which means it is journalists and not translators who conduct translation. Hence, translation in journalism may be different from other text types. Nevertheless, it can be investigated as a type of translation in the field of translation studies. The current study is a pioneering one because news translation in Japanese is an understudied area. It is also new because news translation in general is relatively new and different, e.g. from literary translation, with regards to faithfulness to the ST and other aspects.

Chapter 4

Research design and analytical approaches

This chapter sets out the research design, beginning with the research questions and continuing with explanations of how the analysis is carried out, in order to answer these questions. Several methodological issues relating to the consistency of the analysis across both English and Japanese texts, including the unit of analysis and the boundary between Theme and Rheme, are also discussed. Finally, systems used for analysis, the analytical guidelines and the annotation principles are presented.

4.1 Corpus and tool

The use of corpus is a well-established method to investigate linguistic features. Two of the most important advantages of using corpora are 'authenticity and frequency' (McEnery *et al.* 2006: 101). A sample of authentic texts in use as 'instance(s) of an underlying system' (Halliday & Mattiessen 2004: 26) enables researchers to discover patterns of linguistic features that frequently occur or rarely in a particular language or language variety. The empirical nature of the approach, i.e. corpus as evidence of language use is 'a more reliable guide to language use than native speaker intuition is' (Hunston 2002: 20).

Given that this study aims to investigate translation shifts in a specific genre of 'hard news' (Bell 1991: 14), a specialised corpus, which is 'domain or genre specific' (McEnery *et al.* 2006: 15), is used as data instead of a general one that 'typically serve[s] as a basis for an overall description of a language or language variety' (*ibid.* 15). The corpus consists of a Ja-En parallel news sub-corpus, an En-Ja parallel news sub-corpus, and comparable L1 English and L1 Japanese sub-corpora. Comparable corpora serve to ensure a systematic analysis that breaks out of the confines of a single pair of texts, thus enabling preliminary consideration of the influence on the target text of typical patterns of a specific genre in the target language (Munday 1998a: 542).

The data must be representative of the genre but it is not easy to decide upon the volume of data required in order to be representative. While Biber claims that 'it depends' on the purpose of analysis and on the type of linguistic features and texts to be analysed, he also asserts that a small corpus size is often adequate (1990).

With respect to research of news articles, Caple (2009) constructed a corpus of 1,000 stories for her research on image-nuclear news stories, while Bednarek (2006) used 100 stories in her book presenting an analysis of evaluation in a newspaper corpus. The size of her corpus is roughly 70,000 words, which she claims to be representative enough given that previous research on news discourse claiming to represent corpus-based quantificational approaches use smaller or only slightly larger corpora than hers. Arguably, then, 70,000 words is an appropriate minimum size for a corpus, although the data size needs to be determined in relation to the purpose.

To be representative in terms of the content, Knox (2009) follows the data collection method of a 'constructed week' (Bell 1991); media texts are collected on one day per week for a period of five weeks (Monday to Friday), excluding the weekend as weekend newspapers differ significantly from weekday ones in terms of content. Bell (1991) affirms that it is reasonable to assume that a constructed week is representative of the content of each newspaper for a 6-month period. Although the present study targets linguistic issues and not content, if this method assures the representativeness of content, it can be argued that linguistic representativeness is maximised too. I therefore employ this method, with a slight modification, so that the issue of randomness is addressed and the minimum size of the corpus is not compromised.

Ideally, data is collected randomly but there are limitations. Here, the articles must be parallel – original and translation – but not all articles are translated, which means the data cannot be gathered completely randomly. Moreover, the number of articles translated on each day differs and collecting data on one day a week can be inefficient for assembling reasonably sized corpora. The adjustment, therefore, is to collect a set volume of data a week, instead of on a specific day. In addition, the subject matter is restricted and only national news articles are collected so that there is some degree of information gap between the readers of the original articles and those of the translations such that translators' intervention is anticipated.

The Japanese-English parallel corpus articles are collected from the *Yomiuri* and *Mainichi Shimbun*, two major Japanese daily newspapers that have online versions. The *Asahi Shimbun* was originally included as another major daily with an online version but was eventually excluded as a registration fee is required to read full texts. For the English-Japanese parallel corpus the US editions of Reuters and CNN were sources from which articles were collected

as they have Japanese language websites. All English texts including the English translation of both *Yomiuri* and *Mainichi Shimbun* are in American English.

Every week 3-4 articles were collected from each source and several considerations were given to data collection. The first is to exclude articles at the weekend from CNN and Reuters as weekend papers are significantly different from the weekly ones in terms of content (Bednarek 2006: 220; Knox 2009: 145). In order to keep the Japanese data consistent with the English data, articles from the Japanese source at the weekend were not collected, even though Japanese newspapers and websites do not differentiate between weekdays and weekend where news is concerned. However, this constraint was dropped after a few months. At the beginning of data collection, there were hardly any articles translated at the weekend on both CNN and Reuters. However, gradually the 24 hour news cycle in journalism became apparent along with the translation workflow as articles written and translated during the weekend started to appear.

Another consideration was not to collect articles translated more than a few days after the publication of the original as they pose a problem with the 'timeliness' (Tuchman 1973) of hard news. However, I dropped this as there were articles, for example poll results, which were translated a week or more after the original was published. This indicates hard news is not necessarily just about timeliness, in particular in translation in which the urgency and relevance of the reported event to the reader are not as immediate or to the extent which the original was to the original reader.

The final consideration, which remained until the data collection was complete, was to avoid news articles that report a succession of updates on an incident. The tenor is presumed to be different compared to one-off news articles, which is likely to be reflected in the language used to report the news.

A total of 85 articles were collected from each source between October 2013 and June 2014 as shown in Figure 4-1.

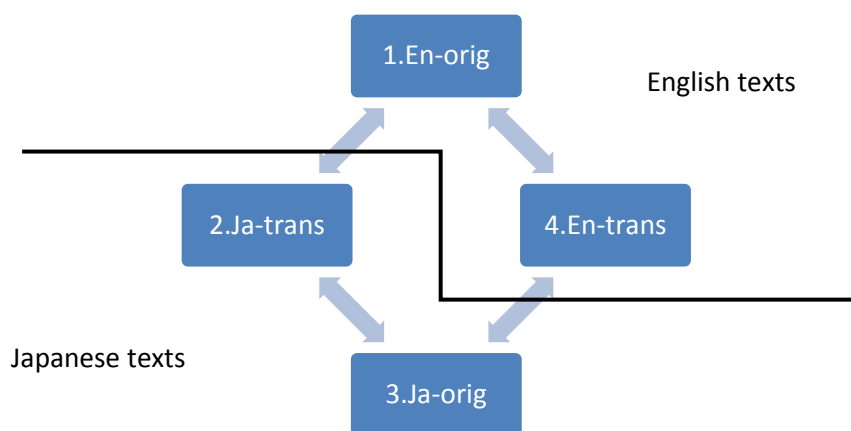
Figure 4-1: Data source and size

Set		Language	Source	Articles	Size
1	En-orig	English (original)	<i>CNN, US edition</i> (36152)	85	71804 words
			<i>Reuters, US edition</i> (35652)	85	
2	Ja-trans	Japanese (translation)	<i>CNN, Japan</i> (44406)	85	81157 characters
			<i>Reuters, Japan</i> (36751)	85	

Set		Language	Source	Articles	Size
3	Ja-orig	Japanese (original)	<i>Yomiuri Shimbun</i> (43262)	85	99,581 characters
			<i>Mainichi Shimbun</i> (56319)	85	
4	En-trans	English (translation)	The Japan News (by <i>Yomiuri Shimbun</i>) (31731)	85	60,171 words
			<i>Mainichi Shimbun, English</i> (28440)	85	

Set 1 consists of English original texts (henceforth En-orig. To identify a particular text C or R plus a number follows that indicates either CNN or Reuters, e.g. En-orig-C1) and Set 2 is the translations of the texts from Set 1 into Japanese (Ja-trans, followed by C or R plus a number as in Ja-trans-C1). They form a parallel En-Ja news corpus. Set 3 consists of Japanese original texts (Ja-orig) and Set 4 translations of those texts into English (En-trans), which form a parallel Ja-En news corpus. The relationship between these sub-corpora is graphically presented below, informed by the CroCo Project corpus map (Newmann & Hansen-Schirra 2005: 6).

Figure 4-2: Relationship of sub-corpora



Sets 1 En-orig and 3 Ja-orig in Figure 4-2 also serve as comparable non-translated corpora of English and Japanese against which the translated texts are compared. By including comparative corpora, unlike simply comparing the source and translated texts, linguistic features depicted in translation can be explained in terms of register.

The corpus I constructed in this project is significant in a number of ways. Firstly, it is a reasonably sized register controlled corpus that is representative of the specific text type of news articles. Secondly, the texts are collected from renowned news providers, in particular in Japanese; *Yomiuri* and *Mainichi* are

two of the three most widely circulated dailies in Japan. Thirdly, all the texts in the corpus are complete articles, paired with their translations.

On the other hand, the corpus presents two possible limitations. Firstly, there are articles in the corpus which are unlikely to be considered typical hard news with regard to topic matter or style (such as how quotations are treated). These remain in the set but a register analysis is conducted on them to establish whether separating them from the rest reveals anything significant in terms of genre or linguistic features.

Secondly, there are articles for which it is difficult to determine if they report national or international news, such as Caroline Kennedy being appointed ambassador to Japan (En-orig-C7) or Toyota's case settlement in the U.S (En-orig-R4 & R30). Although they give the impression that they may be international news because Japan is involved, I determined them to be national/American news and hence they remain in the corpus.

The difficult cases are mostly from CNN or Reuters and hardly any similar cases are observed in the data from *Yomiuri* or *Mainichi Shimbun*. Given that the audience of translated CNN or Reuters news is mainly Japanese, it is understandable that news involving Japan is translated into Japanese as a priority.

On the other hand, a wide range of hard news articles are translated from Japanese into English on The Japan News (by *Yomiuri*) or *Mainichi Shimbun*. According to the reader profile *Yomiuri Shimbun* (2016: 8) provides, 40% of The Japan News readers are foreign nationals while 60% are Japanese. The survey also notes that 58% of the users of The Japan News access from overseas (*ibid.* 15). This partially confirms that the audience for The Japan News consists of English speaking people likely to be living in Japan and Japanese who are studying English. *Mainichi Shimbun* possibly has a similar audience profile.

Thus, the data used in this study is articles of hard, national news. The analysis is straightforward if the data consists of only typical articles, but there are several which appear not to fit this profile. Given the necessity of establishing a reasonable size of corpus, a pragmatic approach was taken and these remained in the data, subject to verification that this atypical data does not significantly differ linguistically. If the register analysis shows them to be different, they can be dealt with separately, potentially offering interesting points for discussion.

Data analysis for Theme and text structure is conducted using the UAM CorpusTool, version 2.8 (O'Donnell 2010), which has useful functionality. Firstly, it supports multilingual analyses as it is capable of dealing with non-alphabetic languages such as Japanese. Secondly, it allows the user to create annotation schemes in the form of system networks, which reflects the vital understanding that the same functions could be realised differently in different languages. Thirdly, more than one descriptive dimension can be included in one project such as Theme realisation and Text unit analyses, but just not at the same time. Fourthly, once a corpus has been annotated, it can yield statistics for relative frequency of realisation. In contrast, data analysis for additions, deletions and moves is conducted manually.

Having explained the construct of the corpus and the tools for analysing it, the following sections explain analytical approaches.

4.2 Unit of analysis

4.2.1 T-unit to analyse Theme realisation

Since it is in the clause that the experiential, interpersonal and textual metafunctions are mapped into a grammatical structure, the clause is the central processing unit in SFL (Halliday & Matthiessen 2004: 10). However, the choice of which unit to employ as the basis of the thematic analysis is not limited to the clause but depends on the analyst's purpose. Thompson (2007: 680) offers the clause, the T-unit and the orthographic sentence as three options.

Analysis by clause will highlight participant continuities in the text, according to Thompson (*ibid.* 680). In relation to Theme for example, Moore (2010) employs the clause as the basis of analysis, claiming it best illustrates the interaction of Participants, Theme and Information. Kim (2011) also adopts the clause to investigate Theme in Korean, stating that it is a more efficient unit for observing Theme patterns than the clause complex. She reasons the clause is better as a unit of meaning, because the way clauses are linked in the clause complex is complicated, and because the topics in those clauses do not remain the same.

An analysis by T-unit or sentence will 'reflect the overall organisation and changes of direction more strongly' (Thompson 2007: 680). The T-unit is 'a clause complex which contains one main independent clause together with all the hypotactic clauses which are dependent on it' (Fries 1995a: 318). For example, focusing on Theme, McCabe (1999) employs the T-unit, which she

terms 'independent conjoinable clause complex', to observe the textual differences in history textbooks written in English and Spanish, claiming that it is the optimal unit of analysis for Thematic progression of a text. Her approach takes the way Spanish behaves into account; a clause in Spanish is considered complete without an explicitly expressed Subject. Because of this, she analyses subjectless verb initial clauses that are continuing clauses in a parataxis as independent clauses, both in Spanish and English to be consistent.

The present study employs the T-unit as the unit of analysis. Investigating shifts in news translation between English and Japanese focusing on Theme requires a unit that accurately illustrates the nature of Theme realisation in each language and also allows comparison between them. Besides the question of what elements are frequently realised as Theme (Chapter 5), how they contribute to the development of texts is also addressed (Chapter 6). The current study adopts, then, the view that the T-unit is the best option, as 'only independent clauses realise moves in the exchange in interpersonal terms and therefore, Themes at this level signal more visibly the way a text is unfolding' (Thompson 2007: 680). Halliday supports the analysis by the T-unit, stating that only independent clauses fully exhibit the option of Theme and that Themes realised in the dependent clauses are recognised as secondary to those in the independent clause (Halliday 1967: 221).

As for Japanese, similar to Spanish, the Subject in a clause is often not explicitly expressed, although the grammatical behaviour is different from Spanish in that Spanish verbs reveal the Subject by conjugation whilst Japanese verbs do not. Also, clause chaining is very common in Japanese (see 4.2.3). It is for these reasons that the T-unit is here considered best for investigating and comparing Theme in English and Japanese. I will demonstrate the suitability of the T-unit for analysing Theme in Japanese in 4.2.3.1. Before discussing the T-unit in Japanese, however, an account of the T-unit in English is presented.

4.2.2 T-unit in English

A T-unit consists of either one clause simplex or a clause complex. By adopting the T-unit as the unit of analysis only the Themes of independent clauses are analysed. Thus, identifying a T-unit involves identifying the interdependency between clauses within a clause complex.

The degree of interdependency of elements is identified by the terms 'parataxis' and 'hypotaxis'. Parataxis at the clause rank is the linking of

clauses of equal status within a clause complex and hypotaxis of unequal status (Halliday & Matthiessen 2004: 374).

If a clause complex consists of more than one independent clause, i.e. paratactically related clauses, there will be more than one T-unit, each of which has its own Theme. However, if the initiating and continuing clauses share the Subject or an element in the Theme position, the continuing clause is considered dependent on the initiating one and they thus form a single T-unit. Below are examples of paratactically related clauses in which the Theme of each T-unit is underlined. Example 4-1 has two independent clauses each with its own Theme, while 4-2 has one independent clause and one Theme for the whole T-unit. Clause boundaries are indicated by //.

4-1 The storm left many of them in ruins, // and now Washington is leaving them out in the cold. En-orig-C11

4-2 Johnson managed to get away from the apartment // and gave a statement to investigators at a local hospital. En-orig-C52

If a clause complex consists of clauses that are hypotactically connected, there will be one T-unit. If the order of the clauses is progressive, i.e. the independent clause followed by a dependent one ($\alpha^{\wedge}\beta$), only one Theme is identified from the independent clause for the whole T-unit as in 4-3.

4-3 Thursday's order from U.S. District Judge James Selna in California federal court put a halt to the lawsuits // while lawyers for Toyota and the plaintiffs try to resolve the cases. En-orig-R30

If the order is regressive and the dependent clause is placed before the independent one ($\beta^{\wedge}\alpha$), the dependent β as a whole clause will be the Theme for the T-unit as in 4-4.

4-4 If the federal government's shutdown ends before then, // the state will receive a refund of unused money, the statement said. En-orig-C10

Taking the whole dependent clause as Theme is reasonable as its function is similar to Circumstances (Thompson 2004: 154-155). In addition, this approach is useful for observing when the dependent β clause is made Theme in English translation in comparison with English original texts or Japanese source texts. With regard to the order of the independent α clause and dependent β clause, English has an option to select from either $\alpha^{\wedge}\beta$ or $\beta^{\wedge}\alpha$. Japanese does not have this option as it has to be $\beta^{\wedge}\alpha$, which means this is one of the places where a shift may occur in translation, in particular in the Ja-En direction.

The final interrelationship between clauses is that of projection. The status of a projected Theme is uncertain as projection involves a different kind of relationship between clauses compared to other types of clause complex (*ibid.* 161). Speeches can be quoted either directly or indirectly. When a speech is quoted directly, combining the projecting and projected clauses with parataxis, the Themes of both projecting and projected clauses are best shown separately, since both of them 'seem to be important in the development of the text' (*ibid.* 161).

With regards to indirect speech, which combines projecting and projected clauses with hypotaxis, there are two possibilities. One is to treat the projected clause as forming a T-unit with its projecting clause. In this case the Theme of the projected clause is not shown separately. The other is to treat the projected clause 'as a separate message on a different "level"' (*ibid.* 162), in which case the Theme of the projected clause should appear separately, as does the Theme of directly quoted speech.

Thompson (*ibid.* 162) prefers the latter approach when dealing with indirect speech. However, in the current study Themes in indirect quotations are not shown separately since they are regarded as the dependent clause within a T-unit, while Themes in direct quotations are shown separately as example 4-5 and 4-6.

4-5 "Given all the things that would have to go right, the probability of recovering the first stage is low" // Shanklin wrote in an e-mail.

En-orig-R41

4-6 A spokeswoman for Toyota, Carly Schaffner, said // the company stood behind the safety and quality of its vehicles.

En-orig-R30

Respecting the structural difference between direct and indirect quotations is intended to reveal any linguistic or functional differences as well as discursal differences between English and Japanese.

Several examples of direct quotations in the corpus are captured in the quotation brackets fragments of the quotation, such as in 4-7.

4-7 Homeland Security said that "out of an abundance of caution" it routinely shares "relevant information" with domestic and international entities, "including those associated with international events" like the Sochi Olympics.

En-orig-C34

Since they are not fully developed as a clause with Theme-Rheme structure, they are not treated as projected clauses.

4.2.3 T-unit in Japanese

In this section I discuss methods of identifying the T-unit in Japanese, covering issues of recognising clauses and the level of interdependency between them in the clause complex. I also present a set of recognition criteria for the T-unit. Before discussing features and issues of identifying the T-unit, however, I attempt to verify that the T-unit is suitable for the current project to analyse the realisation of Theme in Japanese, which is to be compared with English.

The unit of analysis influences what is identified as Theme. Therefore, it is useful to discuss it in relation to how I identify Theme. To do so, I will bring forward some of the discussions from 4.3.3. and argue here for adapting the initial position over the particle *wa*, which is sometimes termed as a Theme marker (Teruya 2006). The discussion commences with an introduction of a unit called the Theme-unit.

4.2.3.1 Unit of analysis, Theme and particle ‘*wa*’

It has been proposed that textual metafunctions of Japanese are organised over a larger unit than the clause because of ellipsis and clause-chaining (Tatsuki 2004; Thomson 2005a). Thomson introduces a unit called the Theme-unit and defines it as ‘a segment of text that forms a co-referential unit in which the referent (that which is retrievable in any non-initial clauses in the unit) is the first constituent in the unit’ (2005a: 16). It can map onto segments larger than the sentence, and this occurs when the co-referential ties extend beyond the full stop as in the example 4-8. A sentence boundary is indicated by the dotted line.

4-8

Theme	Rheme
SeifuWA 政府は	1 2 nichi [[[[chiikigentei DEdaitannakiseikanwa O susumeru]]kokkasenryakutokkuDE 12日、[[[[地域限定で大胆な規制緩和を進める]]国家戦略特区で、 gaikokujinrōdōshaNO u ke i rekakudaiNITSUITEkentōsuru]]kotoOkimeta 外国人労働者の受け入れ拡大について検討する]]ことを決めた。
The government	on 12 th , [[[[area restricted DE bold deregulations O promote]] national strategic special zone DE, foreign worker NO accept expand ABOUT consider]]-NOM O decide-pst.

<p>-</p> <p>(It)</p>	<p>[[[[TokkunaiDEkigyōYAkajishiennadoOokona u]]gaikokujin O taishōNI aratana [[[[特区内で起業や家事支援などを行う]]外国人を対象に、新たな zairyūshikaku O ataer u]]kotoOsōteishiteiru 在留資格を与える]]ことを想定している。</p> <p>[[[special zones inside DE starting business and housework support, etc. O do]]] foreigner O target NI, new residency status O give]]-NOM O suppose-ASP.</p>
----------------------	---

Ja-orig-Y73

Example 4-8 illustrates how more than one sentence can be included in a Theme unit. The Subject of the second sentence is not explicitly realised but it is ‘the government’, which is a co-referent and shared with the first sentence. This ties the two sentences together as a unit. Thomson asserts that the Theme unit ‘functions to signal a shift in Theme’, indicating ‘one phase of continuous thematic progression’ (2005a: 25-26). Indeed, by its capacity to manage implicit Subjects, it appears an ideal unit of analysis to observe how a text develops via Theme realisation.

However, it raises a fundamental question and, when it is applied to actual analysis, several problems too. The essential problem relates to the concept of Theme, that is, Theme as a grammatical category. Since grammar does not exist beyond sentences (Halliday & Matthiessen 2004: 586), it may not be a valid unit to analyse Theme. As a unit of meaning it is useful but it seems associated with cohesion which is a non-structural resource more than Theme which is structural/grammatical.

Major practical problems with this unit are the difficulty of deciding a unit boundary when a Circumstance is the co-referential element, or deciding a Theme between an implicit Subject and explicit Circumstance. Both problems can be solved by employing the T-unit because it does not go beyond the sentence boundary. How I identify Theme within the T-unit is discussed in 4.3.3.

Theme-unit has been introduced to manage Subject ellipsis and clause chaining (Thomson 2005a: 16). Although I do not employ it for the reasons above, I agree that a unit larger than the clause is better to analyse textual development in Japanese, which helps to keep to the global development of text and avoid being distracted by local orientations, in particular when non-Subject Participants are realised as Theme.

Prior to describing the suitability of the T-unit in combination with the initial position for analysing Theme, I will present a textual analysis focusing on the particle *wa*. If I were to take *wa* as a Theme marker in the narrowest sense,

that is, to recognise any element marked by *wa* as the Theme, this automatically determines the unit of analysis. This ignores the sentence boundary and therefore poses the same problem as the Theme-unit, but I will disregard it here for the sake of comparison with the initial position of the T-unit.

Our observation uses text Ja-orig-M69: Osaka woman stabbed to death in suspected stalking-murder case. In order to locate the elements with *wa* in the text, within the sentence and the paragraph, the original Japanese text is presented. All the elements marked by *wa* are underlined, followed by their translations in brackets. Numbers on the left indicate the paragraphs. The sentence ending is marked by the 句点 *kuten* '。'.

4-9 女性死亡:路上で刺され ストーカー相談の男確保 大阪

- 1) 2日午前2時半ごろ、大阪市平野区長吉長原西3の路上で、「女性が倒れている」と近くの男性から110番通報があった。女性は (The woman) 腹部に刃物が刺さった状態で、病院で死亡が確認された。大阪府警は (Osaka Prefectural Police) 女性を刺したとみられる男(57)=平野区=を確保、殺人容疑で逮捕する方針。女性は (The woman) 男から殺意をほのめかすメールを受け取り、ストーカー被害を府警に相談していた。
- 2) 府警によると、死亡したのは (that who died) 大阪府松原市大堀3の飲食店アルバイト、井村由美さん(38)。同日午前2時半ごろ、近所の住人が男女が言い争う声を聞き、井村さんが歩道脇の植え込みに倒れているのを見つけた。井村さんのものとみられる自転車もあった。
- 3) 井村さんは (Ms Imura) 首や腹部を数カ所刺されていた。府警は (The Prefectural Police) 井村さんが帰宅途中に襲われたとみている。
- 4) 現場から白い車が逃げるのが目撃され、府警は (the Prefectural Police) 午前6時15分ごろ、現場の約1.2キロ先の平野区のコンビニ店駐車場で似た車を見つけた。捜査員が車内の男に職務質問したところ、女性を刺したことをほのめかした。男は (The man) 指を負傷し、病院で治療を受けた。
- 5) 府警によると、男は (the man) 2013年8月、井村さんのアルバイト先の飲食店に客として訪れた。男が井村さんに一方的に好意を寄せ、店の女性経営者が今年2月、男に店に来ないよう連絡したが、井村さんへの電話やメールは (calls and e-mail to Ms Imura) やまなかった。
- 6) 井村さんは (Ms Imura) 3月1日、男から「殺される前に警察に電話してや」と殺意を示唆するメールを受け取り、翌日、府警松原署にストーカー被害の相談をした。
- 7) 松原署は (Matsubara police station) 同日、ストーカー規制法に基づき男に電話で口頭で注意し、同12日には (On 12th of the same month)、男を署に呼び出して文書で警告した。4月2日、署員が連絡したところ、井村さんは (Ms Imura) 「その後は

(After that) 何も起きていない」と話したという。**府警は (The Prefectural Police)** 署の対応などに問題がなかったかを含め詳しい経緯を調べている。

8) **現場は (The place (of crime))** 大阪市営地下鉄谷町線・八尾南駅の西約1.5キロの住宅街。

There are 17 elements here marked by the particle *wa*. The locations of *wa* do not correspond to any unit such as clause, clause complex or paragraph. This means that the unit of analysis is automatically decided by *wa*, which inevitably means frequent ellipsis is contained. The elements with *wa* are mainly the Subjects; the victim, the suspect and the police, which illustrate how the text develops reasonably well.

However, there are two concerns for assigning *wa* as the Theme marker. Firstly, it seems ill-advised to exclude all the clauses or sentences that do not have *wa* as if they did not contribute to the development of the text. Secondly, there seems a strong tendency that the elements marked by *wa* are mostly the Subject.

With respect to the first question, let us examine the first sentence of the text, the lead (Chapter 3), which does not have an element marked by *wa*. This sets the scene introducing the time and place of the event and how the victim was found. It could be argued that a clause Theme is not necessary here since the whole sentence is a ‘macro-Theme’ (Martin 1992: 437) and functions as introduction to the whole text. This would be rational if it were to occur regularly, but it is not uncommon for the lead sentence to have an element marked by *wa* (Chapter 6). Nonetheless, it is useful to identify a Theme as a point of departure for the lead sentence irrespective of the status of a macro-Theme.

In summary, in order to investigate how a text develops, it seems apt to make use of as many points of departure as possible. This is also beneficial for examining translation shifts. Given that different languages have different means of indicating Theme and Rheme (Rose 2001), the more Themes available for comparison the better. The elements marked by *wa* provide important points of departure. However, employing the T-units is better because they offer more of them, regardless of the presence of an element marked by *wa*. It is reasonable to recognise all the T-units as contributing to the textual progression and identify a Theme. Moreover, employing the same unit of analysis in English and Japanese is useful for the simplicity and validity of comparison.

The second question, that of a bias towards the Subject, concerns the type of elements that are marked by *wa*. In example 4-9 above, following the elements marked by *wa* means mainly focusing on the Participants/Subjects; 15 elements out of 17 are such. Therefore, tracking the elements plus *wa* as Theme means the focus is primarily on Participants, infrequently on Circumstances and hardly ever on textual or interpersonal Themes. This comes very close to the early definition of Theme; 'it is what the clause is going to be about' (Halliday 1985: 39), hence the suggestion that *wa* is the Theme marker.

However, we have seen that Theme is not necessarily what the clause is about (Chapter 2) but that it functions as 'an orienter to the message'. In order to observe the development of text, focusing on the elements marked by *wa* only is not sufficient, although Participants are profoundly important.

Because of the morphological restriction as to what *wa* can mark, i.e. noun groups, Circumstances are often not marked by *wa* as they are not always noun groups. They have different textual functions from the Participant's and both are important for the text's development. Moreover, textual and interpersonal Themes, though not present in example 4-9, need to be included in the analysis. Teruya's criteria for identifying Theme in Japanese therefore, include additional items to allow these to be recognised as Theme (Chapter 2 Section 3.2).

Focusing on *wa* also raises the question of whether Themes in Japanese must be Given, related to the system of INFORMATION. *Wa* is used to mark the Topic and Topics are characteristically definite (Li & Thompson 1976; Shibatani 1990; Teramura 1991 amongst others) and, thus, Given.

Although the sequence of unmarked informational structure is Given followed by New (Halliday & Matthiessen 2004: 89), conflating the Theme with Given and Rheme with New, it does not always have to be this way. The system of INFORMATION is related to but different from the system of THEME (*ibid.* 87-94). Fries (1983) acknowledges two different approaches; the 'combining approach', equating Theme with Given, or the 'separating approach', disassociating the two. Teruya claims he separates THEME and INFORMATION (2006) but suggesting *wa* to be a Theme marker risks them integrate with each other to a certain extent.

The separating approach is more suitable for the present project and I do not consider *wa* as a Theme marker. This permits Themes to be New, which is marked by *ga*, the Subject marker. Teruya covers this issue in his criteria for

identifying Theme, stating if an element marked by *wa* is absent, the Theme extends up to and includes the topical Theme which may be marked by *ga* (2006: 96).

His criteria are comprehensible but they have several layers involving conditions. His efforts to include *wa* seems to be the cause for that. The question arises then, whether or not *wa* indeed has a significant role in the system of Theme, and if conditions are needed to recognise elements as Theme that could be with or without *wa*.

In addition, it is doubtful whether all the elements marked by *wa* are equally qualified to be Theme if there are more than one in a clause or a clause complex. In example 4-10 two elements are marked by *wa* but the second does not seem to function at the same level as the first.

4-10 日本での特許出願は(wa) 05年は(wa) 約 420,000 件と世界最多だったが、06年に米国、10年に中国に抜かれ、

Patent applications in Japan (wa) in 2005 (wa) about 420,000, the highest number in the world but in 2006 by the United States and in 2010 by China (Japan was) overtaken, ...

Ja-orig-Y74

The second *wa* in '05-nen wa' indicates a contrast with the following years 2006 and 2010. Since all these years relate to 'patent applications in Japan', the first element marked by *wa*, it is practical to assume that is the Theme, while '05-nen wa' is not, or at least it does not function at the same level as the first. Although Teruya (2006: 96) regards all that are marked by *wa* as Theme, it is best to interpret the first element marked by *wa* in the unit of analysis as the Theme but not the following ones.

In summary, elements marked by *wa* indicate how the text develops reasonably well but relying only on them misses important aspects of textual development, i.e. Circumstances as well as textual and interpersonal Themes are better to be recognised contributing to text development. It is also important not to include all the elements marked by *wa* in the Theme as not all function as such.

Having established that *wa* is not to be regarded as a Theme marker, I will next demonstrate that the initial position is suitable for observing textual development. A text analysis, using the same text as 4-9, shows that the initial position provides Theme more than *wa* does. It is notable that most of the elements marked by *wa*, 15 out of 17 recognised in 4-9, remain in this approach, appearing at the beginning of the T-unit.

All the elements that appear at the beginning of the T-unit are presented in 4-11. For reference it includes several that are marked by *wa*, thus recognised as Theme in 4-9 but which are disqualified in the current approach in 4-11. To differentiate between the two approaches of 4-9 and 4-11, underlines and borders are applied. Underlined elements are the newly recognised Themes because they are in the initial position, which have not been Themes in 4-9. Bordered elements were Themes in 4-9 but are no longer so as they are not in the initial position. Numbers on the left indicate the paragraphs.

Several Themes are realised in direct speech, which are in the quotation marks 「 」.

4-11

- 1) 2日午前2時半ごろ、大阪市平野区長吉長原西3の路上で (About 2:30 AM (May) 2 in the street of Osaka Hirano-ku Nagayoshi Nagahara Nishi 3)、
「女性が (A woman)…」…。
女性は (The woman WA)…。
大阪府警は (Osaka Prefectural Police WA)…。
女性は (The woman WA)…。
- 2) 府警によると (According to the police、死亡したのは (that who died WA))…。
同日午前2時半ごろ (On the same day around 2:30 AM)…。
井村さんのものとみられる自転車も (A bicycle that seems to belong to Ms Imura)…。
- 3) 井村さんは (Ms Imura WA)…。
府警は (The Prefectural Police WA)…。
- 4) 現場から白い車が逃げるのが (That a white car running from the site)…。
府警は (the Prefectural Police WA)…。
捜査員が車内の男に職務質問したところ (When police officers questioned the man in the car)…。
男は (The man WA)…。
- 5) 府警によると (According to the police)、男は (the man WA))…。
男が (the man)…。
店の女性経営者が (the female owner of the restaurant)…。
井村さんへの電話やメールは (calls and e-mail to Ms Imura WA)…。
- 6) 井村さんは (Ms Imura WA)...

「殺される前に (Before (you are) killed)...」...

7) 松原署は (Matsubara police station WA)...、

同12日には (On 12th of the same month WA) ...。

4月2日、署員が連絡したところ (When a police officer contacted (her) Apr 12)、

井村さんは (Ms Imura WA)

「その後は (After that WA)...」...

府警は (The Prefectural Police WA)...

8) 現場は (The place (of crime) WA)...

There are 25 elements that are realised as Theme at the beginning of the T-unit. This exceeds the 17 elements marked by *wa* and it naturally means more signposting is provided for the textual development analysis. They involve a variety of Themes ranging from the Participant, to different types of Circumstances, to temporal Location, spatial Location, Angle, and to dependent hypotactic clauses. The majority of the Participants are marked by *wa*, hence as far as the Participant is concerned, those identified as Theme in this approach are almost the same as when Subjects are identified as being marked by *wa*.

While elements marked by *wa* display the textual development primarily around the Participants, Themes realised in initial position exhibit the textual development more broadly taking Circumstances both realised as group or clause into account. Initial position does not discriminate what can be Theme in terms of classes of group or rank, unlike *wa* which is morphologically restricted. For the purpose of this study, to analyse translation shifts, employing the initial position is more useful as it gives a more general perspective of the development of the text compared to *wa*.

4.2.3.2 T-unit and the suspensive form

In the current section and the one which follows I will describe how clause complexes are structured in Japanese, how I recognise clauses and identify T-units. I will begin with the general notion of clause complex, then move on to specific issues with the suspensive form, which creates clause chaining that relates to the boundary of the T-unit.

A T-unit consists of either a clause simplex or a clause complex. With regard to the clause complex recognising the interdependency between the clauses is important since only the Themes in the independent clauses are analysed. Clauses in the clause complex have a fixed order in Japanese. If they are paratactically related, the order is the initiating clause followed by a continuing

one as in 1[^]2. This is the same order as English. If they are hypotactically related, the dependent clause comes first and is followed by the dominant one as in β[^]α. In English there is an option of α[^]β but not in Japanese.

The basic word order of Japanese is Subject-Object-Verb and clauses are linked in a clause complex by changing the form of the verb in the secondary clause; either in the initiating clause 1 in 1[^]2 as in 4-12 or in the dependent β clause in β[^]α in 4-13.

4-12

[[2, 3回被害を受けた]] 留学生も	<u>おり</u> 、	同校が	鳥栖署に	被害届を	出していた。
[[2,3-kai higai O uketa]] ryūgakusei MO	<u>ori</u> .	dōkō GA	Tosusho NI	higaitodo ke O	dashiteita
[[2,3 times assault O receive-pst]] foreign student MO	exist- <u>SUSP</u>	same school GA	Tosu police NI	assault report O	file-ASP-pst
There are foreign students who were assaulted 2-3 times and the school has filed a report to the Tosu branch of the police.					
1 clause		2 clause			

Ja-orig-M82

4-13

M8.5 以上が発生すれば、	…首都圏の広範囲を強い揺れが襲うことになる。
M8.5 ijō GA <u>hasseisureba</u>	...shutoken NO kōhan'i O tsuyoi yure GA osou koto ni naru
Magnitude 8.5 over GA <u>occur-COND</u>	capital NO extensive area O violent quake GA attack NOM NI become
If a magnitude 8.5 or over temblor were to occur, it would violently shake extensive areas in and around the capital.	
β clause	α clause

Ja-orig-M21

The secondary clause will never be an independent one because the verb in the secondary clause is always bound in terms of the form. Hence, clauses in a clause complex do not have an equal status in Japanese in a strict sense. However, it is conventional in Japanese linguistics to deem them to have an

equal status in terms of meaning (Masuoka 1997: 4). Thus, both parataxis, the linking of clauses of equal status, and hypotaxis of unequal status are possible.

From the perspective of meaning, secondary clauses are dependent when they indicate, for instance, the time when the event in the primary clause occurs or conditions for the event in the primary clause to take place, being in a hypotactic relationship with the primary one. These meanings are generally expressed by particular forms. Teruya groups them into two categories depending on how clauses are linked in a clause complex from the structural point of view.

1. Adnominalization, or 連体形 *rentaikei*, i.e. an addition of a structural conjunction such as the conjunctive nominal *toki* 時 “when”, the conjunctive postposition *made (ni)* まで(に) “until, (by)”, or the conjunctive particle *noni* のに “despite”;
2. Verbal conjugation, such as ‘conditionalization’ as in *sur-e-ba* “provided”, *sur-u-to* すると “if, when(ever)”, *sur-u-nara* するなら “if...”, or ‘infinitivization’ (S. Martin, 1975:238) such as the suspensive form as in *shite* “and, -ing”.
(Teruya 2006: 333-334)

If clauses are linked by the ways that are described in 2 above, they are in a hypotactic relationship. In this case the verb in the dependent clause conjugates morphologically to indicate what relationship it is in with the dominant clause. The verb, and thus the clause too, is bound and it has no option to select for Mood, as shown in example 4-14.

4-14

[[捜査員が車内の男に職務質問した]]ところ、	[[女性を刺した]]ことをほのめかした。
[[Sōsain GA Shanai NO otoko NI <u>shokumu-shitsumon shita</u>]] tokoro	[[josei O sashita]] koto O honomekashita
[[Investigator GA car inside NO man NI question-pst]] upon	[woman O stab-pst]-NOM O suggest-pst
When investigators questioned the man in the car, (he) suggested that (he) had stabbed the woman.	
β clause	α clause

Ja-orig-M69

If clauses are linked by the ways illustrated in 1 above, they could be either paratactic or hypotactic. Teruya says that ‘paratactic structural conjunctions

are attached to the finite form of ... verbs that select tense relative to the speech event ... while hypotactic ones do not freely specify tense in the same way' (2006: 335). In example 4-14 above, clauses are hypotactically connected to each other and the verb in the β clause is bound to be in the past tense.

Most of the forms in the secondary clause are the realisations of particular meanings. However, there is a verb form whose relationship to meaning is not clear or specific, and that is the form called the suspensive form. The suspensive form is 'used to keep the flow of events in suspense until such a time that their logical relationship to what comes next becomes clear' (Teruya 2006: 38). A sentence does not stop where it can but continues on to create a compound sentence by a series connection (Morita 1985: 752). 4-12 above is an example of this.

In order to consistently identify the T-unit, two issues in relation to the suspensive form need to be clarified. Firstly, whether or not the suspensive form always functions at the clause rank, and, secondly, how the suspended clauses are decided whether they are parataxis or hypotaxis. The first issue is addressed in this section while the second is discussed in section 4.2.3.3, where the relationship between the suspensive form and taxis is examined.

Before considering the first issue, it is useful to give some details about the suspensive form. There are two types, the first being 第一中止め the first *nakadome* [suspensive] or 連用形 *ren'yō* [continuative] form, and the second being 第二中止め the second *nakadome* [suspensive] or the テ形 *te*-form.

Figure 4-3: Suspensive form

verb	Suspensive form	
	First <i>nakadome</i> (<i>ren'yō</i> form)	Second <i>nakadome</i> (<i>te</i> -form)
食べます <i>tabe-masu</i>	食べ <i>tabe</i> suspended + zero suffix	食べて <i>tabe-te</i> suspended + suffix <i>-te</i>

'(T)he clauses or predicates that are connected in this way are similar to the relationship of nouns connected by the particle *to* [and] and form a relationship of equal status between the elements' (Kokuritsu-kokugo-kenkyūjo 1981: 18, my translation). Creating parallel relationships via verbal conjugation is the main function of the suspensive form.

Historically, the *ren'yō* form, which means 'linked to the predicate', was the original one to fulfil this purpose. However, its connecting function became weak over time, and was later reinforced by the added connecting particle *te* that explicitly shows that the verb in this form is to be connected to the final verb of the sentence (Morishige 1973: 239).

The general understanding is that the *ren'yō* form indicates paratactic relationship of verbs, while the *te*-form shows signs of dependency on the verb that follows (Kokuritsu-kokugo-kenkyūjo 1981; Masuoka 2012 amongst others). In other words, although the *ren'yō* form and the *te*-form are both considered to connect verbs or clauses creating a series relationship, the degree of interdependency between the connected clauses is different. This will be discussed in relation to taxis (4.2.3.3).

Based on the above, the first question regarding whether or not the suspensive form functions at the clause rank all the time is addressed. There are two types that are better treated at the word/group rank. One is postpositional phrases and the other is so-called 'adverbial usage'.

Postpositional phrases are derived from verbs and are similar to minor processes (Teruya 2006: 318-319). Outside SFL of Japanese, Morita & Matsuki (1989: xi) call these '複合辞 *fukugōji* [compound particles/auxiliaries]'. They explain that these expressions are amalgamated words which serve specific grammatical functions as their meanings are more than the combination of each of the words. Examples of possible postpositional phrases in my data are below, including several that pose uncertainty as to whether they are prepositional phrases or clauses.

Figure 4-4: Possible postpositional phrases in the corpus

			Meaning	Alternative suspensive form
1	～について	- <i>ni tsuite</i>	about -	～につき
2	～に関して	- <i>ni kanshite</i>	regarding -	～に関し
3	～に対して	- <i>ni taishite</i>	towards -	～に対し
4	～を巡って	- <i>o megutte</i>	concerning -	～を巡り
5	～に応じて	- <i>ni ōjite</i>	depending on -	～に応じ
6	～に加え	- <i>ni kuwae</i>	in addition to -	～に加えて
7	～と比べて	- <i>to kurabete</i>	compared to -	～と比べ
8	～を含め	- <i>o fukume</i>	including -	～を含めて
9	～を踏まえ	- <i>o fumaie</i>	taking - into account	～を踏まえて

10	～と見て	- <i>to mite</i>	seeing as-	～と見
11	～として	- <i>to shite</i>	taking as -	～とし
12	～と併せて	- <i>to awasete</i>	combined with -	～と併せ

Similar to Morita & Matsuki, Teruya (2006: 319) explains that postpositional phrases normally have a particular circumstantial function. As to how to define them, there is no established definition (Morita & Matsuki 1989: xi). However, they suggest two major reasons for determining an expression to be a compound particle. One is the level of tightness of the united components and the other is the level of formalisation.

For example, - *ni tsuite* [about] (1) in Figure 4-4 is a combination of the particle *ni* and the suspensive form of 付く [attach]. As a unit the original meaning of the verb is unrecognisable. In addition it is so tightly united that no other word can be inserted between the particle *ni* and the suspensive form *tsuite*. The form is also too fixed to be used in other forms, i.e. the verb cannot conjugate without losing the derived meaning 'about'.

There is no general consensus as to which to take as postpositional phrases and which not to take. There is a different degree of amalgamation of components in each expression, which is a reason for disagreement. Teruya's account of the ambiguous cases is that 'they may be in the process of grammatical transition, moving along the rank of group to postpositional phrase' (2006: 319). Based on the assessment making use of the two criteria described above, i.e. the level of tightness and formalisation, I judge that expressions 1-7 in Figure 4-4 are postpositional phrases, while 8-12 are clauses.

The second type of suspended verbs whose operational rank needs to be considered have the function of adverbs. There are not many examples of these in the corpus, which is indicative of the fact that they are rare in the register of news articles.

4-15

しかし	警察は	…情報として	「 」と	誤って	入力し、
<i>Shikashi</i>	<i>keisatsu</i> WA	… <i>jōhō TO-shite</i>	“…” TO	<i>ayamatte</i>	<i>nyūryokushi</i>
But	police WA	information as	“…” PROJ	by mistake	input-SUSP
	Actor	Circ. (postpositional)	Projected	Circ.	Process

Ja-orig-M75

Thomson (2001: 67) presents two types of suspended verbs used as one word; one being 急いで *isoide* [in a hurry] and 黙って *damatte* [in silence] as rankshifted clauses functioning as Circumstances of Manner, and the other すると *suruto* [then] and そう言って *sō itte* [saying so] also rankshifted clauses functioning as Connectives. I have a different view on the example of そう言って *sō itte* [saying so], which I would treat as a clause because of its productivity; *sō* in *sō itte* can be replaced by any projected clause that is followed by the quotation particle *to* as in *wakatta to itte* [saying ‘understood’]. In addition, the level of formalisation, employed to assess postpositional phrases, is not very strong as it can be a sentence final verb and it can select for MOOD.

Regarding these Circumstances of Manner expressed by the verb *te*-form, there is as yet no agreement as to whether to treat them as adverbs or as suspended clauses serving an adverbial function. While Teruya offers *damatte* [in silence] as an adverbial group (2006: 146), Kokuritsu-kokugokenkyūjo (1964 cited in Hayashi 2007: 56) does not categorise either this or *aratamete* [on another occasion] as adverbs. Possible reasons for this are that they are not fixed as units, the meaning of the verb in the *te*-form is not derived, and *aratamete* can take a constituent such as *hi o aratamete* [on another day] (Hayashi 2007: 56). They are moving on the cline of lexis and grammar in a similar way that postpositional phrases are.

The present project recognises the verbs in the *te*-form as adverbs only if they are on their own with no attendant Participants and Circumstances, just as Thomson (2001) does. There are several examples of the verb *te*-form serving the meaning of Circumstance with attendant Participants and/or Circumstances in the present corpus. They are interpreted as clauses that modify or define the Process or the whole clause as an adverb does.

4-16

この印紙は	[[詐取した]]通帳を用いて	購入していた。
<i>Kono inshi WA</i>	[[<i>sashushita</i>]] <i>tsūchō O mochiite</i>	<i>kōnyūshiteita</i>
This revenue stamp WA	[[obtain-pst (by fraud)]] bank account book O <u>use-SUSP</u>	purchase-ASP-pst
α Goal	β clause (Goal + β Process) suspended	α Process

Ja-orig-M72

This example has a Goal as Theme. It conflates with the Topic of the Topic-Comment structure so the rest of the sentence, the Rheme, is Comment about ‘*Kono inshi* [this revenue stamp]’. The clause in question ‘*sashushita tsūchō o mochiite*’ is dependent on the α Process ‘purchased’, explaining the way the revenue stamps were obtained.

To conclude, not all the suspended verbs function at the clause rank but some operate at the word/group rank. They are postpositional phrases with the verbs in the *te*-form functioning as adverbs. Once the suspensive forms that operate at the clause rank have been determined, whether they form a paratactic relationship or a hypotactic one with the primary clause will be examined.

4.2.3.3 The suspensive form and taxis

This section is concerned with the level of interdependency between clauses linked by the suspensive form in the clause complex. The relationship of the suspensive form and taxis must be clarified as it affects where a T-unit ends, and consequently what is realised as the Theme of each T-unit.

As seen earlier, there are two types of suspensive form; the first suspensive form (*ren'yo* [continuative] form) and the second (the *te*-form). Traditionally they are treated as the same, both creating paratactic relationships between verbs or clauses. However, not all linguists conform to this and claim different forms mean different meanings appear. They argue that the *ren'yo* form is typically paratactic while the *te*-form is hypotactic (Kokuritsu-kokugo-kenkyūjo 1981; Morita 1989; Masuoka 2012).

4-17

[[2, 3回被害を受けた]] 留学生も	<u>おり</u> 、	同校が	鳥栖署に	被害届を	出していた。
[[2,3-kai higai O uketa]] ryūgakusei MO	<u>ori</u>	dōkō GA	Tosusho NI	higaitodoke O	dasiteita
[[2,3 times assault O receive-pst]] foreign student MO	<u>exist-</u> <u>SUSP</u>	same school GA	Tosu police NI	assault report O	file-ASP-pst
There are foreign students who were assaulted 2-3 times and the school has filed a report to the Tosu branch of the police.					
1: suspended by the ren'yo form		2			

Ja-orig-M82

4-18 (repeat of 4-16)

この印紙は	[[詐取した]]通帳を	<u>用いて</u>	購入していた。
<i>Kono inshi WA</i>	[[sashushita]] tsūchō O	<u>mochiite</u>	<i>kōnyūshiteita</i>
This revenue stamp WA	[[obtain-pst by fraud]] bank account book O	<u>use-SUSP</u>	purchase-ASP-pst
This revenue stamp has been purchased using the bank account books obtained by fraud.			
α Goal (Topic)	β clause: suspended by the <i>te</i> -form		α Process

Ja-orig-M72

Within SFL, Teruya too regards the *te*-form as creating a hypotactic relationship and the *ren'yō* [continuative] form paratactic. He explains that the *te*-form is used when two events take place consecutively and that the verb in the *te*-form is subordinate to the verb in the primary clause. With the *ren'yō* form each clause in a clause complex is highly likely to have its own Subject instead of having a shared one (Teruya 2006, 2015). Hence, the *te*-form is hypotactic and the *ren'yō* form paratactic. Thomson takes a slightly different position and locates the *te*-form as intermediary, stating they are 'hypotactic in terms of structure yet paratactic in terms of order or sequence' (2001).

Thomson's approach is pragmatic, with which I agree in theory but a practical solution for delimiting the T-unit is needed. It is tempting to separate the two suspensive forms as Teruya does, simply by the form which supposedly reflects the functional difference. However, there are clauses suspended by

the *ren'yō* form which can be replaced by the *te*-form. In her research of the suspensive form, Sawanishi (2003) claims that in newspapers the *ren'yō* form is selected 85% of the time when both the *ren'yō* form and the *te*-form are equally appropriate. This suggests that the two suspensive forms may be used differently from the general description presented above in the present register.

In order to have a better understanding as to how the two suspensive forms are used in newspapers, I conducted a small-scale case study using a portion of the Ja-orig data and observed how the suspensive forms were distributed and how suspended clauses were linked to the primary clause. The data includes 24 articles collected in May 2014. The number of clauses was 242, out of which 129 were clause complexes. I examined whether or not the *ren'yō* form always created parataxis and the *te*-form hypotaxis. Only the most extensive clause nexuses that were direct constituents of the clause complex were investigated.

To assess the level of independency between clauses I employed two criteria. Firstly, whether or not each clause had its own Subject. If it did, it increased the independency of the secondary clause. Secondly, if the secondary clause modified/defined the primary clause or expressed a separate event that occurred concurrently with or successively following the primary one.

The most frequent means of connecting clauses is the *ren'yo* form (98 instances), followed by verbal forms such as conditionals (19), then the *te*-form (7), and then a small number of projections (ideas and indirect quotes, 5). As Sawanishi (2003) indicates, there are indeed more clauses suspended by the *ren'yō* form than the *te*-form. However, not all of them create parataxis.

In fact, 57 out of the 98 clause complexes connected by the *ren'yo* form share a co-referential element such as a Participant or a Circumstance, by which they form one T-unit rather than two separate ones.

4-19

男は	指を	負傷し、	病院で	治療を	受けた。
<i>Otoko WA</i>	<i>yubi O</i>	<i>fushōshi,</i>	<i>byōin DE</i>	<i>chiryō O</i>	<i>uketa</i>
Man WA	finger O	injure-SUSP	hospital DE	treatment O	receive-pst
The man injured his fingers and received treatment at the hospital.					
1 clause (suspended by the <i>ren'yo</i> form)			2 clause		

The number of the *ren'yo* form paratactically connecting clauses, mostly with a separate Subject in each clause, is 39. This means each clause forms one T-unit.

There are only seven instances of clause complexes suspended by the *te*-form, none of which are paratactically related with a separate Subject in each clause. Two clause complexes share a co-referential element by which they form one T-unit as in 4-20.

4-20

同署は	//[[薬物を使用していた]]可能性もあると//	見て	調べる。
<i>Dōsho WA</i>	[[<i>yakubutsu O shiyōshiteita</i>]] <i>kanōsei MO aru TO</i> //	<i>mite</i>	<i>shiraberu</i>
Same police station WA	//[[<i>drug O use-SUSP-pst</i>]] possibility MO exist-PROJ//	<u>see-</u> <u>SUSP</u>	investigate
The same police station presumes there is a possibility (he) was taking drugs and will investigate.			
α Actor/β Sensor	β Clause (suspended by the <i>te</i> -form)		α Process

Ja-orig-M74

'*Dōsho wa* [The same police station]' is the co-referential element made the Topic of the clause complex, which is both the Actor of the primary Process '*shiraberu* [investigate]' and the Sensor of the secondary Process '*mite* [look (suspensive)]'.

Out of the seven clause complexes suspended by the *te*-form, five are hypotactically related as in 4-21, creating one T-unit together. The suspended clause here functions as Circumstance of Cause; the reason for the three boys to be arrested, which is described in the dominant α clause. What is placed in initial position distinguishes 4-21 from 4-20. It is not impossible for the Actor in the α clause in 4-21 to be in initial position, in which case it will have the same structure as 4-20, forming one T-unit because of the shared Subject.

4-21

//佐賀県鳥栖市の路上で留学生に生卵を投げつけたと//して	県警鳥栖署は	26日	…少年3人を	暴行容疑で	逮捕した。
//Saga-ken Tosu-shi NO rojō DE ryūgakusei NI namatamago O nagetsuketa TO//shite	kenkei tosu-sho WA	26-nichi	shōnen 3-nin O	bōkōyōgi DE	taiho shita
//Saga-prefecture Tosu-city NO street DE foreign student NI raw egg O throw-pst PROJ//take-SUSP	prefectural police Tosu branch WA	on 26 th	3 boys O	suspicion of assault DE	arrest- pst
Taking as (they) threw eggs to foreign students in the street of Tosu-city, Saga, the Tosu branch of the police arrested three boys on suspicion of assault on 26 th .					
Projection + Process	Actor	Circ.	Goal	Circ.	Process
β clause	α clause				

Ja-orig-M82

Suspended clauses in initial position often function as Circumstances of Cause or Manner in relation to the primary clause. Unlike the other forms such as conditionals whose meaning is associated with the form, the meanings of Cause or Manner with the *te*-form are only inferred by the relationship between the clauses in the clause complex (Kokuritsu-kokugo-kenkyūjo 1981: 36). Because the form and meaning are not fixed, the *te*-form at times poses a difficulty in interpreting the relationship it creates with the primary clause.

Although the data size is not extensive, we can infer approximately how the two suspensive forms are used in the register of news. Clauses in the clause complex are connected by the *ren'yo* form more frequently than the *te*-form. The *ren'yo* form is generally recognised as creating paratactic relationship between connected clauses, although they often share a co-referent that unites them as one T-unit instead of two separate ones.

However, these co-referents cause a problem when they are implicit Subjects. Given that I do not recognise implicit elements as Theme (4.3.3.1), the implicit Subjects in these cases are not going to be analysed. Instead, if clauses in the clause complex are paratactically connected sharing an implicit Subject, they are not considered as forming one T-unit but separate ones, each having its own explicit experiential element as its Theme.

4.2.3.4 Identifying T-unit in Japanese

A set of recognition criteria for the T-unit in Japanese is presented here as a summary of the discussion from the previous sections.

- A T-unit consists of either a clause simplex or a clause complex;
- Clauses that are paratactically related to each other in a clause complex form one T-unit each;
- The first suspensive form (*ren'yō* form) generally creates parataxis;
- However, when paratactic clauses in the clause complex share a referent, they form one T-unit in the same way as branching clauses do in English;
- If, however, the referent shared by paratactically related clauses is implicit, it does not unite the clauses as one T-unit and therefore, each T-unit has its own Theme;
- Clauses that are hypotactically related to each other in a clause complex form one T-unit together;
- The second suspensive form (*te*-form) generally creates hypotaxis;
- However, when each clause connected by the *te*-form has its own Subject, they are considered paratactically related to each other;
- Postpositional phrases are those whose meaning and form are firmly established as one, such as *ni tsuite* (Figure 4-4);
- There are verbs in the *te*-form that operate as adverbs, without any accompanying Participants or Circumstances;

After recognising the T-unit, Themes are identified. Before discussing how to identify Theme, however, the next section introduces another unit of analysis this project uses to investigate thematic progression of text.

4.2.4 Discourse unit to analyse thematic development

This section explains how I carry out the task of segmenting the news articles into discourse units, applying the understanding of the text structure presented in Chapter 3 Section 2.3. A hard news article comprises one textual nucleus and a set of satellites as discourse units. The purpose of segmenting the texts into the nucleus and satellites is to prepare them for the investigation of translation shifts in terms of the text structure; the relationships between the discourse units and additions/deletions/moves of informational content of news, and between the discourse units and thematic progression (Chapter 6).

4.2.4.1 Identifying the nucleus and satellites

Segmenting a text involves identifying discourse spans as a text nucleus and satellites. The text nucleus is identified first in the text as it is positioned at the beginning.

Identifying the end of nucleus is relatively straightforward in En-orig texts. It is because the break after the nucleus, where satellites start specifying the meanings of the nucleus, typically coincides with either the end of the first sentence, i.e. the lead, or that of the first paragraph.

However, it is occasionally difficult to decide whether to interpret a sentence as specification of the nucleus or an extension of it. An example of such a case is presented in 4-22, in which the underlined sentence poses a problem. It introduces the 'Sensible Tobacco Enforcement' bill which is not presented in the first paragraph. Thus, it is possible to treat it as newly added information to extend the meaning of the first paragraph as part of the nucleus.

However, the bill can also be interpreted as an elaboration/exemplification of NYC Council getting tough on tobacco, i.e. a satellite, which is how I annotate it. This decision is supported partly by the fact that the discourse span under discussion appears in a separate paragraph from the lead sentence that comprises the first paragraph. Very few texts in the En-orig corpus have the nucleus/lead appearing across separate paragraphs.

4-22

Nucleus	NYC Council gets tough on tobacco, approves raising purchase age to 21 New York (CNN) -- The New York City Council voted on Wednesday night to approve an anti-tobacco law that will raise the tobacco-purchasing age from 18 to 21.
Satellite 1	<u>In addition to the "Tobacco 21" bill, which includes electronic cigarettes, the council also approved a second bill, "Sensible Tobacco Enforcement."</u> It will prohibit discounts on tobacco products and increase enforcement on vendors who attempt to evade taxes.

En-orig-C24

The same type of difficulty, i.e. specification or extension of the text nucleus, is present in the Ja-orig data. In fact, there are more when compared to English and the paragraph boundary poses problems in Japanese, while it usually aids in English to decide the end of the nucleus. In 4-23, a sentence has been identified as a possible satellite (underlined). However, it is in the first paragraph together with the lead sentence.

4-23 リベンジポルノ:男逮捕 「写真まく」と元交際相手脅す

Revenge porn: Man arrested, '(I will) distribute photos' threatened ex-girlfriend

元交際相手の女性を「交際を続けなければ写真をばらまく」などと脅したとして警視庁青梅署は7日、東京都青梅市沢井2、無職、坂本龍馬容疑者(30)を強要未遂容疑で逮捕した。同署によると容疑を認めている。

(Taking as he) Threatened his former girlfriend saying 'unless (you) continue (our) relationship, (I would) distribute (your) photos', the police arrested Ryoma Sakamoto (30), Sawai 2, Oume-city, Tokyo, unemployed, for attempted coercion on 7th. According to the police (he) admits to the allegation.

Ja-orig-M28

The question is whether to interpret the underlined sentence as a satellite or part of the text nucleus. It functions more like a satellite to contextualise a situation after the arrest that the suspect admits to the allegation. The same type of sentence frequently appears in Japanese news articles and they are placed at various locations in the text, not very often in the nucleus, although it is not impossible as in this example. Thus, the function and usual location in the satellite block support the idea that the underlined sentence is a satellite.

However, the very argument that the given sentence can be at other locations in the text offers a counter argument; why it is in the first paragraph with the lead sentence when it can be outside. It is likely to be the result of the writer's or editor's choice, to treat it as part of the nucleus rather than as a satellite.

The claim that the text nucleus in the Japanese news articles tend to be more summary-like compared to English (Thomson *et al.* 2008) supports the decision to include the sentence in question in the nucleus. Providing a subsequent situation of the arrest in the nucleus to offer the reader with a general summary of events makes sense from the way the nucleus is perceived and defined in the Japanese news articles (Chapter 3 Section 2.3).

Thus, although it is the function of the discourse unit that is most important in segmenting texts, the paragraph boundary between the nucleus and the block of satellites' is given significance and the boundary will be observed as a stronger indicator to identify the end of the nucleus in the Japanese news articles. This does not contradict the principles of segmentation, however, since the textual location itself is a genre feature that is functional.

While the location indeed indicates a function, as is seen in the case of a nucleus above, there are also exceptions. The lead sentence in the nucleus is generally placed at the beginning of the body of the news article to fulfil its function to provide the reader with the news point. However, it can be delayed as in 4-24.

4-24 Alzheimer's disease ravages the brain, robbing its victims not only of their memories but often their ability to do things as basic as swallowing.

Now, a study of aging patients suggests its true toll may top half a million lives a year -- a figure that would put Alzheimer's just below heart disease and cancer on the list of America's top killers.

En-orig-C45

The most important news point in this article is presented in the second paragraph (underlined), not right at the beginning. Evans (2000) explains in his practical account of journalistic writing that this is a story-telling technique applied to news in order to give some structural variation. If the lead sentence is delayed, I will annotate everything up to and including the delayed lead sentence as the nucleus.

Once the nucleus is identified, the remaining part of the text is divided into individual satellites, each of which holds a certain semantic relationship with the nucleus. Example 4-25 shows the text nucleus and the first three satellites that immediately follow it. Each satellite refers back to the nucleus and specifies an aspect of information presented in it.

4-25

Nucleus	NYC Council gets tough on tobacco, approves raising purchase age to 21 New York (CNN) -- The New York City Council voted on Wednesday night to approve an anti-tobacco law that will raise the tobacco-purchasing age from 18 to 21.
Satellite 1	In addition to the "Tobacco 21" bill, which includes electronic cigarettes, the council also approved a second bill, "Sensible Tobacco Enforcement." It will prohibit discounts on tobacco products and increase enforcement on vendors who attempt to evade taxes.
Satellite 2	"Tobacco 21" will take effect 180 days after it is enacted, according to the council's news release.

Satellite 3	New York City has now become the largest city to have an age limit as high as 21. Needham, Massachusetts, raised the sale age to 21 in 2005, according to the New York City Department of Health.
-------------	---

En-orig-C24

A satellite frequently conflates with a paragraph as in 4-25 both in English and in Japanese (White 1998; Thomson 2005b; Washitake 2009) but not always. Difficulties in recognising the boundary between the adjacent satellites are often caused by cohesive ties amongst segments of the text. A long satellite is also problematic since satellites are typically fairly short, consisting of 1-3 sentences (White 2000). An example of a long satellite is presented below to illustrate how this long span is united as one by cohesive ties and why it is better to accept it as such.

4-26

nucleus	<p>Polio-like illness found in five California children</p> <p>(CNN) -- Doctors are looking for more information about a "polio-like syndrome" that has caused paralysis in a few children in California.</p>
Satellite	<p>¹ <u>Enterovirus 68</u> was first identified in a California lab in 1962, after four children came down with a severe respiratory illness. Between 1970 and 2005, only 26 cases of enterovirus 68 in the United States were reported to the CDC. Since 2000, the government agency has kept a closer watch and has seen 47 cases, Oberste said. Outbreaks have occurred over the years in Asia and Europe, but it's still one of the rarest types of enterovirus.</p> <p>² <u>More common -- and more concerning</u> to health officials -- is <u>enterovirus 71</u>, which was discovered by <u>the same</u> California lab in 1969, Oberste said. Enterovirus 71 is usually associated with severe neurological issues, including aseptic meningitis, polio-like paralysis and encephalitis. Waubant is not sure if the samples from <u>this</u> latest group of patients were tested for enterovirus 71.</p> <p>³ According to a CDC report, several outbreaks of paralysis caused by <u>enterovirus 71</u> were seen in Europe in the 1960s and 1970s. In the late 1990s and early 2000s, fatal encephalitis was a big problem in Malaysia and Taiwan.</p> <p>⁴ "Ever since <u>then</u>, the virus has circulated in high levels in Southeast Asia," Oberste said.</p>

	<p>⁵ In recent years, <u>the outbreaks</u> have spread to Australia; a cluster of cases near Sydney drew media attention in 2013.</p> <p>⁶ Between 1983 and 2005, 270 cases of <u>enterovirus_71</u> were reported in the United States. But none has resulted in a larger outbreak, despite the virus's infectious nature.</p> <p>⁷ "<u>That's</u> the really odd thing," Oberste said. "We see cases from time to time in the United States. Occasionally they'll be severe. Basically it's identical to what's circulating in Asia ... but it doesn't cause the same big outbreak in disease. And we really don't know why."</p>
--	--

En-orig-C43

Although smaller segments can be recognised as (1-2), (3-5) and (6-7) in 4-26, the seven paragraphs are perceived as one satellite. The subject matter is enterovirus 68 and 71 throughout this text span and the cohesive ties unite it as one such as reference chains (Halliday & Matthiessen 2004: 535) with 'enterovirus' or demonstrative and comparative references. If this satellite is broken into smaller units, they will not function to specify any element that appears in the nucleus. Therefore, it is reasonable to accept this segment as one unit that elaborates 'polio-like illness' which appears in the nucleus. The decision to take this long unit as one instead of breaking it into smaller segments is because I am interested in the function the unit has in relation to the text nucleus (rather than how this unit is rhetorically structured), which is to investigate the relationship between the Theme realisation and text structure. The relationship of each satellite with the nucleus is essential as that is how the news discourse is conventionalised.

A useful strategy in distinguishing a satellite is to observe if a text span functions on its own to specify the nucleus, and also if it can move around without compromising the coherence. Functionality as a satellite and the unity created by referential chains are both important, but the former is prioritised. Therefore, simply because there is a word such as 'meanwhile' at the beginning of a paragraph, that itself is not taken as an exclusive reason to connect the segments before and after that. If the segment after 'meanwhile' functions as a self-contained specifier of the nucleus, it is treated as an individual satellite.

A text span's self-containedness and its function as an individual specifier are applied to identify a satellite that involves quotations. Quotations can be independent satellites but they are often better combined with the previous sentences, which in a way provides a context for the speech as in 4-27.

4-27 During the main part of the trial, several witnesses testified that Trujillo was prone to unprovoked violence when drinking.

"I have problems expressing myself," Trujillo told jurors after taking the stand in the penalty phase of the trial. En-orig-R63

So far, I have discussed general principles of deciding the boundary between the nucleus and the block of satellites, i.e. recognising where the nucleus ends, and the boundaries between satellites, i.e. recognising where one satellite ends and another starts. However, there are several irregular texts in the corpus – texts that do not have the nucleus and satellite structure of typical hard news articles. The beginning of such an example is presented in 4-28. The headline is underlined.

4-28 Hope for survivors of landslide dims as death toll rises as high as 24

Darrington, Washington (CNN) -- Brenda Neal was still at the firehouse at midnight, watching as rescuers caked with mud returned from the search for survivors of a massive landslide in rural Washington state.

But they had no answers for her about her missing husband, Steven.

There was despair on their faces, she said. En-orig-C56

The first few paragraphs presented in 4-28 do not read like typical news in that they neither offer the news focus nor a summary of news, which is often done by presenting the same information that appears in the headline. In 4-28 the number of dead presented in the headline finally appears in the 6th paragraph, three paragraphs further down. In cases like this, it doesn't seem that the delayed lead device is at work, but rather simply that the structure is not that of nucleus and satellite. A tag of 'not the nucleus and satellite structure' is used to annotate these.

The above issue is mainly with English, but a different irregularity appears with the Ja-orig data. There are several articles that have an additional section after the news itself ends after the writer's name. They are mostly extra information, such as excerpts of an interview conducted with someone who appears in the news article, a list of examples that are relevant to the news, or an extensive description of the background of the news. The additional parts are neither considered to be the nucleus nor the satellite. They may or may not have the nucleus and satellite structure of its own, but in either case they are labelled as one segment as an additional section.

When all the satellites are identified and boundaries between adjacent satellites are established, the segmentation of text is complete for this project.

Although a more sensitive annotation of satellites classifying their relationships to the text nucleus is possible (Chapter 3 Section 2.3), it is not conducted here since this lies outside the scope of this investigation. In addition, although it is possible that the detailed analysis might shed extra light on the relationship between the Theme and text structure, there is also concern it may not provide accurate results to be profitable.

This concern is based on two issues in correctly determining the relationship that a satellite holds with the text nucleus. Firstly, it is not always straightforward to recognise the semantic relationship of a satellite with the nucleus as this is not always explicitly realised by the lexicogrammar. The major relationships are Elaboration, Cause & Effect, Justification, Contextualisation, and Appraisal (Chapter 3 Section 2.3). Satellites in example 4-25 above are regarded as Elaboration but they scarcely offer clear grammatical signals as to why they are interpreted as such.

The second type problem is that a satellite's semantic relationship with the nucleus is not always a distinctive one but can be mixed, e.g. Elaboration and Cause & Effect.

4-29

Nucleus	Two boys arrested, accused of putting poison in teacher's water bottle New York (CNN) -- Two New York City fourth-graders were arrested Tuesday after allegedly putting poison in a teacher's water bottle, police said.
Satellite 1	<u>1The two boys, ages 9 and 12, allegedly poisoned the teacher's water</u> <u>Monday afternoon, 2causing her to feel nauseated the next day. 3She was</u> <u>treated by her doctor and is now in good condition, said Sgt. John Buthorn</u> <u>of the New York City Police Department.</u>

En-orig-C77

The satellite in 4-29 elaborates on the nucleus, stating an effect that the incident presented in the nucleus has caused and also contextualising one of the participants' conditions after the incident. This satellite may be divided into three parts (1-3 in the example) according to the functions they fulfil. The first part is Elaboration, the second Cause & Effect and the third Contextualisation. Despite the clear division of functions, this section is established as one satellite because the minimum rank for a discourse unit is the T-unit in this project. Moreover, the meaning of the third part is better understood when placed right after the second part rather than away from it. The third part may be regarded as a satellite on its own. However, without the second part 'feeling

nauseated', there seems to be a gap between 'getting poison in her water' and 'being treated by her doctor'.

Theoretically, this could be solved by adopting the dimensions approach, considering the degree of the semantic relationships involved, instead of deciding on a particular one over another. Practically, however, this is difficult to reflect on the annotation in the UAM CorpusTool. The tool allows multiple-layered annotation but specifying several dimensions at the same time is apparently not supported. Without specific linguistic features that can be used as a guide, accuracy and consistency of annotation are potentially compromised. The consequences of these are unreliable analysis and results.

Hence, this project does not extend the annotation task to assign semantic relationships to satellites that they hold with the nucleus. The two problems of satellites' relationships to the nucleus being either implicit or mixed will pose a problem in terms of the accuracy and consistency of annotation, which in turn influences the reliability of the analysis results. Such a full annotation of this level, whilst valuable in itself, is outside the scope of this study and the research questions can still be addressed.

4.2.4.2 Annotating the nucleus and satellites

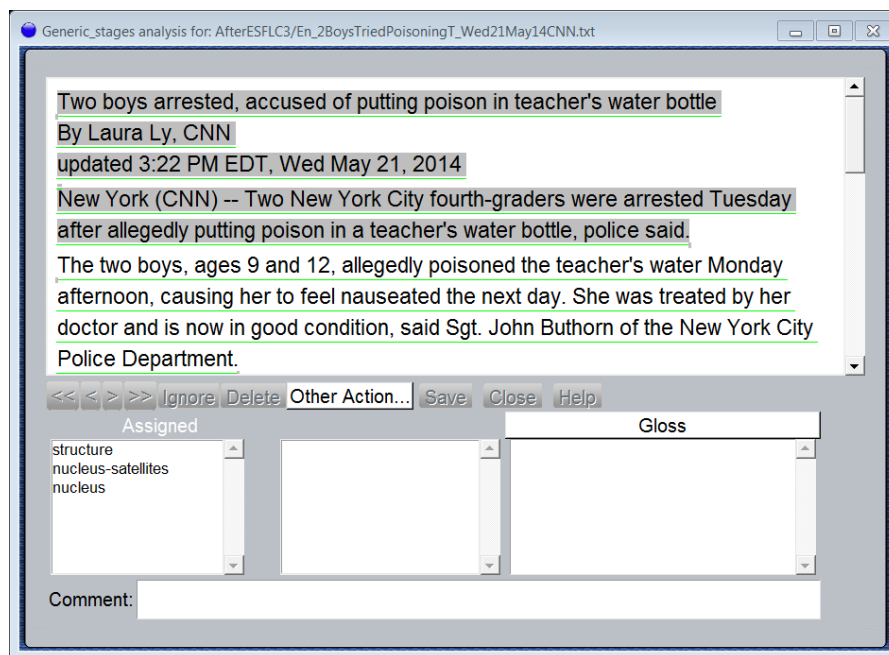
The general principles for identifying the nucleus are summarised below. Principles 1 to 7 are from White (1998, 2000), while 8 is from Evans (2000) and 9 is my own.

1. Each text has one nucleus;
2. The nucleus contains the headline and the lead;
3. The headline and the lead in the nucleus generally present the same informational contents, thus the lead ends after presenting what is introduced in the headline;
4. The nucleus normally ends when the lead ends; the lead typically contains one sentence although it could stretch to more than one sentence;
5. The end of the nucleus/lead often coincides with the end of the first paragraph;
6. The experiential Theme of the first satellite can be used as a signal of the beginning of the specification because it refers back to an element introduced in the nucleus, making the Theme as Given information;
7. The end of the nucleus/lead normally coincides with the sentence ending, and not the middle;

8. The lead sentence can be delayed, instead of appearing at the beginning of the text;
9. A news article may not have the nucleus-satellite structure.

The figure below is a screenshot of the annotation task of nucleus conducted on the UAM CorpusTool. The highlighted part is the nucleus while the underlined part is the first satellite. The text span assigned as the nucleus meets the criteria from the general principles described above.

Figure 4-5: Annotation example of the nucleus on the UAM CorpusTool



En-orig-C77

Bylines giving the name of the reporter, date and place are placed between the headline and the lead in the news articles published by both CNN and Reuters, as seen in the above figure. Technically, they are not part of the nucleus but I take the practice of not excluding them from the nucleus in order to keep the nucleus as one text span. They do not affect the analysis because they are disregarded.

The general principles of identifying the spans of satellites are presented below. Principles 1 to 4 are either directly or indirectly borrowed from White (1998, 2000), while 5 to 7 are my additions.

1. A satellite is self-contained and functions on its own to specify the meanings of the nucleus;
2. Each satellite has a particular relationship, i.e. function of specification, with the nucleus;
3. A satellite can be moved around without compromising the coherence;

4. A satellite is normally short with one to three sentences, often corresponding with a paragraph but it can be long, involving several paragraphs;
5. The end of a satellite often coincides with that of a paragraph's although it is not impossible to have a boundary in the middle of the paragraph;
6. The minimum size for a satellite is the T-unit;
7. Quotations can be independent satellites but they are often combined with the previous sentences, which in a way provides a context of situation for the quotations.

4.3 Identifying Theme

4.3.1 Boundary of Theme and Rheme

As the choice of analytical unit depends on the purpose of the analyst, so does how far the Theme extends in the unit. According to Halliday, the Theme of a clause contains only one experiential element, which is either Participant, Circumstance or Process, and ends with it (Halliday & Matthiessen 2004: 79). Although their discussion is based on English, it applies to Japanese as well as other languages.

If a circumstantial element comes at the beginning of the T-unit either with or without a textual and/or interpersonal Theme in front of it, the line between Theme and Rheme is drawn after the circumstantial element and the Subject is not included in the Theme. According to this approach just one experiential element exhausts the thematic potential as in 4-30 and 4-31.

4-30 In July, Selna gave final approval to a settlement valued at more than \$1.6 billion to resolve economic-loss claims, ... En-orig-R30

4-31

景気回復への手ごたえから	企業の採用意欲は	高まっている。
<i>Keiki-kaifuku E NO tegotae KARA</i>	<i>kigyō NO saiyō-iyoku WA</i>	<i>takamatteiru</i>
Economic recovery E NO feel KARA	Company NO recruit willingness WA	rise-ASP
Because of the feel that the economy is recovering, companies' willingness to recruit is rising.		
Circumstance	Medium	Process
Theme	Rheme	

Ja-orig-Y28

There are discussions that push the boundary of Theme up to and including the unmarked Subject Theme or everything that precedes the verb (Berry 1995). If this approach is taken, the Theme in 4-30 will include the unmarked Subject 'Selna' and '企業の採用意欲は [companies' willingness to recruit]' in 4-31. Thus, it places a weight on the Subject as the Subject Theme shows the continuity of the Participants in the text. Thompson (2004: 173) states it is meaningful to include both marked circumstantial Theme and unmarked Subject Theme because they are both important for the text to develop given that they fulfil different textual functions.

However, the present study follows Halliday & Matthiessen (2004: 79) and adopts the approach of including only one experiential element in Theme. Given that the aim of this study is to compare Theme realisation between English and Japanese news, and in particular that Japanese often makes use of Subject ellipsis, this approach is better suited as it enables the comparison to be straightforward and simple.

Moreover, the label of 'experiential Theme' is employed instead of 'topical Theme' to avoid confusion with the Topic from the Topic-Comment structure, in particular in Japanese, as in Thompson (2004) and Thomson (2013).

As for how Theme is realised in Japanese, as discussed in 4.2.3.1, I choose the first position rather than the particle *wa*, similar to Thomson (1998) and unlike Teruya (2006). If more than one experiential element marked by *wa* appears in initial position, only the first is taken as the Theme. If the unmarked Subject is implicit, an explicit experiential element is Theme. Theme realisation in Japanese will be discussed further in 4.3.3.

If a dependent clause (minor as well) in the T-unit is placed in initial position followed by an independent one, the boundary of Theme comes after the dependent clause, which is taken as Theme for the whole T-unit.

4-32 If they meet the deadline, the legislation will then head to Governor John Hickenlooper, a Democrat, who is expected to sign the bill into law.

En-orig-R73

The same applies to Japanese and it will be discussed more in detail in 4.3.3.5.

4.3.2 Theme in English

Thematisation relates to the word order in the clause and there are natural candidates for the Theme, i.e. unmarked Theme. Theme in English has a close link with the MOOD system (Halliday & Matthiessen 2004: 71) and the natural unmarked candidates for the Theme are:

1) the Subject in declarative clauses;

The duke has given my aunt that teapot. (unmarked)

(*ibid.* 66)

2) the Finite verbal operator in yes/no interrogative clauses;

Has the duke given my aunt that teapot? (unmarked)

3) Wh-element in Wh-question clauses;

Who has given my aunt that teapot? (unmarked)

4) and the Predicator in imperative clauses.

Give my aunt that teapot. (unmarked)

Unmarked Themes are naturally Theme because of the word order. One way to move away from the naturally realised Theme and manipulate which constituent is realised as Theme is passivisation as in 5) below. Selecting an uncommon element to be marked Theme is another way to change the word order as in 6) below.

5) My aunt has been given that teapot by the duke. (unmarked Theme)

6) That teapot the duke has given to my aunt. (marked Theme)

(*ibid* 66)

An additional example of marked Theme I include in the system of THEME for English is another verbal Predicator. It is different from the Predicator Theme in imperative clauses, which is an unmarked choice. The marked verbal

Predicator occurs in the projecting clause that is placed after a projected clause in which the verbal Predicator (underlined) is positioned before the Subject as below.

4-33 “This is the final piece to our pot puzzle,” said Singer, a Democrat.

En-orig-R73

In addition to passivisation and marked Theme, special thematic structures are resources that allow unnatural or uncommon elements to be made Theme so that a clause fits in the text and keeps the flow of discourse. Or they allow an element to be placed towards the end of the clause so that it has a focus of information as New in Rheme. In short, these structures concern the order of introducing an element in a clause as either New or Given information.

Thompson (2004) introduces five special thematic structures as ‘enhanced Theme’. This study adopts Thompson’s position and recognises all of them as special thematic devices because they are usefully treated separately when analysing translation from/to Japanese. Each type is briefly examined below and reasons why they are thought to be useful for this study are offered.

The first enhanced Theme is ‘predicated Theme, traditionally called cleft sentence’ (Thompson 2004: 151; Halliday & Matthiessen 2004: 95). A single element in a clause is selected and given prominence as Theme.

4-34-A but it was a group of hunters who heard Penaflor’s cries for help and eventually found him in a canyon...

En-orig-R11

4-34-B a group of hunters heard Penaflor’s cries for help and ...

Comparing the two versions above makes it clear that an emphasis is placed on the predicated Theme (double underlined), which is an Identified in a relational clause. This structure allows Theme to conflate with New by delaying its introduction (Thompson 2004: 152). Although Japanese does not have a specific structure like this, the same device exists, which makes use of a relational clause with nominalisation. This has the same function of controlling the position of New or Given information in a clause and identifying this structure in both English and Japanese is useful in discussing translation shifts.

Next is ‘thematic equative’ (*ibid.* 149-150), which groups two or more elements in the clause to form a single constituent and manipulates the clause structure as an identifying clause, creating the Theme=Rheme relationship (the Wh-clause is the Value). This is traditionally called pseudo-cleft.

4-35 But what we've seen over the last several years is their capacities
continue to develop and expand... En-orig-R3

The starting point in a thematic equative is often a question that the speaker/writer imagines the hearer/reader might want to ask at this stage. Accordingly, this pattern is used to indicate a transition or staging in the text (*ibid.* 150). Again Japanese does not have a special structure for this meaning but a relational clause with nominalisation will produce the same meaning. The interesting thing is that the clause structure in Japanese for this is the same as it is for predicated Theme. Whether or not Japanese somehow distinguishes the two types is observed in Chapter 6.

Next is 'preposed Theme' (*ibid.* 153), a nominal element that is picked out as a separate constituent and then substituted by a pronoun in the appropriate place in the following clause.

4-36 The parent, he just broke down. En-orig-C75

In order to mean this in Japanese, topicalisation will be effected by moving an element towards the clause initial position and marking it by the particle 'wa'. As a similar device is available in Japanese, it is worthwhile recognising it as a special structure in English.

Comment is similar to the predicated Theme in terms of the structure but the comment part is not a meaning component of the following clause. The speaker/writer starts the message with their Comment on the value or validity of what they are about to say. Halliday & Matthiessen (2004: 97) recognise only 'it' as Theme and do not consider this structure as a Thematic device, whereas Thompson (2004: 153) argues that it makes more sense to include the speaker's comment in the Theme, otherwise it obscures the method of development in the text.

4-37 It's important to understand that we are going to have to have a policy
discussion that is bigger than any individual intelligence agency of either
Europe or the United States. En-orig-R3

There is no strong reason to identify this structure in English from the point of view of how Japanese makes this kind of meaning, as the empty 'it' does not exist in Japanese. However, considering that it may prove to be useful in the activity of translation to distinguish 'it' in the Comment structure from that in a clause, such as 'it rained yesterday', this study follows Thompson in including the comment part in the comment structure, in contrast to Halliday & Matthiessen (2004), who take only 'it' as Theme.

The final enhanced Theme is the 'there' in an existential clause, which is the Subject but fulfils no experiential function. 'There' is a dummy Subject that indicates something is about to be introduced as an existent, which is New information. Halliday & Matthiessen (2004: 257) call it the 'presentative construction' and identify only 'there' as Theme. Thompson (2004: 161), on the other hand, includes the existential process in the Theme as a meaningful experiential component.

4-38 There have been roughly 200 proposed class actions and more than
500 individual cases filed against Toyota ... En-orig-R4

Existential clauses in Japanese take the form of a relational clause and, depending on the information status of the Existent, the word order and the particle used to mark the Existent change. Whether or not to include the existential process in Theme in English is not a significant issue from the viewpoint of Japanese. However, taking Thompson's point that the existential process is the first experiential element, which is obligatory in Theme, 'there + existential process' is considered to be Theme in the existential clause.

The system network of THEME in English is presented in 4.3.4 alongside one in Japanese.

4.3.3 Theme in Japanese

Theme is realised at the clause initial position in Japanese as in English. I have established that the particle *wa* is not considered as a Theme marker (4.2.3.1) though it is at times taken as such (Tatsuki 2000; Teruya 2006).

Japanese is said to have a 'free word order' (Martin 1975: 35). As grammatical relationship is marked by postpositions, the linear organization of the clause is less significant than in languages such as English (Teruya 2006: 87). However, 'it is possible to identify a default experiential sequence' (*ibid.* 87), and if the order is changed by putting forward an element to the clause-initial position, that moved element is given a prominence. Therefore, it is reasonable to assume that the clause-initial position plays a key part in Theme realisation in Japanese as Thomson demonstrates (1998). Teruya (2006) also recognises the significance of the initial position although he gives prominence to the particle *wa* too.

In what follows I will explain principles for annotating Theme in this project. Firstly I present the systems of THEME by Teruya (2006) and Thomson (2001). The system I created for this project is adapted partly from each. With regard to several decisions I make in identifying Theme, I refer to them and

describe similarities and differences between my position and theirs, as well as reasons for my decisions.

Figure 4-6: The most general system of THEME (Teruya 2006: 133)

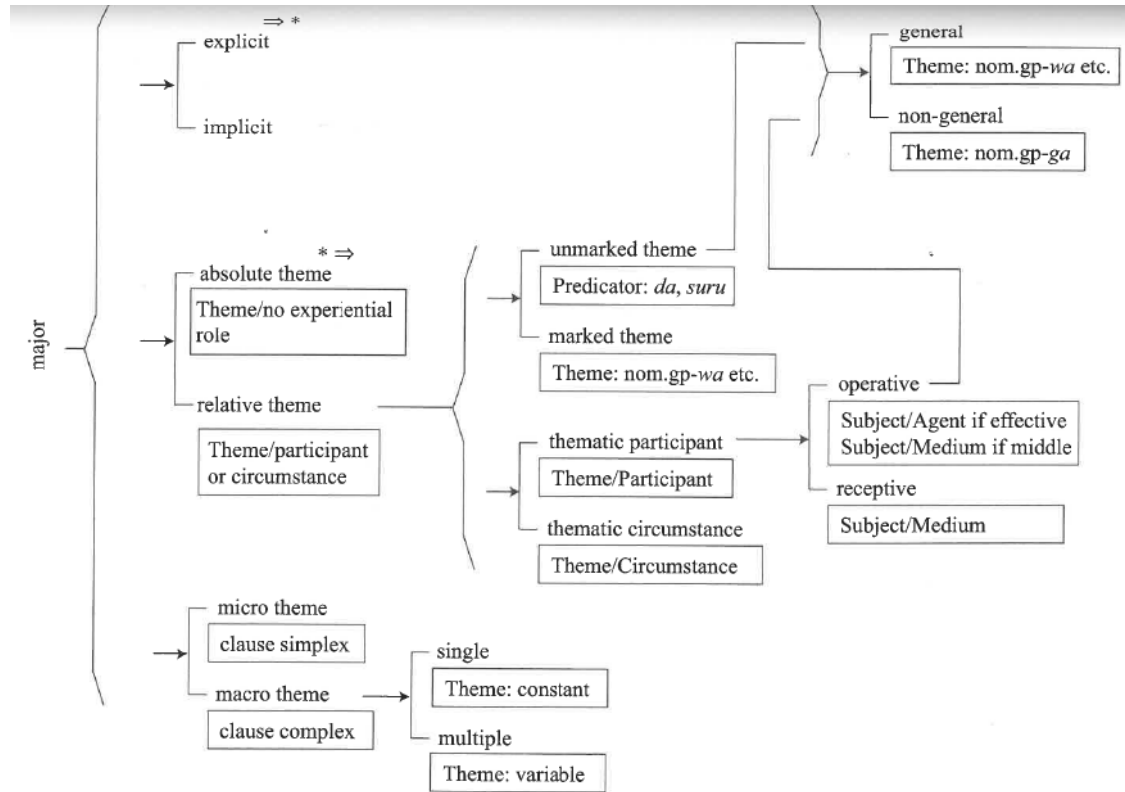


Figure 4-7: System network of THEME (Thomson 2001: 119)



4.3.3.1 Implicit or explicit Theme

Both Thomson (2001) and Teruya (2006) include the distinction between implicit and explicit Themes in their systems of THEME. This means they allow Themes to be implicit, as a way to handle ellipsis.

Ellipsis is a type of cohesive relation, an omission of an item; when something is left unsaid but understood nevertheless (Halliday & Hasan 1976: 142). It is essentially the same process as substitution. Substitution is the term for replacing an item with another and ellipsis for replacing an item with nothing (*ibid.* 88). There is a similarity in the way substitution in English and ellipsis in Japanese work.

Thomson explains that implicit Themes consist of only ideational elements and predominantly Participants, which become implicit because they are redundant information (2001: 125). Focusing on how Theme can be implicit, Teruya states,

... while Theme is an integral part of the clause as message, its explicitness depends on its thematic continuity. Theme recedes into background, or is ellipped, until such a time when a given Theme is reset by new one ^{sic}.

(Teruya 2006: 51-52)

When discussing Korean, Kim (2011) offers a similar view on the explicitness of Theme. She claims that sometimes the elements that appear in clause initial position do not function as orienter of the message, and that allowing implicit Theme is more appropriate to examine the method of development of texts.

I agree with Teruya and Kim that being implicit does not mean nothing exists. However, I differ from them in that I assign the status of Theme to elements that appear overtly in initial position in the T-unit, and do not allow implicit Themes. What is important in my project is to distinguish explicitness or implicitness so that the similarities and differences of Theme realisation between English and Japanese are clearly displayed. Allowing implicit Theme may be one way. However, it is more useful to show what appears in initial position as Theme rather than simply indicating something is missing, in particular in relation to translation practice or training.

Halliday & Matthiessen stress the importance of distinguishing the realised and implicit with regard to conjunction.

The presence or absence of explicit conjunction is one of the principal variables in English discourse, both as between registers and as between texts in the same register; this variation is obscured if we assume conjunction where it is not expressed.

(Halliday & Matthiessen 2004: 549)

For Spanish, which has a tendency to drop Subject pronouns, it is proposed that the Process is the Theme when the Subject is ellipped (McCabe 1999; Munday 2000; Aruz 2010). Their rationale is Theme should be an explicit element. Moreover, they claim the Subject may be located after the Process when it is not ellipped, which means there is no reason why Subject, whether it is implicit or explicit, should be prioritised as Theme over the Process. They also state that even when the Subject pronoun is ellipped, tracking the Participant's identification is made possible by verbal inflections.

Tracking Participants' identities is an important feature of the textual metafunction, and in relation to reference, Halliday & Matthiessen claim that 'the unmarked anaphoric strategy is to use the pronoun' (2004: 554).

This explanation is about reference and not ellipsis, but there is a 'strong tendency for reference items to be thematic' (*ibid.* 554) so the Theme and pronouns in English are tied closely. In contrast, in Japanese there is a tendency of making use of ellipsis where English employs pronouns. By leaving out parts of a structure when they can be presumed, ellipsis 'indicates continuity, allowing speaker and addressee to focus on what is contrastive' (*ibid.* 535). Teruya states;

Once it has been established as Given/Theme it often continues to serve as Given/Theme in subsequent clauses and is left implicit (Thomson 1998, Tsukada 2000) as an implication of this continuity until the identity of the Theme is 'reset' by an explicit Theme, which may be another given Theme or a new topical Theme. ...whether it is Given or New will contribute to cohesion in text.

(Teruya 2006: 121)

His explanation indicates that implicit Themes occur when they are Given information. Thomson too calls them 'co-referential ellipsis' (2001: 73). Their accounts raise the question of whether or not it is reasonable to assume it is truly Theme that is implicit. Aruz (2010) claims what is implicit in Spanish is not Theme but Given.

Since the way Theme is realised can vary in different languages (Rose 2001), how closely it is linked with other systems such as INFORMATION or REFERENCE may also vary. Teruya explains, '(w)hen the topical Theme of a clause is implicit because it is referentially recoverable, it is also Given and the clause is made up of the New element only' (2006: 118) suggests that he considers the systems of THEME and INFORMATION close.

However, given that this project takes the separating approach (Fries 1983), I maintain that there will be a Theme even if there is not Given information. Example 4-39 illustrates how I annotate Theme when the Subject as Given information is ellipsed. Due to Subject ellipsis a dependent clause is the Theme in 4-39-2 and so is a Circumstance in 4-39-3. Example 4-39 presents a preceding T-unit 4-39-1 before the two successive clauses with Subject ellipsis to indicate the ellipsed Subject in them.

4-39-1

このため、	チームは	事前に	CT のデータを元に、	女性と夫それぞれの・・・を	3D プリンターで	作成。
<i>Kono tame</i>	<i>chīmu WA</i>	<i>jizen NI</i>	<i>CT NO dēta O motoni</i>	<i>josei TO otto sorezore NO ... O</i>	<i>3D purintā DE</i>	<i>saku sei</i>
Due to this	team WA	before -hand NI	CT NO data O base NI	woman TO husband each NO ... O	3D printer DE	create
Theme (Actor)		Rheme				

4-39-2: clause Theme (Subject ellipted)

これらを	使って	綿密な手術手順の検討などを	行い、
<i>Korera O</i>	<i>tsukatte</i>	<i>menmitsuna shujutsu-tejun NO kentō nado O</i>	<i>okonai</i>
These O	use-SUSP	detailed operation procedure NO discussion, etc. O	do-SUSP
Theme (clause)		Rheme	

4-39-3: Circumstance Theme (Subject ellipted)

今年 3 月下旬に	移植を	行った。
<i>Kotoshi, 3-gatsu gejun NI</i>	<i>ishoku O</i>	<i>okonatta</i>
This year March late NI	transplant O	do-pst
Theme (Circumstance)		Rheme

Ja-orig-Y72

In both 4-39-2 and 4-39-3 the Subject is co-referential チームは [the team] and ellpted and the first experiential elements are realised as Theme. Explicit Theme strategy determines the way the T-unit is identified, i.e. this clause complex will be interpreted as two paratactically related clauses. If implicit theme were allowed, they would have formed one T-unit with the implicit shared Subject.

Subject ellipsis is a major reason for Teruya and Thomson to employ the implicit Theme. However, it is not the only situation and there is another type of 'omission' which leads them to adopt implicit Themes. This is not ellipsis

since the presupposed information is not an item whereby a grammatically appropriate form fills the empty slot in the structure (Halliday & Hasan 1976: 145). Example 4-40-2 illustrates such a case in which a situation or what is expressed in the previous Rheme in 4-40-1 is presupposed.

4-40-1

維新の会と結いの合流協議を巡っては、	石原氏が	//政策合意に「自主憲法制定」を盛り込むよう//	主張し、	結いの江田代表が	反発。
<i>Ishin-no-kai TO Yui NO gōryūkyōgi O-megutte WA</i>	<i>Ishihara-shi GA</i>	// <i>seisaku-gōi NI 'jishukenpō-seitei' O morikomu yō//</i>	<i>shuchō shi</i>	<i>Yui no Eda-daihyō GA</i>	<i>hampa tsu</i>
<i>Ishin-no-kai party & Yui party NO merger O concerning</i>	Mr Ishihara GA	//policy agreement NI 'own constitution establish' O include-PROJ//	insist-SUSP	Yui NO Eda leader GA	oppose
Concerning the discussion on the merger of <i>Ishin</i> and <i>Yui</i> parties, Mr Ishihara insisted they included 'establishing own constitution' in the policy agreement (but) Mr Eda opposed.					
Circumstance	Sayer	Projected	Process	Actor	Process
1 clause: Projecting			Projecting	2 clause	
Theme	Rheme				

4-40-2

[[合流協議が進まない]]状況に	陥っていた。
[[<i>Gōryūkyōgi GA susumanai</i>]] <i>jyōkyō NI</i>	<i>ochiitteita</i>
[[Merger talk GA forward-neg]] situation NI	fall-ASP-pst
(It) fell into a situation that the merger talks would not go forward.	
Attribute	Process
Theme	Rheme

J-orig-Y77

Assigning the status of Theme to the first explicit element as creating a local context in 4-40-2 is not difficult. However, it leaves the Rheme very short; in this case it is only Process. Given that the Rheme is where the text develops

from the point of departure and indicates a global concern, the goals of the text or a text portion (Fries 1994), it may be questionable. However, it includes core information such as polarity, and I accept short Rhemes as an inevitable consequence of not allowing implicit Themes.

4.3.3.2 Simple and multiple Theme

Themes can be simple or multiple. An experiential element is obligatory in Theme and if it stands on its own as Theme, that makes a simple Theme. Multiple Themes consist of an experiential element and textual and/or interpersonal elements.

Both Teruya and Thomson acknowledge these. However, Teruya does not include them explicitly as options in the system of THEME but he states everything that precedes the element marked by *wa* is unmarked Theme (2006: 96), which includes textual and interpersonal Themes. Thomson's system includes them but she refers to simple Theme as 'single' Theme (2001: 119). The present project incorporates them in the system as an important aspect of Theme realisation but keeps the labels of simple and multiple Theme, rather than single and multiple that Thomson employs.

Thomson's single Themes can be textual or interpersonal elements if they are realised on their own without an ideational Theme. I do not adopt them because I do not allow implicit Themes and assign the thematic status to explicit elements only.

A category of her multiple Theme, on the other hand, presents a useful type for the current project, that is, univariate Theme. A typical multiple Theme contains different constituents serving different metafunctional roles such as a compulsory experiential Theme plus either textual Theme or interpersonal Theme, or both. A univariate multiple Theme, Thomson explains, 'consists of two circumstantial Themes, that is, two constituents serving the same function' (2013: 119).

This category may be questionable given that the Theme in this project is defined as containing only one experiential element. Nevertheless, it occurs relatively frequently in the corpus and as Thomson notes, is likely to be a feature of Japanese news articles (2013: 119). Often employed 'to set the scene of the action' (*ibid.* 119), it will be a usefully recognised linguistic feature with respect to translator training.

As for the combination of Circumstances in univariate Themes, I will limit it to two types; that of the temporal and spatial Locations as in example 4-41, as Thomson does, and that of the Location and Extent. The latter is included because of the relative similarity of the Extent with Location.

4-41

新潟県新発田市真野原の雑木林で	7日夕、	… 女性の遺体が	見つかり、
<i>Niigaga-ken Shinhotta-shi Manohara NO zōkibayashi DE</i>	<i>7-ka yū</i>	<i>... josei NO itai GA</i>	<i>mitsukari</i>
Niigata prefecture, Shinhotta-city, Manohara NO wood DE	7 th evening	woman NO body GA	find-psv- SUSP
A body of woman was found in the evening of the 7 th in the wood in Shinhotta-city, Manohara, Niigata prefecture			
Circ.: spatial Location	Circ.: temp.	Medium	Process
univariate Theme:		Rheme	

Ja-orig-M59

Although not as common as that of temporal and spatial Locations, other combinations do occur. However, I do not recognise them as univariate Themes. Although both are Circumstances, they do not appear to fulfil the function of Theme as one with equal weight unlike the univariate of temporal and spatial Locations. Instead, the Circumstance in initial position takes the full responsibility and the second falls in the Rheme. Hence, in 4-42-A '宇都宮市では [in Utsunomiya city]' (Location) is the Theme while the subsequent 'パソコン約 1500 台について [about the 1500 PCs]' (Matter) falls into the region of Rheme.

4-42-A

宇都宮市では	…パソコン約 1500 台 について	//「…ウイルス対策ソフトで 対応する」と//	判断し、
<i>Utsunomiya-shi</i> <i>DE WA</i>	<i>… pasokon yaku</i> <i>1500-dai NI-tsuite</i>	//‘… uirusu-taisaku-sofuto de taiōsuru’ TO//	<i>handanshi</i>
Utsunomiya city DE WA	PC about 1500 NI- tsuite	//‘… anti-virus software DE deal with’-PROJ//	judge-SUSP
In Utsunomiya city, (they) judged they would deal (the situation) with the anti-virus software about 1500 PCs,			
Circ.: Location	Circ.: Matter	Projected	Process
Theme	Rheme		

Ja-orig-Y60

In 4-42-A it may appear that the particle *wa* that marks the first Circumstance may have a part in the decision. However, in this case it does not. If the two Circumstances are reordered, the first without *wa* will still stand as Theme on its own as in 4-42-B. This is because *wa* in 4-42 indicates a general contrast between Utsunomiya city and other cities which may take a different approach.

4-42-B

…パソコン約 1500 台に ついて	宇都宮市では	//「…ウイルス対策ソフトで 対応する」と//	判断し、
<i>… pasokon yaku 1500- dai NI-tsuite</i>	<i>Utsunomiya-shi</i> <i>DE WA</i>	//‘… uirusu-taisaku-sofuto de taiōsuru’ TO//	<i>handanshi</i>
About 1500 PCs, (they) judged they would deal (with the situation) with the anti-virus software in Utsunomiya city,			
Circ.: Matter	Circ.: Location	Projected	Process
Theme	Rheme		

Ja-orig-Y60 (reordered version)

Let us clarify here the two functions *wa* has; to mark the Topic on one hand and to create a contrast on the other (Teramura 1991; Noda 1996 among others). They are related to, rather than distinct from, each other. In a typical situation to create a contrast there will be more than one explicit item to be contrasted while it is also possible for one item to be present and the other(s)

implicit as in 4-42-A. When the sense of contrast is so reduced that it becomes as one item against all the other imaginable items possible, the item marked by *wa* is the Topic. *Wa*'s function, therefore, is to show paradigmatic relations between the given item and another specific one when indicating a contrast or all the other possible items that are not explicitly present when it is marking the Topic.

Returning to the example in question, if two Circumstances appear in initial position, unless they are two Locations or a Location and an Extent, only the first will be recognised as Theme. The particle *wa*, marking the first Circumstance, does not affect the assignment of Theme. However, if it is the second Circumstance that is marked by *wa*, a consideration is given as to whether it marks the second element only to create a contrast or it marks both Circumstances to topicalise them together. If it is the latter, though unlikely, then the univariate Theme tag will be used.

In the following example, however, the particle *wa* needs to be recognised as influencing the boundary of Theme and Rheme in between the univariate Theme. This is because in this case *wa* marks the Topic which creates a break between the Topic, the first Circumstance, and the rest of the T-unit, which is the Comment. This break overrides the functional univariateness of the two Locations.

4-43

コロラド州では、	今年 1 月に	販売が	開始された。
<i>Kororado-shū DE WA</i>	<i>kotoshi 1-gatsu NI</i>	<i>hanbai GA</i>	<i>kaishisareta</i>
Colorado State DE <u>WA</u>	this year January NI	sales GA	start-psv-pst
In Colorado the sales was started in January this year.			
Circ.: spatial Location	Circ.: temp. Location	Medium	Process
Theme=Topic	Rheme=Comment		

Ja-trans-R73

The sense of a boundary is reinforced by the comma placed after *wa*, which indicates a pause between the Topic/Theme and the Comment/Rheme. Although I still maintain *wa* does not mark Theme, its function as the Topic marker of the Topic-Comment structure and the fact that Topic often conflates with Theme mean the significance of *wa* must be recognised here. The present decision however, is not a general rule. Similar instances are handled

case by case. Whether or not *wa* divides a univariate Theme of two Circumstances, singling out the first as a sole Theme, will be decided taking into account the particle *wa*, presence or absence of a comma, and the flow of discourse.

4.3.3.3 Experiential Theme

An experiential element is obligatory in a Theme. As noted in 4.3.3.1, I allow only one experiential element in the Theme and take explicit elements as Theme, not allowing implicit Themes.

Experiential Theme in both Teruya's (2006) and Thomson's (2001) systems is either Participant or Circumstance. In addition to these, I incorporate verbal Predicator. It is a necessary arrangement due to not allowing implicit Themes, unlike Teruya and Thomson. Almost all the verbal Predicators appear in the projected clause as in example 4-45.

4-44

沖山容疑者は	//「だまして	いない」と//	否認している。
<i>Okiyama-yōgisha WA</i>	//'damashite	<i>inai' TO//</i>	<i>hininshiteiru</i>
<i>Okiyama WA</i>	//'trick-ASP	-neg'-PROJ//	deny-ASP
Okiyama denies (allegation) 'I haven't tricked (him)'.			
Theme	Theme: verbal Predicator	Rheme	Rheme
Projecting	Projected		Projecting

Ja-orig-Y13

Verbal Predicators are difficult examples of Theme as they often mean dividing a single word into Theme and Rheme as in 4-44 'だましていない [(I) have not tricked (him)]'. This is where implicit Themes would be usefully employed. However, since I do not allow them, a part of the verbal Process is assigned as Theme. An alternative approach is to treat it as a case without the Theme and Rheme structure. I opt for the first option.

Halliday and Matthiessen (2004: 111) offer an example of a functional element 'fused' into a single word although not in relation to Theme. They explain in the system of MOOD in English the Finite element indicating the present tense and the lexical verb are fused into one word 'loves'. It may not be directly applied to the system of THEME in Japanese, and in particular, it is

questionable whether a departing point and the developing part of the T-unit could be fused into one.

Nonetheless, I take this approach mainly because it is the consequence of two decisions I have made, i.e. not allowing implicit Themes and assigning part of verbal Process Theme, and partly because neither Teruya nor Thomson do. An attempt on a different means of identifying and analysing Theme may potentially lead to new discoveries about Theme in Japanese. Or it may offer reasons to accept implicit Theme.

In the case of Participant Theme, Teruya makes a distinction between operative and receptive clauses, in which Themes are either Subject/Agent or Subject/Medium, while Thomson (2001) adopts Participant types.

I am closer to Teruya in that I employ Subject and non-Subject division instead of Participant types. I need to include an option for non-Subject Theme that he does not because he allows implicit Themes and I do not. This approach enables the Subject Theme to be distinguished from the non-Subject Theme which occurs due to Subject being implicit as well as non-Subject being topicalised. Adopting the Participant types would effectively illustrate transitivity types. However, it will not show whether a Goal is realised as Theme by topicalisation or simply because the Subject is implicit for instance, which is an important distinction to make in light of translation training. In addition, setting the same distinction in Japanese and English if possible supports direct comparison.

In Theme realisation distinction between marked and unmarked Themes is important as this will affect the textuality of the translated texts. In English, a Subject-prominent language (Li & Thompson 1976) whose word order is relatively fixed, the Subject is an unmarked Theme in declarative sentences. Other elements are marked Themes.

In Japanese it is not as clear as that. An informal pilot study I conducted using a part of the Ja-orig corpus (113 independent clauses found in the articles collected in May 2014) revealed the order of Circumstance^Subject occurred in 57% of cases while Subject^Circumstance occurred in 43%, when both Circumstance and Subject were present. Statistically they are not significantly different ($X^2=1.991$, $df=1$, ns), which means it is difficult to affirm which order is more natural, i.e. unmarked.

Teruya includes markedness in his system of THEME. However, what exactly is unmarked or marked is not entirely clear. He states that the element marked by *wa* is an unmarked Theme but he determines non-Agent elements that are

marked by *wa* in the receptive clauses to be marked Themes (Teruya 2006: 98-101). In addition, his system shows that a marked Theme can be realised with the particle *wa* (4.3.3), which seems to contradict his definition of an unmarked Theme.

In relation to translation, Naganuma (2000) emphasises the importance of paying attention to markedness of Theme but without explicitly stating what constitutes a marked Theme in Japanese.

Thomson discusses the issue making use of the frequency of realisation in her data, stating Participant as unmarked Theme and Circumstance as marked (2001: 129). However, she does not include it in her system. I also do not include the options of marked and unmarked Theme in mine. However, I will discuss it in relation to the frequency of realisation in Chapter 5, in the same manner as Thomson.

Although I neither recognise the particle *wa* as a Theme marker nor relate it with unmarkedness, there are times when *wa* marking is usefully distinguished. It occurs when the Subject is implicit and non-Subject Participant is realised as Theme. Whether it is marked by *wa* or not, i.e. topicalised, strongly influences the way text develops or is interpreted, and therefore, is an important factor in relation to translation training. In 4-45 non-Subject Participant is realised as Theme that is also the Topic, and in 4-46 it is not topicalised.

4-45

依頼は	24 時間	受け付ける。
<i>Irai WA</i>	<i>24-jikan</i>	<i>uketsukeru</i>
Order <u>WA</u>	24 hours	take
Orders (they) take 24 hours.		
Goal (topicalised)	Circ.	Process
Theme	Rheme	

Ja-orig-M60

4-46

「女の人を	刺した」
'on'na NO hito <u>Q</u>	sashita'
'woman <u>Q</u>	stab-pst'
(I) stabbed a woman.	
Goal	Process
Theme: non-Subject Participant, not topicalised	Rheme

Ja-orig-M37

This approach enables me to distinguish whether a non-Subject Participant is Theme simply because the Subject is ellipted or because it is topicalised and fronted.

4.3.3.4 Absolute Theme

Thomson's system of THEME does not include absolute Theme but Teruya recognises it as a 'contextualizing element for the clause as a message (2006: 94). An absolute Theme does not have an experiential role as it is outside the transitivity structure of the clause (*ibid.* 94). Teruya explains that the absolute Theme is marked by *dewa* but it can be marked simply by *wa* (Noda 1996: 75-82). I take both as marking absolute Themes.

It is not always clear whether or not an initial element is an absolute Theme, as no experiential role does not mean no semantic relationship between the absolute Theme and the rest of the clause. Also the Topic-Comment structure allows the Topic, marked by *wa*, to have no 'selectional relationship' with the verb (Li & Thompson 1976: 468). An example of absolute Theme is 4-47.

4-47

児童手当は、	子供の年齢に応じて	1人当たり月額1万～1万5000円を	支給。
<i>Jidōteate WA</i>	<i>kodomo NO nenrei NI-ōjite</i>	<i>hitori atari tsukigaku 1-man - 15000-en O</i>	<i>shikyū</i>
child-care allowance WA	child NO age NI <i>depend on</i>	one (child) per monthly amount 10000-15000 yen O	supply
As for child-care allowance, (they) supply 10000-15000 yen a month per person (child) depending on the child's age.			
N/A	Circumstance	Goal	Process
Theme	Rheme		

Ja-orig-M11

The Theme 'child-care allowance' in example 4-47 does not have an experiential role in the clause but it presents a Topic, on which the Rheme provides information as Comment.

4.3.3.5 Logical Theme

A dependent clause is placed in initial position in Japanese as a rule, and that is a Theme.

4-48-A

撃たれた後、	平岩巡査部長は	銃の弾倉をつかみ	発射できないようにした。
<i>Utareta-ato</i>	<i>Hiraiwa-junsabuchō WA</i>	<i>jū NO dansō O tsukami</i>	<i>hasshadekinai yōnishita</i>
shoot-psv-pst after	Hiraiwa police sergeant WA	gun NO magazine O grab-SUSP	fire-ABLE-neg-aim-pst
After being shot, Hiraiwa police sergeant grabbed the magazine of the gun and made it unable to fire.			
β	$\alpha 1$		$\alpha 2$
Theme	Rheme		

Ja-orig-M74

However, if there is a shared Subject that is marked by *wa*, i.e. made the Topic, and placed initially, that is the Theme. This is because everything that follows the Topic gives an account of it as in the Topic-Comment structure.

4-48-B

平岩巡査部長は	撃たれた後、	銃の弾倉をつかみ	発射できないようにした。
<i>Hiraiwa-junsabuchō</i> WA	<i>utareta-ato</i>	<i>jū NO dansō O</i> <i>tsukami</i>	<i>hasshadekinai yōnishita</i>
Hiraiwa police sergeant WA	shoot-psv-pst after	gun NO magazine O grab-SUSP	fire-ABLE-neg-aim-pst
Hiraiwa police sergeant, after being shot, grabbed the magazine of the gun and made it unable to fire.			
β : Goal, α : Actor	β	$\alpha 1$	$\alpha 2$
Theme	Rheme		

Ja-orig-M74 (reordered version)

Teruya does not assign the status of Theme to a dependent clause unless it is highlighted by the particle *wa* (2006: 107), explaining that in Japanese it always comes before the independent one on which it depends. Thomson on the other hand includes the logical Theme in her system of THEME, which is labelled either as β clause or Projection. She further assigns it labels of Temporal, Manner, Cause or Condition. I recognise the dependent clause in initial position as Theme but unlike Thomson I do not analyse it further than that as this is outside the scope of the study and the research questions can be addressed without that level of detailed analysis.

4.3.3.6 Projection and Theme

Projection involves both speech and ideas. In Japanese, speech can be quoted (directly) or reported (indirectly) while ideas are always reported. Thompson, talking about English, states that when a speech is quoted directly, the Themes of both projecting and projected clauses are best shown separately as both of them 'seem to be important in the development of the text' (2004: 161).

Teruya (2006: 103) agrees it is the same in Japanese, and explains that there are two lines of thematic development and the Theme of the projecting clause is related to the narrative sequence while the Theme of the projected clause is related to the universe of discourse that is created by the projection. Themes in both projecting and projected clauses are shown in 4-49-A.

4-49-A Theme in direct quotations

末松氏は	//「以後、	注意してください」と//	小松氏に	苦言を	呈した。
<i>Suematsu-shi</i> WA	//‘igo	<i>chūishitekudasai’</i> TO//	<i>Komatsu-shi</i> NI	<i>kugen O</i>	<i>teishita</i>
Mr Suematsu WA	//‘from now on	careful please’- PROJ//	Mr Komatsu NI	warning O	give-pst
Mr Suematsu warned Mr Komatsu ‘from now on, please be careful’.					
Theme: Sayer	Theme	Rheme	Rheme		
Projecting	projected		Projecting		

Ja-orig-M55

Projected clauses can be placed in initial position. If they are direct quotations as in 4-49-B, a double layered annotation is conducted as a special case.

4-49-B

//「以後、	注意してください」と//	末松氏は	小松氏に	苦言を	呈した。
//‘igo	<i>chūishite kudasai’</i> TO//	<i>Suematsu-shi</i> WA	<i>Komatsu-shi</i> NI	<i>kugen O</i>	<i>teishita</i>
//‘from now on	careful please’- PROJ//	Mr Suematsu WA	Mr Komatsu NI	warning O	give-pst
‘From now on, please be careful’, Mr Suematsu warned Mr Komatsu.					
Theme	Rheme				
Theme		Rheme			

Ja-orig-M55 (reordered version)

The Theme in the projected clause is ‘igo [from now on]’ and simultaneously the whole projected clause is Theme for the clause complex.

With regard to reported speech, as in English, the Theme of the projected clause is not analysed in this project, given that the projected clauses are hypotactically related to the projecting clause. Teruya does not explicitly explain about the Theme and indirect speech. However, the examples he uses show he treats reported speech as dependent clauses and without their own Theme.

Similar to English, there are examples of quoted speech that are not fully developed clauses. In such cases, as in English, they are not treated as projected clauses as in 4-50.

4-50

その内容は	「[[眠気も吹っ飛ぶ]]うれしい内容」(同課長)	だった。
<i>Sono naiyō WA</i>	<i>'[[nemuke MO futtobu]] ureshii naiyō' (dō kachō)</i>	<i>datta</i>
Its content WA	'[[sleepiness MO fly away]] pleasing content' (same section manager)	be-pst
The content was 'the pleasing content that even sleepiness goes away' ((by) the same section manager).		
Carrier	Attribute	Process
Theme	Rheme	

Ja-orig-M66

In relation to Projection, there are two issues that need clarifying. One is split Rheme. In 4-51 the Rheme is split into two by the projected clause.

4-51

自民党の石破幹事長は	会談後の記者会見で、…	//「いかなる事態でも対応できるように態勢を整えておく」と//	述べた。
<i>Jimintō NO Ishiba-kanjichō WA</i>	<i>kaidan-go NO kashakaiken DE</i>	<i>//'ikanaru jitai DEMO taiōdekiru yō taisei O totonoeteoku' TO//</i>	<i>nobeta</i>
Liberal Democratic Party NO Ishiba General Secretary WA	meeting after NO press conference DE, ...	//'Whatever situation DEMO respond-ABLE-aim. preparation O make ready-ASP'-PROJ//	state-pst
General Secretary of Liberal Democratic Party Ishiba stated in a press conference after the meeting '(we) will be prepared to be able to respond in whatever situation.'			
Theme	Rheme: Split front	Projected	Rheme: Split rear
Projecting			projecting

Ja-orig-Y26

The Rheme in the projecting clause in 4-51 is split into two, namely front and rear. The front split Rheme in 4-51, a Circumstance of Location, can appear

after the projected clause. In such cases, the Rheme can be simple. It is also possible for a Rheme to be split into three, front, middle and rear, if more than one projected clause by the same Sayer is included in one T-unit.

The second issue is the distinction between projection and an expression of evidentiality ‘*to iu* [said to be]’. Teruya (2006: 219) explains that the latter belongs to the system of EVIDENTIALITY; the verbal process *iu* [say] and the projection marker *to* pressed into service as an expression of evidentiality. It describes the ‘source’ to report the proposition as hearsay (*ibid.* 219). Because of the combination of the particle *to* which generally marks projected clauses and the verb *iu* [say], it is at times difficult to distinguish the particle *to* as a projection marker from that in the expression of evidentiality.

To in 4-52 is an example of evidentiality. The Subject ‘被害男性は [the man/victim] is part of the proposition rather than the Sayer; the sentence means ‘the man/victim is said to be...’. Although it is not impossible, depending on the situation, for the sentence in question to mean ‘the victim says’, taking the particle *to* as marking the projection, it does not mean such in this case.

4-52

被害男性は、	…2週間のけがと	診断されたという。
<i>Higai dansei WA</i>	... 2-shūkan NO kega TO	<i>shindansareta to iu</i>
Victim male WA	... 2 weeks NO injury TO	diagnose-psv-pst-h.SAY
The male victim is <u>said to be</u> diagnosed of having an injury which takes 2 weeks to recover.		
Medium	Phenomenon	Process
Theme	Rheme	

Ja-orig-M2

The key then, is to identify the Sayer. Teruya (2006: 219) explains that the source is often implicit with the evidentiality ‘*to iu*’ but implied as reliable. It is indeed when the Sayer is implicit but can easily be identified in the preceding discourse, the confusion arises. These instances are tested if ‘*to iu*’ can select for MOOD, i.e. if it can, it is a projection since evidentiality ‘*to iu*’ never changes the form. In addition, when the source of information is included in the T-unit in question but not as Sayer, the same test is applied. The Circumstance of Angle in 4-53 is in fact agnate as Sayer. However, the Process type is relational, instead of verbal because ‘*to iu*’ expresses evidentiality.

4-53

CEAによると、	[[最低賃金を受け取っていない]]労働者は	10%に	上ると <u>いう</u> 。
CEA NI-yoruto	[[saitei chingin O uketotteinai]] rōdōsha WA	10% NI	noboru <u>to iu</u>
CEA according to	[[minimum wage O receive not]] worker WA	10% NI	reach-h.SAY
According to CEA workers who do not receive the minimum wage are said to reach 10%.			
Circumstance	Medium	Attribute	Process
Theme	Rheme		

Ja-trans-R54

4.3.4 Annotating Theme

4.3.4.1 Systems of Theme and general principles for annotation

The system to analyse Theme of English is presented in Figure 4-8 and 4-9 Japanese, which are followed by the general principles for annotation.

Figure 4-8: System of THEME in English (adapted from Thompson 2004)

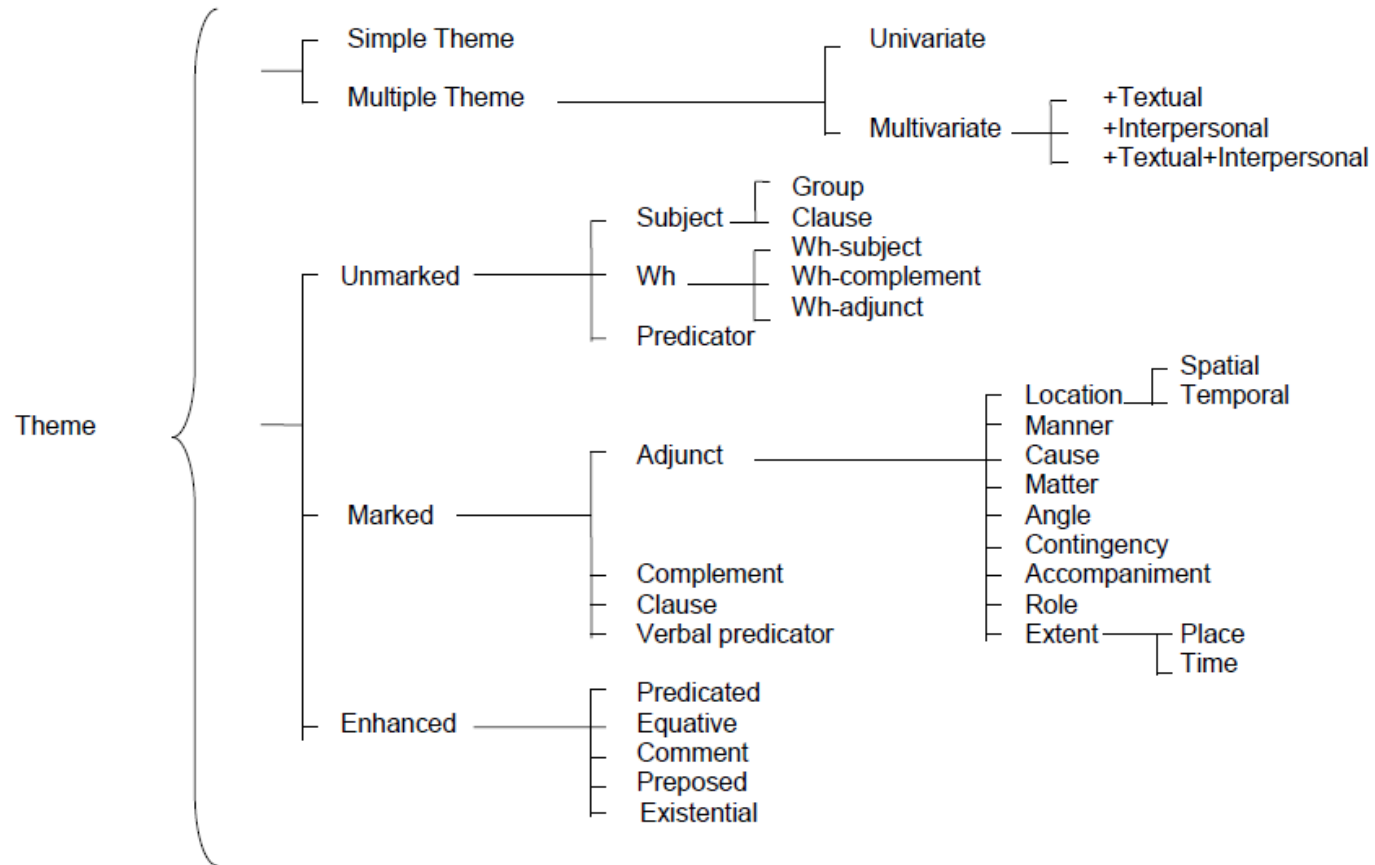
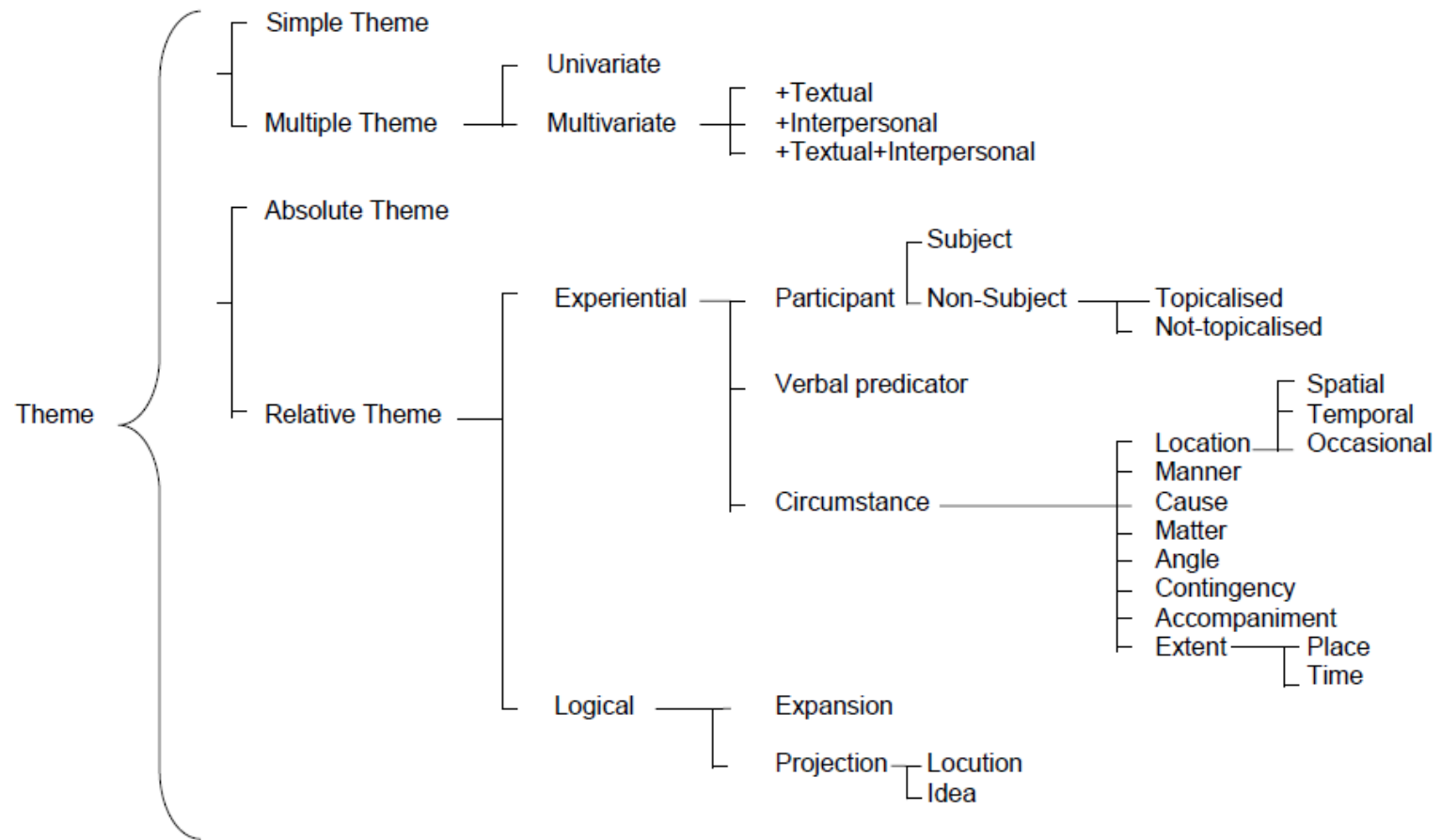


Figure 4-9: System of THEME in Japanese



The general principles for annotating Theme both in English and in Japanese are as follows.

1. The basis for analysis is the T-unit;
2. Themes are realised in the initial position of the T-unit;
3. There are parallel systems of simple or multiple Theme and ideational Theme;
4. An obligatory experiential element on its own makes a simple Theme;
5. A multiple Theme consists of one experiential element plus textual and/or interpersonal elements;
6. A multiple Theme can be univariate that consists of two circumstantial Themes;
7. The univariate multiple Theme, originally suggested for Japanese, is extended to English to allow comparison of them;
8. Only one experiential element is included in Theme;
9. Clauses that are paratactically related to each other in a clause complex form one T-unit and each has its own Theme;
10. If paratactic clauses share a referent, they form one T-unit and only one Theme is identified for the T-unit;
11. Clauses that are hypotactically related to each other in a clause complex form one T-unit;
12. Quoted speech is regarded as paratactically combined with the projecting clause and therefore Themes in both projecting and projected clauses are identified;
13. Reported speech is regarded as dependent clauses within the T-unit and only one Theme in the projecting clause is identified.

4.3.4.2 Guidelines specific to English

In addition to the general principles that apply both to English and Japanese, guidelines specific to English in identifying Theme are as follows.

1. A distinction is made between marked and unmarked Theme;
2. Predicator Theme in imperative clauses is an unmarked Theme;
3. Verbal Predicator is different from Predicator in imperative clauses and is marked Theme. It is the verbal Predicator positioned before the Subject in the projecting clause that immediately follows the projected clause;
4. If the order of the hypotactically connected clauses is independent clause followed by dependent clause ($\alpha^{\wedge}\beta$), the Theme is the element of the independent clause in initial position;

5. If the order of the hypotactically connected clauses is dependent clause followed by independent clause ($\beta^{\wedge}\alpha$), the dependent β clause as a whole is the Theme;
6. With regard to comment structure and existential structure as enhanced Theme, comment and 'there + existential Process' are recognised as Theme rather than simply 'it' or 'there'.
7. Multiple layered annotation is generally not conducted.

4.3.4.3 Guidelines specific to Japanese

The following are the general principles of Japanese specific guidelines to identify Theme. Nos. 2 and 4 are adopted from Thomson (2001), no.6 is drawn from Thompson (2004), and the rest are my own.

1. Only explicitly realised elements are Theme (implicit Theme is not allowed);
2. The particle 'wa' is not considered as a Theme marker;
3. However, topicalised non-Subject Themes marked by 'wa' in initial position are distinguished from non-Subject Themes that are not topicalised (not marked by *wa*);
4. A dependent β clause placed in initial position is the Theme;
5. However, if the Subject (marked by *wa*) is placed in initial position that is shared by both the α & β clauses in the clause complex, the Subject is the Theme, not the whole dependent β clause;
6. Themes are shown separately in direct quotations but not in indirect reports since the latter are regarded as the dependent clauses within the T-unit;
7. If a projected clause that is a direct quotation is in initial position, a double layered annotation is conducted as an exception; a Theme is identified in the projected clause, and the projected clause itself is identified as Theme for the whole T-unit.

4.4 Identifying additions, deletions and moves

4.4.1 Rank of additions, deletions and moves

Translations are manually annotated against their STs to be segmented and assigned with labels of addition, deletion and move. According to Bielsa & Bassnett (2009: 64) important background information is added, unnecessary information is eliminated and paragraphs are reordered in news translation to suit the news to the new audience (Chapter 3). By examining what is actually

added, deleted and moved, it is possible to see what is considered important, unnecessary and an appropriate order in the translated news in English and in Japanese.

Once those instances of translation shifts are identified, their ranks are decided according to the 'semantic ranks' (Halliday & Matthiessen 2004: 589); text unit, sequence, figure and element. However, since semantic meanings are realised by grammar, e.g. the semantic rank of figure is realised by the grammatical rank of clause, grammatical ranks are also presented in Figure 4-10 for reference. The rank of figure is the smallest at the rank of semantics but largest at the rank of grammar as it corresponds to the clause (*ibid.* 661). Larger than the figure is sequence which at the grammar level is the clause complex (T-unit). Above that is the functional discourse unit that is the text unit.

Semantic ranks are employed because the focus of investigation is the functional unit of meaning rather than grammatical. Moreover, the translation shifts of additions, deletions and moves occur beyond the rank of clause or clause complex. Semantic ranks are not as clearly defined as the grammatical ranks other than the text, that is the most extensive, and the figure, the smallest. Intermediate ranks are likely to vary from one register to another and it is possible for different registers to operate with different compositional scale (Halliday & Matthiessen 2004: 591). This is why I draw on White (1998) with respect to the text unit to identify functional discourse segments of news articles (Chapter 3 and 4.2.4).

The ranks are set taking into account the domain of translation shifts' contribution to meaning making. This approach helps decide the rank of a given translation shift by considering the direct domain of its contribution. For example, it is possible for a figure (clause) to be a sequence (T-unit) too as in 'She is the first female US ambassador to Japan.' In such cases the larger rank is employed because its contribution of meaning making is for the text unit and not the sequence. It is also possible for the same example to be a text unit, though it is not so in the actual text. If it were, it would be assigned as a text unit as its meaning making contribution is for the text.

Figure 4-10: Semantic rank of translation shift

Semantic rank of translation shift	Realised by	Examples from 4-54 below
text unit		10
sequence	T-unit	2, 6, A, C, E, F
figure	clause	3
element	phrase, group, word, morpheme	1, 4, 5, 7, 8, 9, B, D

Below the rank of figure is element, which is phrase/group/word/morpheme at the grammar level. Teich (2003) includes only phrases and groups as element, but I add words and morphemes to include examples such as 米 from 米国 [US] as in 米東部 [US eastern part] or 米大統領 [US President].

Since the investigation is focused on the text meaning, any additions, deletions and moves below the rank of figure (clause) are assigned as element, without classifying them further. Embedded clauses are also included in the category of element because they are 'rankshifted' clauses and function within the structure of a group or phrase (Halliday & Matthiessen 2004: Ch7). This rank enables distinguishing and comparing global and local translation shifts, which are above and below the figure.

In the example 4-54-ST and TT below, instances of additions, deletions and moves are displayed. Deletions are indicated by pink font in the ST, and additions by orange font in the TT, and moves are highlighted in grey in both the ST and TT to illustrate the movement. The corresponding moved items are specified by the alphabet letters A to E. Blue lines indicate the boundaries of text units.

4-54-ST

Blind man beaten on Philadelphia street, police video shows
 By Julia Talanova, CNN
 October 9, 2013 -- Updated 1336 GMT (2136 HKT)

(CNN) – ^A A video showing a blind man being repeatedly kicked and stomped by an unknown assailant ^A was released by the Philadelphia Police Department Tuesday.

^B The 33-year-old victim was walking down the street around noon on October 2 in Philadelphia's Southwest Germantown neighborhood when the suspect looked in his direction as he was approaching, according to police⁽¹⁾.

The **grainy**⁽¹⁾ police surveillance video shows the suspect laying down his backpack near a corner shop after he spots the victim.

The video goes black⁽²⁾, and the next moment shows the victim being pushed to the ground and repeatedly punched, stomped on and then kicked about four times by his attacker as witnesses watched.

One man is seen crossing the street mere steps from the beating, and he continues walking.

After the suspect finishes assaulting the blind man⁽³⁾, **he** picks up his backpack and walks away, leaving the victim helpless on the ground **with** injuries to his head and face, according to police.

Philadelphia⁽⁴⁾ **police** spokeswoman Officer **Tanya Little**⁽⁵⁾ said it was not clear whether the suspect knew that the victim was blind. **The** victim told police he did not recognize his attacker's voice.

Police are asking the public to help identify the attacker, **whom**⁽⁶⁾ they described as a 20-year-old black man wearing a tan baseball cap, tan shirt and khaki pants.

En-orig-C1

4-54-TT

路上で盲人に暴行、通行人ら素通り 米フィラデルフィア
2013.10.10 Thu posted at 17:57 JST

(CNN) **米東部**⁽⁷⁾フィラデルフィアの警察は10日までに、盲目の **33歳**男性が市内の路上で男に襲われ、繰り返し殴られたり足で踏み付けられるなどの暴行を受ける事件が起きたと報告した。**被害者は頭部と顔面を負傷した。**

A 地元警察は犯行の様相を収めた監視ビデオの画像も公表。**E** 警察の報道担当者は、被害者が目が不自由であったことを容疑者が知っていたのかどうかは不明としている。

ビデオ画像によると、容疑者は被害者を見かけた後、通りの角にある店の近くにバックパックを置き、襲いかかっていた。路上に目撃者がいるにもかかわらず、4度蹴るなどの暴力を加えた後、**C** 被害者を放置し、バックパックを拾い上げて立ち去っていた。

ビデオには、通りの暴行現場から数歩のところを横断する男性1人が事件を無視して歩き続ける場面も映されていた。

警察によると、暴行犯は20歳**前後**⁽⁸⁾の**アフリカ系**⁽⁹⁾(黒人)で、褐色の野球帽子にシャツ、カーキ色のズボンの姿だった。**F** 被害者は容疑者の声に聞き覚えはないと証言した。

暴行の動機や、被害者が所持品を奪われたのかなどは伝えられていない⁽¹⁰⁾。

Ja-trans-C1

The following section describes annotation principles.

4.4.2 Annotating additions, deletions and moves

Additions, deletions and moves are analysed in the main body of the text. Changes involving the headline are not included in the analysis as described in Chapter 3.

In addition, changes triggered by differences in the grammar of English and Japanese are not considered either. This includes instances such as adding a Subject pronoun in a sentence in the English translation that is implicit in the Japanese ST, or deleting one in the Japanese translation. The rationale is that it is not the choice of the translator/editor but an inevitable change due to the grammatical difference. The following illustrates what is targeted for investigation in the present project as additions, deletions or moves.

Additions are annotated when items appear that do not exist in the ST in the translation. For example, 米東部⁽⁷⁾ [US eastern part] in 4-55-TT is an addition at the rank of element. The ST simply states ‘Philadelphia’ but in the TT the country it is located in and its approximate region within the country are added.

At times words chosen in the TT indicate a sense of addition although nothing visible is actually added. For example ‘the victim told...’ in 4-55-ST is translated as ‘被害者は...証言した [the victim testified]’. Because of the lexis a more specific meaning is conveyed in the TT which may have been missed otherwise. This type of addition is termed as ‘implicit explanation’ by Aktan & Nohl (2010) as a form of explicitation in their research of typical patterns of trans-editing from English to Turkish. These implicit explanations are not analysed in this project, not because they are insignificant but because the focus is given to the items of information content that can be annotated as added items at the level of lexicogrammar.

Also excluded from analysis on the same basis is ‘semantic shift’, (Aktan & Nohl 2010), which is similar to implicit explanations. A translation of ‘North Korea’ into ‘The Pyongyang administration’ is an example of this. Other examples from my own corpus include a day of the week in the ST changed to a day of the month in the TT, or a personal pronoun translated as ‘the suspect’. According to Aktan & Nohl (2010) reasons for attempting these shifts range from adhering to the TL rules or the genre conventions to improving comprehension.

A final example of exclusion from addition is that caused by a change of syntactic structure which triggers some form of addition as in 4-55. A relational Process is translated into a verbal Process and the new structure requires additional constituents.

4-55-ST

別れ話が	あり
<i>Wakarebanashi GA</i>	<i>ari</i>
Breakup talk GA	exist-SUSP
Existent	Process
There is a breakup talk	

Ja-orig-M9

4-55-TT

She	said	she	wanted	to break up with him.
Sayer	Process	Senser	Process	Phenomenon
Projecting		Projected		

En-trans-M9

This type of addition due to the choice of different structure in the TT is not considered as a case of addition.

I now turn to deletions, annotated when items that exist in the ST do not appear in the TT. In 4-56 for instance, a proper noun ‘Tanya Little₍₅₎’ is deleted as well as ‘Philadelphia₍₄₎’. They are easily identified as deletion because the TT structure preserves the other elements intact in the same Process type.

4-56

Philadelphia ₍₄₎ police spokes-woman Officer Tanya Little ₍₅₎	said	it was not clear whether the suspect knew that the victim was blind.
警察の報道担当者は、	被害者が目が不自由であったことを容疑者が知って いたのかどうかは不明と	している。

In annotating deletion, projection often draws attention, with similar types of structure change occurring regularly. For example, an English sentence comprised of a projecting and a projected clause are translated into Japanese not as projection but with the expression of hearsay ‘*to iu* [it is said]’. Most of them are categorised as deletion because it is the Sayer that is deleted. If the Sayer is expressed differently, say, as the Circumstance of Angle (example 4-

53), it is not annotated as deletion. A structural change from quoted speech to reported speech itself does not count as deletion.

Finally, moves are annotated when items in the ST do not appear in approximately the same position in the TT. Bielsa and Bassnett (2009) note only paragraph reorganisations as a move but I include movements of items both larger and smaller than the paragraph because they do occur. What I do not recognise as a move is that which takes place within the same clause or clause complex. For example, if a dependent β clause of the α^{β} order in an English sentence is translated in the order of β^{α} in Japanese, it is not a move. This case will not count in any case as this is also an obligatory change that is triggered by the grammatical constraint; Japanese always place the dependent β clause in initial position.

In deciding the rank of moves, annotation is approximate although the aim is to achieve complete accuracy. For example, there are instances in which part of a T-unit is deleted and the rest is moved as in 4-58 taken from 4-55-ST above.

4-57 Philadelphia⁽⁴⁾ E police spokeswoman Officer Tanya Little⁽⁵⁾ E said it was not clear whether the suspect knew that the victim was blind.

The rank of move here is the sequence (T-unit). With two underlined elements deleted, the remaining part is still a complete T-unit.

Another example concerns a conjunction (double underlined). In 4-58 the underlined clause is deleted and the conjunction is also removed. The rank of deletion is the figure (clause) only and the deletion of conjunction is treated as part of the deletion of figure, instead of in its own right.

4-58 Police are asking the public to help identify the attacker, whom⁽⁶⁾ they described as a 20-year-old black man wearing a tan baseball cap, tan shirt and khaki pants.

Taking the above as general rules, additions, deletions and moves are annotated.

Chapter 5

Theme selection and translation shift

This chapter concerns Theme selection from the viewpoint of register and aims to answer research questions 1-2 which are set out in Chapter 1. The focus is on the type of elements that are frequently realised as Theme in translated news articles. This chapter also presents the results of the quantitative analysis of Theme realisation in En-orig and in Ja-orig (5.1) to identify contrasts between them as possible indicators of shifts. Then it continues to describe Theme realisation in En-trans in comparison to En-orig and to Ja-orig to investigate where thematic shifts might occur and whether Theme realisation in translation adheres to the source text or moves closer to the target language convention possible patterns of shifts (5.2). The same procedure is applied to the description of Theme realisation in Ja-trans in comparison with Ja-orig and En-orig (5.3). There follows a summary of patterns of shifts in terms of Theme selection for En-trans and Ja-trans (5.4).

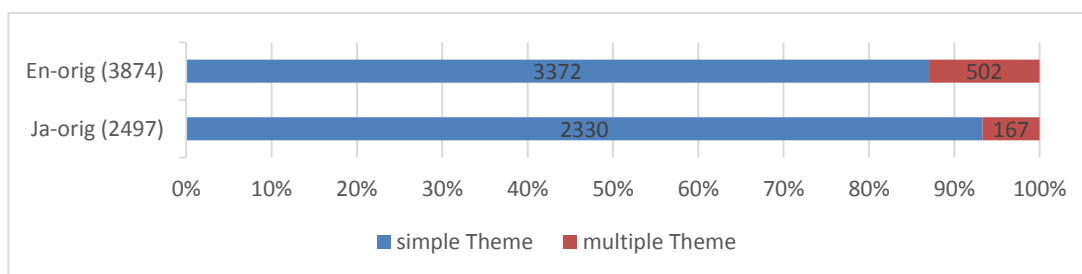
5.1 Theme in En-orig and Ja-orig

This section describes Theme realisation in the En-orig and in the Ja-orig in order to observe contrasts and similarities between them. This contrastive account is then used as a basis for analysing Theme realisation in English translation (5.2) and in Japanese translation (5.3). Statistics provided in 5-1 are obtained using an online statistical analysis programme js-STAR, version 9.0.6j (Tanaka & Nakano 2018).

5.1.1 Simple and multiple Theme

Figure 5-1 presents the proportion of simple and multiple Themes in En-orig and Ja-orig. Simple Theme constitutes the majority in both datasets at around 90%; in En-orig 87% and in Ja-orig 93%. The ratio of 0.9: 0.1 is relatively common in the system of language for the distribution of probabilities between unmarked and marked choices (Halliday 1992).

Figure 5-1: Simple & multiple Theme (En-orig & Ja-orig)



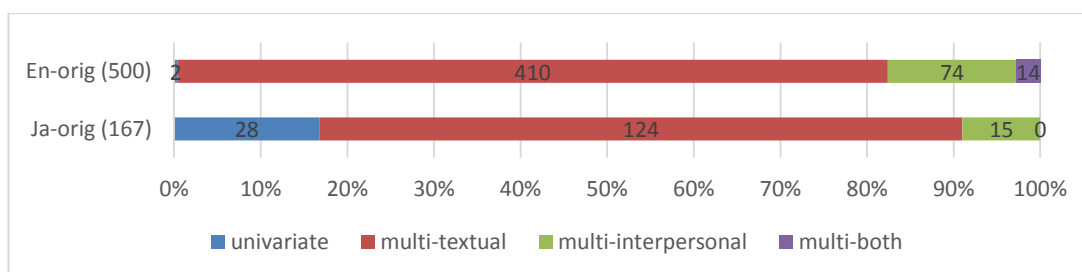
The data used for this study show that news articles both in En-orig and in Ja-orig develop the text mainly by means of simple Theme as the departing point of the T-unit. However, Ja-orig makes use of simple Themes more than En-orig does and the two datasets are found to be significantly different ($X^2=60.162$, $df=1$, $p<.01$).

Multiple Themes that include textual and interpersonal Themes are marked and rare in both En-orig and Ja-orig. However, if they are employed, they constitute an interesting contrast. Major differences are observed in the use of univariate multiple Theme and different types of textual Theme.

A typical multiple Theme consists of one experiential Theme plus textual Theme and/or interpersonal Theme. However, as explained in Chapter 4, this study draws a distinction between typical multivariate multiple Theme and univariate Theme in English in order to offer direct comparison with Japanese. As discussed in Chapter 4 Section 3.3.2, univariate multiple Themes consist of two Circumstances; either two Locations or one Location and one Extent.

Figure 5-2 displays the proportion of univariate and multivariate multiple Themes. Multivariate Themes are further divided into three types; experiential + textual (multi-textual in the figure), experiential + interpersonal (multi-interpersonal) or experiential + textual + interpersonal (multi-both).

Figure 5-2: Multiple Theme type (En-orig & Ja-orig)



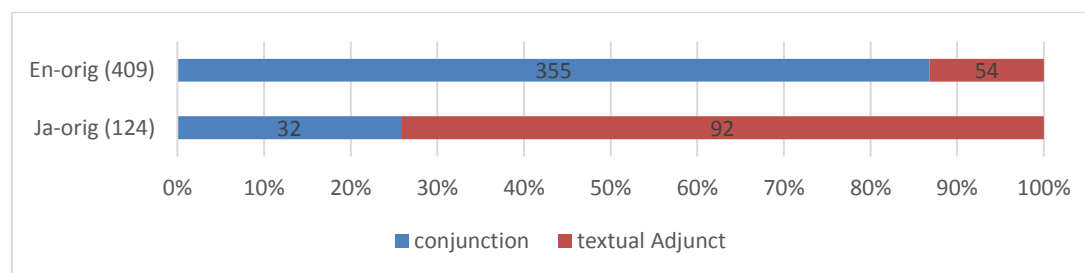
It is clear from Figure 5-2 that univariate multiple Themes rarely occur in En-orig (2 instances, 0.4%), while in Ja-orig they account for 18%. The two datasets are significantly different in terms of the choice of univariate or multivariate Themes ($X^2=79.389$, $df=1$, $p<.01$). Thomson points out univariate Themes are a possible genre feature of Japanese news reports (2013) and

here I have uncovered a correlation of register and data analysis. Since this feature is not shared in English, a shift is anticipated both from and into Japanese translation.

As for multivariate multiple Themes, the combination of an obligatory experiential Theme and a textual Theme accounts for the majority both in En-orig and Ja-orig (Figure 5-2). The difference between the datasets on this account is weak ($X^2=5.912$, $df=2$, $p<.10$). In terms of the other two multivariate multiple Themes, i.e. the interpersonal + experiential combination and the combination of experiential + textual + interpersonal, En-orig and Ja-orig turn out to be different ($p<.05$).

When we look at the types of textual Theme, a choice between Conjunction and textual Adjunct, a difference emerges as in Figure 5-3.

Figure 5-3: Textual Theme type (En-orig & Ja-orig)



En-orig employs Conjunctions more than it does textual Adjunct while Ja-orig employs textual Adjunct more than Conjunctions. The two datasets are significantly different in the use of these ($X^2=171.796$, $df=1$, $p<.01$). A possible reason for the proportional difference of Conjunctions is that certain types of Conjunctions appear towards the end of the clause in Japanese as in 5-2, while in English they appear in initial position. An example from English is presented first (5-1) and its translation in Japanese follows to indicate the positional difference of a Conjunction. Clause boundary in 5-1 is indicated by //.

5-1 The test was nearly successful, // but the rocket's spinning (En-orig-R41)

5-2

実験は成功したように見えた <u>が</u> 、	機体が回転して
The test Wa succeed-pst-appear NI look-pst <u>BUT</u>	body GA spin-SUSP
The test looked like it was successful but the body spun...	
1 clause	2 clause

Ja-trans-R41

The Conjunction ‘but’ in 5-1 appears at the beginning of the second T-unit as a textual Theme. On the other hand in 5-2, the equivalent Conjunction ‘が GA [but]’ is attached to the verb in the first T-unit, thus not to be a textual Theme in the second T-unit.

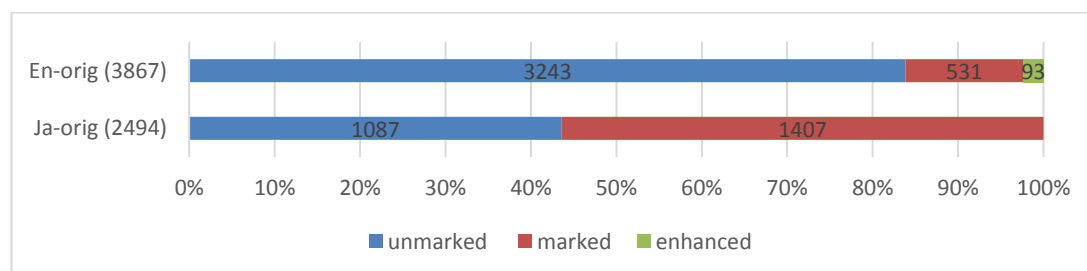
Conjunctions that appear at the beginning of a T-unit in Japanese are ‘etymologically agnate with structural conjunctions that represent logico-semantic relations’ (Teruya 2006: 82). Examples from the corpus includes *しかし shikashi* [but], *また mata* [also], *だが daga* [but] and *そして soshite* [and].

5.1.2 Ideational Theme

The systems of analysis for English and for Japanese are not exactly the same. In English the first distinction of ideational Theme is unmarked, marked or enhanced Theme (Chapter 4). In Japanese, it is either absolute or relational. Inevitably, some comparisons are approximate. However, the systems are not distinctively different and they have categories that are comparable such as Subject Theme. In addition, non-Subject participant, Circumstance and logical Themes in the system of THEME for Japanese are comparable to Complement, Adjunct and clause Themes in English respectively. Because of the similarity of categories, comparisons can be said to be reasonably accurate.

The first variable for a comparison is the unmarked Theme. It is a category in English that includes the Subject, wh-elements and the Predicator in imperative clauses. The comparison here is between the count on the unmarked Theme in En-orig on the one hand and the count on the Subject in Ja-orig on the other as the system of THEME in Japanese does not have a distinction between the marked and unmarked Theme. The result can be said to be sufficiently accurate since in over 99% of cases unmarked Theme is the Subject in En-orig.

Figure 5-4: Unmarked or other Theme (En-orig & Ja-orig)

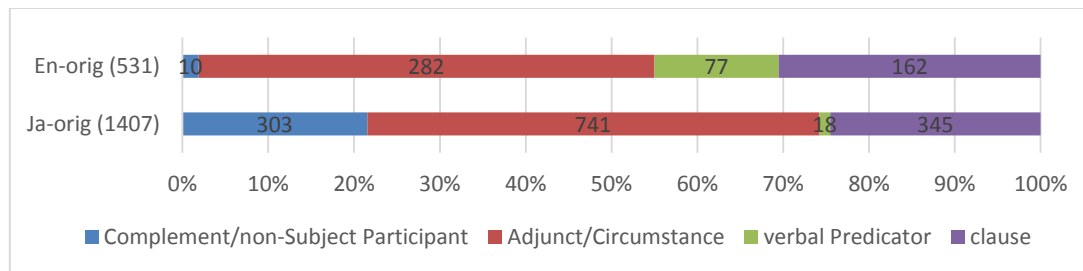


With regards to the choice of unmarked or other Theme, En-orig and Ja-orig are found to be significantly different at $X^2=1115.239$, $df=1$, $p<.01$. English predominantly starts the T-units with an unmarked Theme while Japanese

utilises Subject Theme and other Themes more or less equally. This is likely to trigger a shift in translation both from and into Japanese.

Moving on to the marked Theme, it includes Complement, Adjunct, verbal Predicator and clause Themes in English. In Japanese, the principally corresponding categories are non-Subject Participant, Circumstance, verbal Predicator and logical Themes. In terms of these variables, En-orig and Ja-orig are found to be significantly different ($X^2=271.413$, $df=4$, $p<.01$).

Figure 5-5: Marked Theme type (En-orig & Ja-orig)



Complement Theme is rare in En-orig, accounting for 2% of the total number of marked Themes, while in Ja-orig non-Subject Participant Theme accounts for 22%. The two datasets differ significantly ($p<.01$) on this account.

The difference lies mainly in the level of Subject explicitness in each language. Every free clause in English requires a Subject in general (Halliday & Matthiessen 2004: 151), while in Japanese subjects can be ellipited. The general word order of SOV in Japanese means that if the Subject is ellipited, the non-Subject Participant, i.e. the O of the SOV, is the subsequent candidate for Theme. The difference is typological, and not caused by the register.

There are two forms of realisation for the non-Subject Participant Theme; one is with the particle *wa*, i.e. topicalised, and the other is without *wa*, not topicalised (Chapter 4). In the corpus of Ja-orig, 21% of them are topicalised and 79% are not. This is unlikely to affect translation into English but may cause a shift into Japanese translation due to several decisions which need to be made, firstly whether or not to omit the Subject, and then whether or not to topicalise the non-Subject Participant.

Next category, the verbal Predicator exists in both English and Japanese. However, although the label is the same, they involve different elements as Theme (Chapter 4), and, therefore, a comparison is not attempted with respect to this account here. An example of the verbal Predicator in English (5-3, underlined) and in Japanese (5-4, also underlined) is;

5-3 "The big issue . . .," said Caplan.

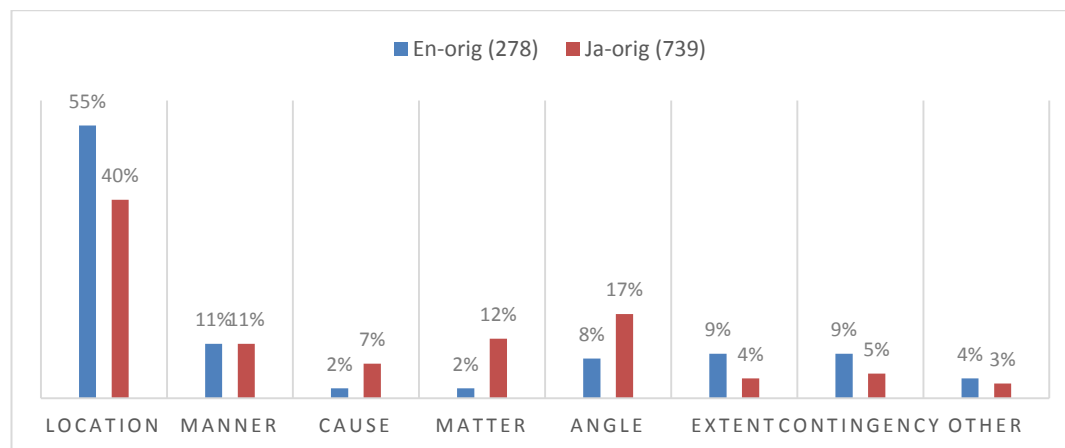
En-orig-C42

5-4 「だましていない」 '*damashite inai*' [(I) haven't tricked (him)] Ja-orig-Y13

Next is the clause Theme, with respect to which En-orig and Ja-orig are weakly different ($p < .10$). Although the level of significance is weak, the structural restriction on the side of Japanese, in which hypotactically related clauses are always organised in the order of $\beta^{\wedge}\alpha$, potentially prompts shifts during translation both from and into Japanese.

The final variable for comparison within the marked Theme is Adjunct/Circumstance Theme types. Regarding the use of Adjunct/Circumstance as Theme, this accounts for roughly 50% in both En-orig and Ja-orig (Figure 5-2) and no significant difference is found between the datasets. However, when it comes to which types of Adjunct are realised as Theme, significant differences emerge ($X^2=59.455$, $df=7$, $p < .01$) with regards to seven variables as in Figure 5-6, but excluding the category 'other'. The label 'other' here contains Accompaniment and Role, whose realisation proportions are small and there are no differences between the non-translation and translation in both English and in Japanese (5.2 & 5.3) with regard to these. The category 'other' also includes 'other', which is an additional category for instances that do not fit in other categories, and this will be discussed in 5.2 and 5.3.

Figure 5-6: Adjunct/Circumstance Theme type (En-orig & Ja-orig)



The two datasets are found to be significantly different in terms of Location, Cause, Matter, Angle and Extent ($p < .01$). With respect to Contingency, they are marginally different ($p < .05$), while no significant difference is found between them regarding Manner (no significance).

5.1.3 Summary of comparison

Table 5-1 summarises the discussions above and indicates contrasts found between En-orig and Ja-orig, including two variables that turn out not to be

significantly different. The level of significance is indicated by the number of +; 99% (+++), 95% (++) , 90% (+) and no significance (ns).

Table 5-1: Summary of contrast between En-orig and Ja-orig

Theme	En-orig & Ja-orig	significance
Simple Theme	Ja > En	+++
Multiple Theme	En > Ja	+++
Univariate Theme	Ja > En	+++
Interpersonal Theme	En > Ja	++
Textual Theme	En > Ja	+
Conjunction	En > Ja	+++
Textual Adjunct	Ja > En	+++
Unmarked Subject Theme	En > Ja	+++
Marked: Complement	Ja > En	+++
Marked: Adjunct	En = Ja	ns
Location	En > Ja	+++
Manner	En = Ja	ns
Cause	Ja > En	+++
Matter	Ja > En	+++
Angle	Ja > En	+++
Extent	En > Ja	+++
Contingency	En > Ja	++

In target-oriented translations such as news translation (Bielsa & Bassnett 2009), differences observed between the source and target languages as above are likely to prompt shifts.

5.2 Theme in En-trans (compared with En-orig & Ja-orig)

This section presents the results of the analysis regarding which elements are selected as Theme in En-trans. To discover similarities and differences of Theme realisation in En-trans with the comparable English non-translated

texts (En-orig) and with the original Japanese STs (Ja-orig), comparisons are conducted in a triangular form between En-trans and En-orig, and between En-trans and Ja-orig. The same format will be taken for Ja-trans in 5.3.

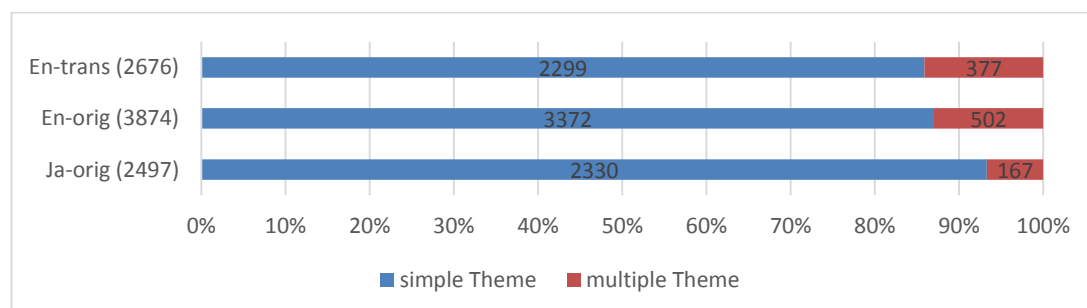
The following sections discuss specific aspects of Theme realisation from the results table obtained from the UAM CorpusTool. The system networks of THEME in English and in Japanese are presented in Chapter 4 Section 3.4. In addition, later in this chapter in 5.2.4 where a summary of the discussion is offered, the system network of THEME in English is provided again to indicate where translation shifts have occurred.

With regard to the statistics provided in this section, Chi-square values and significance levels are computed by the UAM CorpusTool for En-orig and En-trans comparisons because they are in the same project in the CorpusTool. For comparisons of En-trans and Ja-orig, which are not in the same project in the CorpusTool, statistics are calculated using js-STAR (Tanaka & Nakano 2018).

5.2.1 Simple and multiple Theme

Figure 5-7 shows the proportion of simple and multiple Themes in En-trans, alongside En-orig and Ja-orig. En-trans presents slightly more than 85% of simple Theme, suggesting that simple Theme is close to having 90% probability of being the unmarked choice (Halliday 1992). It can be said that the data used for this study show that news articles in English, both originals and translations, developed the text mainly by means of simple Theme as the departure point of the T-unit. En-trans and En-orig are found to differ, although the significance level is weak ($df=1$, $p<.10$) in terms of the choice of simple and multiple Theme.

Figure 5-7: Simple & Multiple Theme (En-trans)

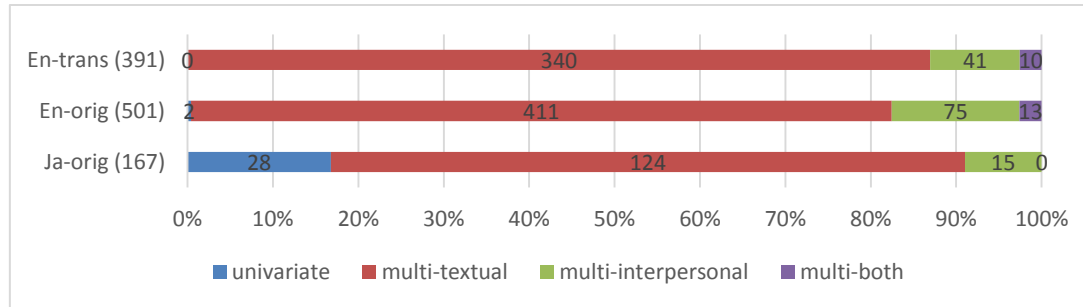


Compared with Ja-orig, En-trans is also significantly different from these ($X^2=79.465$, $df=1$, $p<.01$). This suggests that a shift has occurred during translation to make use of more multiple Themes in En-trans than appear in the Ja-orig. Given the proportional difference between En-orig and Ja-orig

($p < .01$) and similarity between En-trans and En-orig, the implication is that En-trans are closer to the TL norm.

As for different types of multiple Themes, a comparison is firstly made between the univariate and multivariate Themes.

Figure 5-8: Multiple Theme type (En-trans)



In En-trans no univariate Theme is realised and all the multiple Themes are multivariate. Similarly, En-orig make very little use of univariate Theme, with only two instances observed. No significant difference is found between En-trans and En-orig in terms of the choice of multivariate and univariate Theme. On the other hand, the univariate Theme accounts for 18% in Ja-orig, from which En-trans are found to be significantly different ($X^2=69.454$, $df=1$, $p < .01$). This suggests that a shift has occurred in translating univariate Themes from Japanese into English, possibly due to the linguistic restriction within English as the TL.

5.2.2 Textual and Interpersonal Theme

There are three types of multivariate multiple Theme; a combination of an experiential Theme and a textual Theme, an experiential Theme and an interpersonal Theme, or an experiential plus both textual and interpersonal Themes (Figure 5-8). The combination of an experiential Theme and a textual Theme represents the majority both in En-trans and in En-orig. En-trans makes slightly more use of it (87%) than En-orig (82%). In terms of the overall use of multivariate Theme, En-trans is different from En-orig although the significance level is weak ($df=2$, $p < .10$), while no significance is found compared with Ja-orig ($X^2=3.650$, $df=2$, ns). This suggests that En-trans maintains the structure of Ja-orig.

I now look at the types of textual Theme. Figure 5-9 shows a clear tendency for En-trans to make use of Conjunctions more than textual Adjuncts, although not to the same degree as in En-orig. In terms of their use the datasets are found to be significantly different ($X^2=81.05$, $df=1$, $p < .02$).

Figure 5-9: Textual Theme type (En-trans)

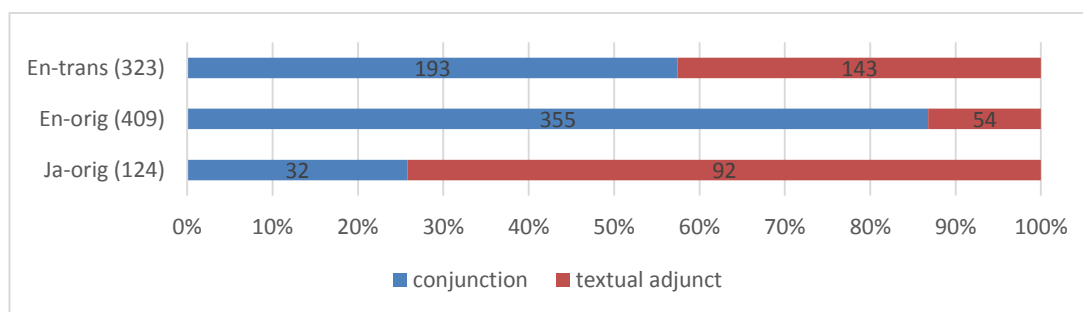


Figure 5-9 shows that En-orig employs Conjunctions most among the three datasets, followed by En-trans, with Ja-orig making the least use of them. This suggests that En-trans is influenced by Ja-orig to make use of textual Adjunct more than En-orig does, but it is also influenced by En-orig to utilise Conjunctions more than Ja-orig does, and so moving closer to TL norms. En-trans and Ja-orig are significantly different ($X^2=33.895$, $df=1$, $p<.01$).

With regard to the choice between the Conjunction and the textual Adjunct, statistics show that En-trans is significantly different from both En-orig and Ja-orig. Due to the linguistic difference mentioned in 5.1.1, En-trans is moving away from the SL tendency and towards TL norms, but not sufficiently close to it to be considered similar.

During analysis, several examples posed difficulties for deciding the type of textual Theme (underlined) as 5-5 and 5-6.

5-5 Authorities did not immediately release the identity of the dead nor did they provide details about where the bodies were found. En-orig-C56

5-6 As such, the board collected the manga in January this year,

En-trans-M51

For 5-5 Geoff Thompson recommends interpreting 'nor' as having both textual and interpersonal meanings (personal communication at ESFLC 2014). Although multiple annotation is possible with the UAM CorpusTool, it is annotated simply as '+ textual + interpersonal', which is the label for multiple Themes that involve one textual element and one interpersonal element such as 'and possibly'. 'Nor' in example 5-5, however, is one element which has two functions. It is important in this study to recognise multiple functions of Theme and ideally they are individually identified. However, the important point here is to distinguish the multiple Theme from the simple. Therefore, further annotation of 'nor' was not conducted. Example 5-6 was annotated as 'textual adjunct' because of its function as creating a link between the current T-unit and the preceding one.

Still staying with the multiple Theme, let us turn to interpersonal Themes, which make infrequent appearances in the corpus compared to textual Themes (Figure 5-8). This is not surprising because news articles are generally written to report facts objectively (Chapter 3). The proportion of interpersonal Themes, both with only the obligatory experiential Theme and combined with a textual Theme, is 13% in En-trans and 17% in En-orig. The two datasets differ in terms of the interpersonal Theme realisation ($X^2=4$, $df=4$, $p<.05$).

There are several possible reasons why En-orig includes more interpersonal Themes than En-trans; 1) En-orig employs direct quotations more that contain interpersonal Themes, and 2) En-orig includes news articles that are not typical hard news and have more space for interpersonal Themes. It is also possible that En-trans is influenced by Ja-orig, which includes significantly fewer instances of them than En-trans.

The relative infrequency of the interpersonal Theme in Ja-orig can be said to have two major causes. Firstly, most of the texts in Ja-orig are hard news and as such, written in a more formal register. Secondly, Mood types, which appear in initial position in English, are indicated towards the end of the clause in Japanese by the Predicator and the Negotiator (Teruya 2006: 135). There is no difference between En-trans and Ja-orig ($X^2=0.293$, $df=1$, ns).

Interpersonal Themes are either a finite or an interpersonal Adjunct and the latter appeared much more often (at 85-88%) than the former in both En-orig and En-trans. The datasets do not differ in terms of the use of interpersonal Adjunct. Examples of interpersonal Adjunct with the interpersonal Theme underlined.

5-7 Ultimately, the President would need to sign off on the decision.

En-orig-C37

5-8 “I think more people will hear his case and say ‘Hey, I never did get that naturalization paperwork’”

En-orig-R75

The latter is an example of the grammatical metaphor, indicating the interpersonal meaning of probability. It was expected that the interpersonal meaning would appear in direct quotations mainly as in 5-8, but they appear in descriptive parts too, as in 5-7.

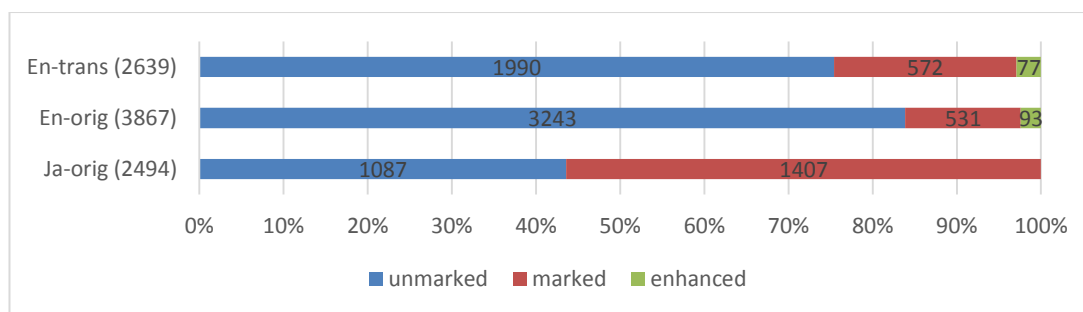
5.2.3 Ideational Theme

Categories for the Ideational Theme in English are unmarked, marked and enhanced Themes.

5.2.3.1 Unmarked Theme

Figure 5-10 displays the proportion of unmarked Theme, marked Theme and enhanced Theme for English. Unmarked Theme constitutes the majority in both En-trans and En-orig, accounting for over 75%, although En-orig indicates roughly 10% more instances than En-trans. The two datasets are found to be significantly different ($X^2=75.90$, $df=2$, $p<.02$).

Figure 5-10: Ideational Theme (En-trans)



In English the typical pattern of Theme realisation in declarative clauses is by the grammatical Subject, and in over 99% of cases unmarked Theme is indeed the Subject in both En-trans and En-orig data. Wh-element and Predicator appeared very little, at under 1% combined.

While the overall realisation of unmarked Theme in En-trans is significantly different from that of En-orig ($X^2=75.90$, $df=2$, $p<.02$), in terms of the type of unmarked Theme, the two datasets do not appear to be significantly different, with the chi-square value for each type between 0.00 and 0.90.

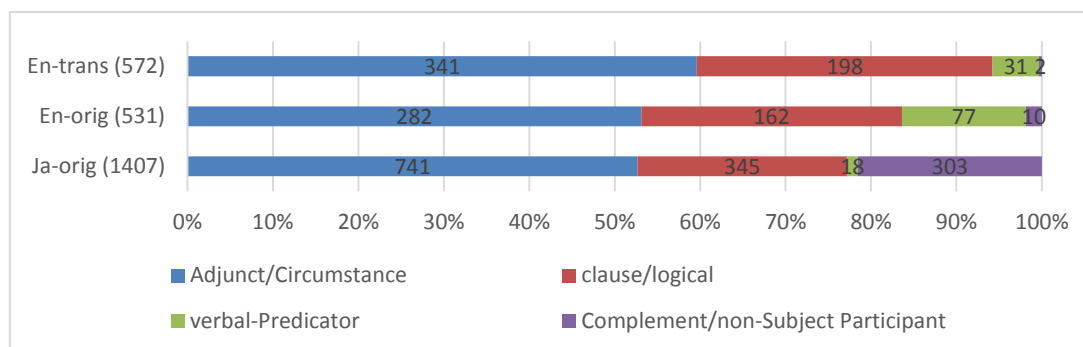
To compare En-trans with Ja-orig, given that the key unmarked Theme in English is the Subject, a proportion of the comparable variable in Japanese, the Subject Theme, is used which accounts for 48% of Theme (Figure 5-10). En-trans and Ja-orig are found to be significantly different ($X^2=704.338$, $df=2$, $p<.01$). Considering that En-trans is proportionally similar to En-orig more than it is to Ja-orig (Figure 5-10), it can be said that En-trans moves away from Ja-orig in its use of unmarked Theme and closer to En-orig.

The proportion of enhanced Theme is small in both En-trans and En-orig and they do not appear to differ from each other in terms of its realisation (no significance). Considering that Japanese does not exploit this kind of special thematic structure to manipulate word order, the fact that En-trans makes use of it indicates a shift has occurred. The specific types of enhanced Theme which are realised in En-trans are observed in 5.2.3.3.

5.2.3.2 Marked Theme

Figure 5-11 shows the realisation of different types of marked Theme. As explained in 5.1, they are not categorised as marked in Japanese but still available and a direct comparison between En-trans and Ja-orig is possible.

Figure 5-11: Marked Theme type (En-trans)



En-trans and En-orig differ significantly in terms of the overall use of the marked Theme types ($X^2=36.783$, $df=3$, $p<.01$). En-trans and Ja-orig are also significantly different ($X^2=165.115$, $df=3$, $p<.01$).

Focusing on individual types, Adjunct and Clause Themes account for a large proportion in all the datasets. They are not rare in Japanese (5.1). Similarly in English, although categorised as marked, they are not as uncommon in this role as Complement (Thompson 2004) or verbal Predicator.

En-trans displays higher percentages for Adjunct Theme than do En-orig and the two datasets are different in this respect ($X^2=4.36$, $df=3$, $p<.05$). On the same account, En-trans and Ja-orig are also found to be different ($X^2=165.284$, $df=3$, $p<.05$). The results can be interpreted neither as TL normalisation nor as ST influence because En-orig and Ja-orig are not different from each other (5.1).

With respect to the clause Theme, En-trans and En-orig are found to be different ($X^2=3.04$, $df=3$, $p<.10$) albeit weakly. On the other hand, compared with Ja-orig, En-trans is different from it ($X^2=4.772$, $df=3$, $p<.05$), more significantly than the difference between En-trans and En-orig. This implies that En-trans does not retain the clause Themes in Ja-orig, thus moving closer to the TL norm.

As predicted in 5.1, the difference in structure between English and Japanese is a likely reason why a shift occurs. Firstly, the hypotactic β clause in Japanese must be placed before the dominant α clause in a clause complex, while in English there is an option for the order of $\alpha^{\wedge}\beta$. Secondly, the clause Theme in Japanese includes projected clause (reported clause), while in

English this does not occur. These may have influenced the way the clause Themes in Ja-orig are translated into English.

Next is the verbal Predicator, which turns out to be the variable that clearly differentiates En-orig and En-trans ($X^2=28.60$, $df=3$, $p<.02$). En-trans employs this much less than En-orig and the reason is possibly because it is a choice, instead of an obligatory change in structure, but it is unclear when to select the verbal Predicate as Theme when translating from Japanese. A possible environment for the verbal predicator Theme (single underlined) is when the Sayer (double underlined) is long, as in 5-9.

5-9 "These little embryos, these are people born with a disease, they can't make power. You're giving them a new battery. That's a therapy. I think that's a humane ethical thing to do," said Caplan, the director of medical ethics at New York University's Langone Medical Center. En-orig-C42

However, there are also instances that involve short Sayers as in 5-10. This example is located several sentences after the example 5-9 in the same text, i.e. the Sayer is already known to the reader.

5-10 "The big issue over the next 5 to 10 years is going to become how far do we go in pursuit of the perfect baby," said Caplan. En-orig-C42

It seems that the verbal predicator Theme is likely to occur after a direct quotation from an individual whose position indicates authority for the first time as in 5-9, and then after more quotations from the same person as in 5-10.

Ja-orig has a category with the same label 'verbal Predicator' in the system. However, as explained in 5.1, the function of this is different from the English counterpart. Hence, comparing En-trans against Ja-orig on this account is not useful. However, the existence of the verbal Predicator Theme in En-trans suggests that En-trans is moving towards the TL norm in its use.

With regards to the next variable, Complement is the marked Theme with the fewest instances among all the marked Theme types and provides another example of greater use in En-orig than in En-trans. The two datasets are significantly different in this respect ($X^2=6.27$, $df=3$, $p<.02$).

5-11 So deep was his commitment, En-orig-C85

5-12 Also unclear is how many people have paid for their policies, En-orig-R53

Complement Themes in English in the data occur due to word order inversion, whose markedness is high compared to Adjunct or clause Theme (Thompson 2004:146). On the other hand, non-Subject participant Themes in Japanese,

which is comparable to Complement Theme in English, is considerably less marked as observed in 5.1. On this account, En-trans and Ja-orig are significantly different ($X^2=11.852$, $df=3$, $p<.01$). Given that Ja-orig includes considerably more Complement Themes than En-orig ($p<.01$) as well as En-trans appearing closer to En-orig proportionally (Figure 5-11), it can be said that En-trans is moving closer to the TL norm.

The significant difference between En-trans and Ja-orig suggests that the translators are aware of the different level of markedness of the Complement Theme in English and non-Subject participant Themes in Japanese. Because of that, non-Subject Participant Themes in Japanese are not translated as the Complement Theme in English, and other ways are employed such as with an additional unmarked Theme or as a Matter of Circumstance. Regarding the use of Adjunct as Theme, En-trans and En-orig are different ($p<.05$), which is marginally significant. However, when it comes to what types of Adjunct are realised as Theme, significant difference emerges between the two datasets, as in Figure 5-12 and Table 5-2.

Figure 5-12: Adjunct/Circumstance Theme type (En-trans)

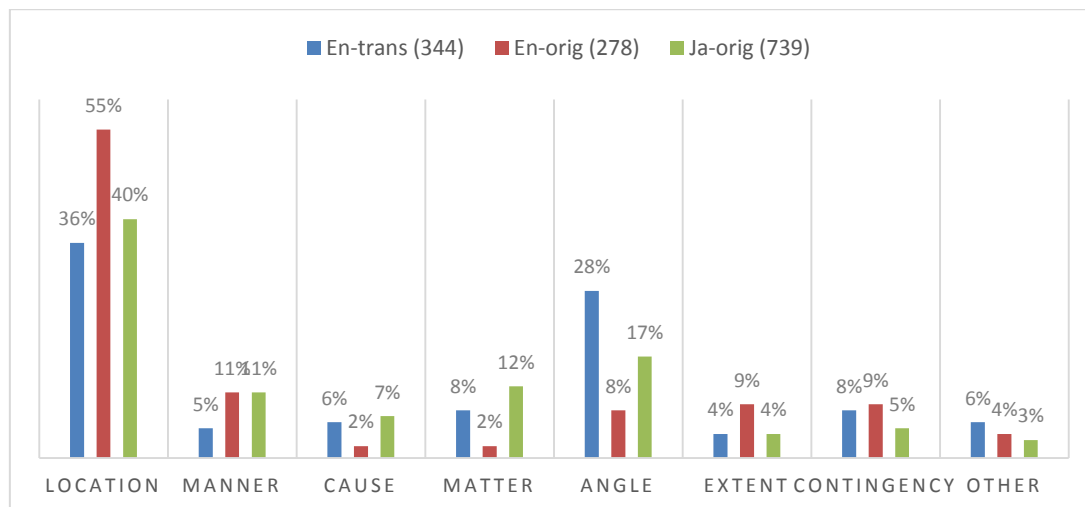


Table 5-2: Adjunct Theme types: significance (En-trans)

	En-orig vs. En-trans		En-trans vs. Ja-orig	
Location	Orig >	+++	=	ns
Manner	Orig >	+++	< Ja	+++
Cause	< Trans	++	=	ns
Matter	< Trans	+++	< Ja	+
Angle	< Trans	+++	Trans >	+++
Extent	Orig >	+++	=	ns
Contingency	=	ns	Trans >	++

En-trans and En-orig show a difference in the way they employ six Adjunct types. Among those, En-trans and Ja-orig turn out to be significantly different with regard to Manner and Angle ($p < .01$), Contingency ($p < .05$) and Matter ($p < .10$), while no significance is found in terms of Location, Cause and Extent. This indicates that En-trans keeps close to the ST with respect to Location, Cause and Extent. Judging also from the way Matter is realised (Figure 5-12), En-trans is moving towards TL norm but it remains closer to Ja-orig.

On the other hand, in terms of Contingency, En-trans and En-orig show no difference, while En-trans is different from Ja-orig ($p < .05$). This suggests En-trans moves away from Ja-orig and closer to the TL norm.

With regard to Manner, En-trans is different from both En-orig and Ja-orig. There is no difference between En-orig and Ja-orig and its realisation is peculiar and difficult to explain.

Angle also draws attention to its realisation in comparison to En-orig and Ja-orig. There is significantly more realisation of Angle in En-trans than En-orig and Ja-orig. If there were no difference between En-trans and Ja-orig, it could have been explained by the direct transfer of the expression in Japanese 'によると *ni yoruto* [according to]' often employed to relate the source of the news. The fact En-trans utilises it more than Ja-orig indicates that En-trans makes use of the expression even when Ja-orig does not. In any case, En-trans is close to the ST in its use of Angle as Theme.

The final variable in this category is 'other' that contains several instances that do not fit in any other categories. Examples (Theme underlined) are;

5-13 Among people aged 100 or older, levels of the hormones are high,
En-trans-Y24

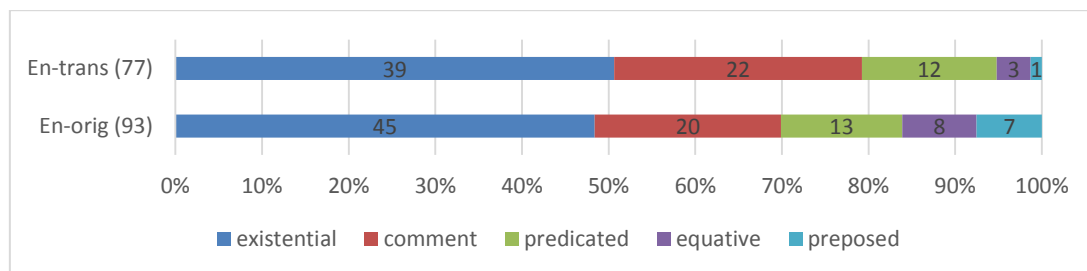
5-14 Among female respondents, 66% backed the approach. En-trans-Y75

There are nine instances of these in En-trans while only one in En-orig and all of them are of this type. Although Mattiessen *et al.* (2010: 223) treat this type as the Absolute Theme (Chapter 4), examples 5-13 and 5-14 appear to function with a circumstantial role rather than outside the experiential structure. Thus, in this project, I take them as Circumstances.

5.2.3.3 Enhanced Theme

Figure 5-10 above shows that Theme is rarely realised as an enhanced type, at 3% or below in both En-trans and En-orig. The datasets differ little (no significance) with regards to different types of enhanced Theme except for Preposed Theme, where they appear to be different, albeit weakly ($df=4$, $p<.10$).

Figure 5-13: Enhanced Theme type (En-trans)



Considering Japanese does not have special thematic structures similar to these available in English, it is interesting that En-orig and En-trans do not differ in this respect, except for Preposed Theme. This set of results is a useful reminder that the absence of specific structures in Japanese does not prevent translators from making use of the enhanced Theme structures available in English to express the same meaning that Japanese conveys using different devices/structures (4.4.3). The only difference is that Preposed Theme is utilised very little in En-trans. An example from En-orig;

5-15 “These little embryos, these are people born with a disease,
En-orig-C42

This part is omitted in the TT and no translation is available for comparison here. However, if it had been translated, it would have been with the particle *wa* such as ‘これらの小さい胎児は [these little embryo WA]’. It is very unlikely for this enhanced structure to be translated as such, though it is not impossible, mainly because it would appear unsuitable for the register of news in

Japanese, even in a direct quotation, as it gives the impression of being theatrical.

Considering translating into English; Preposed Themes are unlikely to be utilised. Firstly, the structure is unlikely to appear in the STs in Japanese, and secondly the noun groups which could be potentially translated into the Predicated Theme such as those above are not marked and, therefore, would be translated with an unmarked Theme. There is a difference between the way the Preposed Theme is realised and other enhanced Themes in En-trans. A possible reason is the different level of markedness of the expression in the register of news in English and in Japanese.

5.2.4 Patterns of Theme selection in En-trans

Figure 5-14 shows where in the system of THEME in English En-trans and En-orig are found to be significantly different. Realisation gaps between En-trans and En-orig are identified in the choices between simple & multiple Theme ($p < .10$), Interpersonal Theme ($p < .05$), Conjunction and textual Adjunct ($p < .02$), unmarked and marked Theme ($p < .02$), types of marked Themes ($p < .02$), and types of Adjunct Themes.

Figure 5-14: Theme realisation gaps between En-trans and En-orig

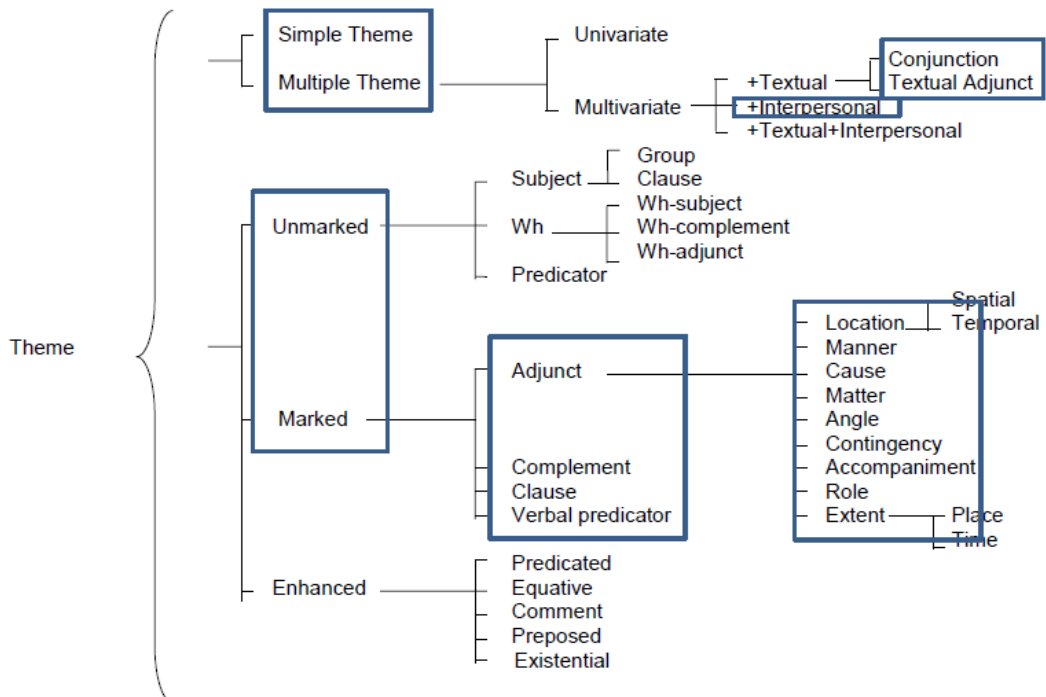


Table 5-3 summarises variables with respect to which En-trans and En-orig are significantly different, indicating the orientation of Theme realisation in En-trans, whether it is similar to the TL (TL normalisation) or to the ST (ST influence).

Table 5-3: Orientation of Theme realisation (En-trans)

	Theme	En-trans	Level of trend
1	simple	TL normalisation	weak
	multiple	TL normalisation	weak
2	interpersonal	ST influence	medium
3	Conjunction	TL normalisation	strong
	Textual Adjunct	ST influence/TL normalisation	strong
4	Unmarked/Subject	TL normalisation	strong
	Marked	No direct comparison made	
5	Adjunct	Neither TL normalisation nor ST influence	
	Complement	TL normalisation	strong
	Clause	TL normalisation	weak
	Verbal Predicator	TL normalisation	weak
6	Location	ST influence	strong
	Manner	Neither TL normalisation nor ST influence	
	Cause	ST influence	medium
	Matter	ST influence/TL normalisation	medium
	Angle	ST influence	strong
	Extent	ST influence	strong
	Contingency	TL normalisation	medium

With respect to what elements are realised as Theme, En-trans indicates influences from both the TL and the STs. In terms of some variables En-trans moves closer to TL norm, while in terms of others it remains close to the ST trend. Hence, both TL normalisation (Baker 1996, Teich 2003) or TL standardisation (Toury 1995) and SL shining through (Teich 2003) or SL influence (Toury 1995) are at work.

TL normalisation is observed towards the left side of the system of Theme (Figure 5-14). Choices between simple and multiple Theme, between unmarked and marked Theme, or among four marked Theme types are examples of this. TL normalisation occurs with general choices of Theme

rather than the more delicate choices, e.g. marked Theme is more general than the Adjunct Theme.

En-trans indicates TL normalisation where there is a systemic difference between English and Japanese, and also where the level of markedness differs. For example, the distinction between unmarked and marked Theme is important in English but Japanese does not distinguish them in the system as such (Chapter 4). In addition, even though all the four categories of marked Theme in English are present in the Japanese system, though the label 'verbal Predicator' represents different elements in each language, those in Japanese are more common as Theme compared to those in English (5.1). The differences in the system and the level of markedness are linked to TL normalisation in the way Theme is realised in En-trans.

TL normalisation in relation to the choice between simple and multiple is not due to differences in the system but partly related to the linguistic difference. The difference is that not all the textual and interpersonal meanings are realised in the initial position in Japanese. Although the reason is neither the system nor the level of markedness, En-trans indicates a process of TL normalisation in terms of the choice between simple and multiple Theme.

On the other hand, ST influence is observed towards the right side of the system of Theme (Figure 5-14), i.e. more delicate/specific choices. ST influence occurs with the interpersonal Theme, in the choice between the Conjunction and textual Adjunct, and with four types of Adjunct Themes.

En-trans indicates ST influence where a similar linguistic feature is available both in English and in Japanese. This means features that are frequently employed in the ST can easily be incorporated into translation, in particular when the level of its markedness is not dissimilar in the two languages. In short, where there is no difference in the system and their markedness in each language is similar, the ST features are transferred in En-trans, creating a gap from En-orig.

This can be related to the tendency Teich observes in the comparison of different languages, stating that 'systems of low delicacy (more general grammatical types) tend to be similar across languages, and systems of higher delicacy (more specific grammatical types) tend to be dissimilar' (2003: 50). Applying her observation to the present study, Theme realisation in En-trans tends to be similar to the TL in the choice of low delicacy, and dissimilar to the TL in the choice of higher delicacy.

In summary, the orientation of Theme realisation in En-trans is a mixture of TL normalisation and ST influence, as expected. What is significant is that I have found that TL normalisation occurs in terms of the general choices of Theme, while ST influence occurs with more delicate choices.

5-3 Theme in Ja-trans (compared with Ja-orig & En-orig)

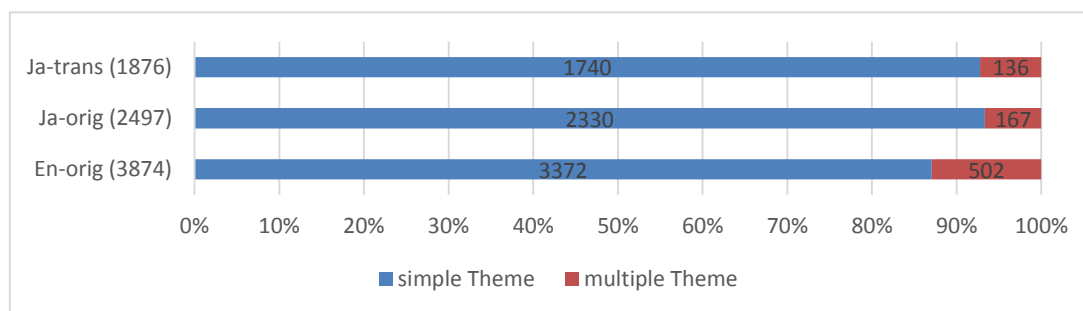
This section presents results of data analysis regarding which elements are selected as Theme in Japanese translation, making comparisons with Ja-orig and En-orig. As with En-trans, comparisons are conducted in a triangular form between Ja-trans and Ja-orig, and between Ja-trans and En-orig.

The following sections discuss specific aspects of Theme realisation from the results table obtained from the UAM CorpusTool. The system networks of THEME in Japanese and in English are presented in Chapter 4 Section 3.4. In addition, in 5.3.4 where a summary of the discussion is offered, the system network of THEME in Japanese is provided again to indicate where translation shifts have occurred.

5.3.1 Simple and multiple Theme

The proportion of simple Themes in Ja-trans and Ja-orig is similar at around 93% and no significant difference is found between them.

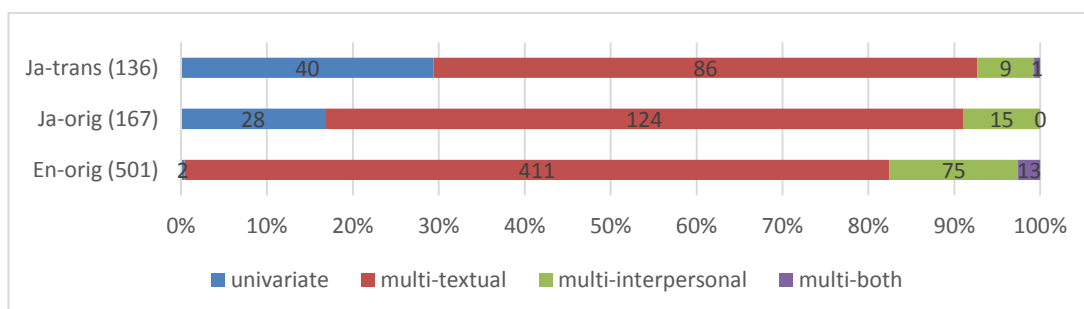
Figure 5-15: Simple & Multiple Theme (Ja-trans)



When Ja-trans is compared with En-orig, a significant difference is found between them ($X^2=41.452$, $df=1$, $p<.01$). This suggests a shift has occurred in translation from English into Japanese, making use of fewer multiple Themes and moving closer to the TL norm.

With regard to multiple Theme types we will first examine the occurrence rate for the univariate Theme against the multivariate.

Figure 5-16: Multiple Theme type (Ja-trans)



The appearance of univariate Theme in Ja-trans is 30% of the multiple Theme, while in Ja-orig contains this stands at 18%. The datasets are significantly different ($X^2=7.18$, $df=1$, $p<.02$). Comparing Ja-trans with En-orig, in which univariate Themes rarely appear, these too are significantly different ($X^2=141.519$, $df=1$, $p<.01$). This means that with respect to the univariate Theme, a shift has occurred during translation into Japanese, moving towards the TL norm. In fact, Ja-trans makes use of it more than the Ja-orig does. It can be taken as a form of confirmation that the univariate multiple Theme is a feature of the Japanese news genre to allow a lot of information to be presented in a limited space (Thomson 2013: 119).

Univariate multiple Themes consist of two circumstantial elements, either two Locations or a combination of one Location and one Extent. The majority of them consist of one temporal Location and one spatial Location, with three instances involving Extent. Two Locations can take the order of either temporal[^]spatial Location (5-16) or spatial[^]temporal Location (5-17), and both are equally frequent in the corpus.

5-16

13 日午後 8 時 40 分ごろ、	大阪市都島区…の市道で	…高齢男性が	…パトカーに	はねられ、
<i>13-nichi gogo 8-ji 40-pun goro</i>	<i>Osaka-shi Miyakojima-ku ... NO shidō DE</i>	<i>... kōrei-dansei GA</i>	<i>... patokā NI</i>	<i>hanerare</i>
13th pm 8:40 about	Osaka-city Miyakojima-ward ... NO city road DE	elderly man GA	police car NI	hit-psv-SUSP
On 13th around 8:40 pm, an elderly man was hit by a police car on the city road in ... Miyakojima-ward, Osaka.				
Circ.: temp	Circ.: spatial	Medium	Actor	Process
Theme (univariate)		Rheme		

5-17

新潟県…雑木林で	7日夕	…女性の遺体が	見つかり、
<i>Niigaga-ken ... zōkibayashi DE</i>	<i>7-ka yū</i>	<i>... josei NO itai GA</i>	<i>mitsukari</i>
Niigata-pref, ... wood DE	7th evening	Woman NO body GA	find-psv-SUSP
In the wood of ... Niigata, a woman's body was found in the evening of 7th,			
Circ.: spatial	Circ.: temp	Medium	Process
Theme (univariate)		Rheme	

Ja-orig-M59

A close observation reveals that Ja-orig employs the order of temporal[^]spatial more (19 instances out of 28), while spatial[^]temporal order appears more in Ja-trans (35 instances out of 40). Ten out of 19 instances which take the temporal[^]spatial order in Ja-orig state a specific time such as 13日午後8時40分ごろ [On 13th around 8:40 pm] as in 5-16. On the other hand, when the temporal Location comes after the spatial, it never indicates a specific time, the most specific being ‘夕 [evening]’ or ‘早朝 [early morning]’ in the corpus.

In Ja-trans, five instances that have the order of temporal[^]spatial indicate a tendency of placing the time before the place when the time is in focus. Example 5-18 appears towards the end of the news article which concerns a convicted murderer, a Kennedy relative is freed on bail. It describes when the murder occurred.

5-18

1975年、	コネティカット州の高級住宅街で
<i>1975-nen</i>	<i>Konetikatto-shū NO kōkyū jūtakugai DE</i>
1975 year	Connecticut state NO high class residential area DE
Circ.: temporal	Circ.: spatial

Ja-trans-C16

The analysis results suggest univariate Themes are used differently in Ja-trans and Ja-orig. This will be discussed in relation to the location of the text in which they appear in Chapter 6.

5.3.2 Interpersonal and Textual Theme

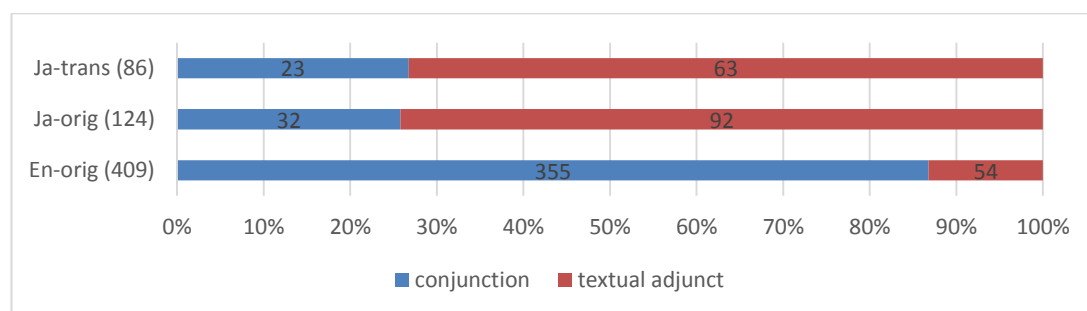
Interpersonal Themes appear just under 10% in both Ja-trans and Ja-orig (Figure 5-16) but the two datasets are found to be marginally different in this

account $X^2=6.894$, $df=2$, $p<.05$). The relative infrequency of interpersonal Themes is similar to English. The low percentage may relate to the idea that in Japanese ‘the lexicogrammar of the interpersonal metafunction of MOOD and MODALITY tends to utilise at the end of the clause where the process is most typically located’ (Thomson 2013: 114). Interpersonal Themes that appear in initial position in the corpus include ‘特に [in particular]’ and ‘例えば [for example]’.

In terms of the realisation of the interpersonal Theme, Ja-trans indicates a smaller percentage of it than does En-orig. The two datasets are found to be different from each other ($X^2=148.649$, $df=2$, $p<.01$). Ja-orig and En-orig are marginally different in this regard ($df=2$, $p<.05$), with Ja-orig displaying a smaller percentage of it than En-orig (5.1.3). Thus, a shift has occurred in Ja-trans during translation and it can be said to be moving away from En-orig and closer to the TL norm.

Textual Themes appear more frequently than interpersonal Themes do in both Ja-trans and Ja-orig. In this account too, no significant difference is found between them. Of the two options of the textual Theme, textual Adjuncts are employed more frequently than Conjunctions in Ja-trans as well as in Ja-orig. The two datasets do not differ significantly in terms of their use. On the other hand, compared with En-orig, Ja-trans turns out to be significantly different ($X^2=137.860$, $df=1$, $p<.01$).

Figure 5-17: Textual Theme type (Ja-trans)



This suggests a shift has occurred during translation from English into Japanese, making use of textual Adjunct more instead of Conjunctions that are employed in En-orig. This implies that Ja-trans moves closer to the TL norm.

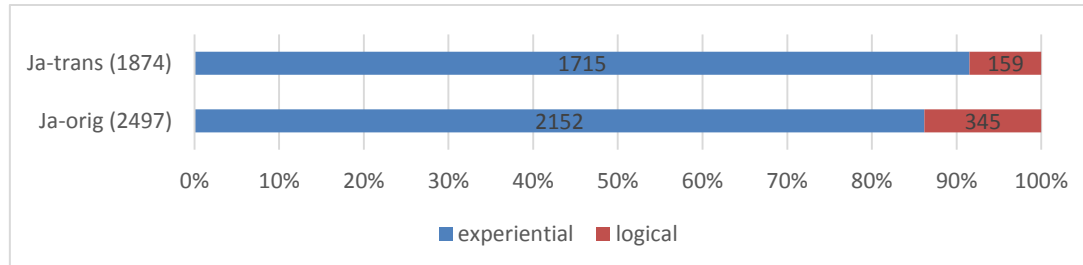
5.3.3 Ideational Theme

The first distinction in this category is between the absolute Theme and the relative Theme. Absolute Themes are very rare with under 1% realisation rate

in both Ja-trans and Ja-orig. However, a significant difference is found between the datasets in this respect, albeit at a low level ($X^2=3.02$, $p<.10$).

More than 99% of Themes in Japanese are relative Themes, which are either experiential or logical Theme.

Figure 5-18: Relative Theme type (Ja-trans)



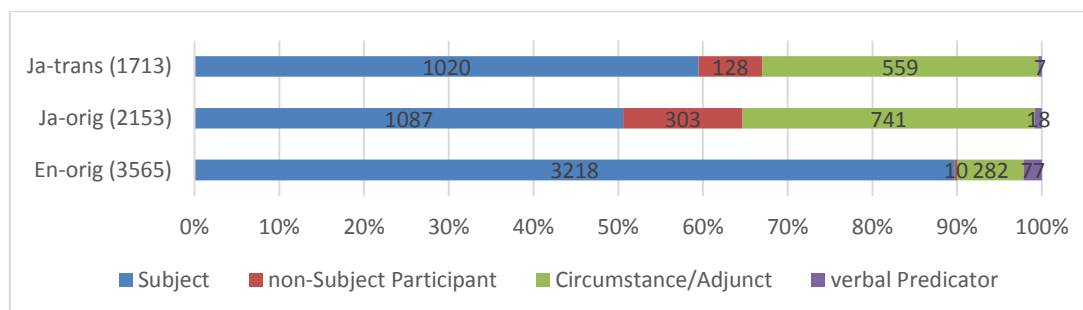
Both Ja-trans and Ja-orig make use of the experiential Theme more frequently than the logical Theme but Ja-trans does so more than Ja-orig. The datasets are found to be significantly different ($df=1$, $p<.02$).

In En-orig the occurrence rate for the logical (clause) Theme accounts for 4% of all the Themes and on this account Ja-trans and En-orig are significantly different ($X^2=48.770$, $df=1$, $p<.01$).

There are three types of experiential Themes; Participant, Circumstance and verbal Predicator. Participant is the most likely candidate for Theme in both Ja-trans and Ja-orig at over 60%, followed by Circumstance at around 35%. However, the two datasets turn out to be different in terms of how the Participant Theme and the Circumstance Theme are realised ($p<.10$).

In order to facilitate a direct comparison between Japanese and English, the category of the Participant in Figure 5-19 is divided into Subject and non-Subject.

Figure 5-19: Experiential Theme type (Ja-trans)



The realisation of different types of experiential Theme is discussed in the next section.

5.3.3.1 Participant Theme

The Participant Theme can be either a Subject or a non-Subject Participant. Between them, the Subject is selected as Theme 89% of the time in Ja-trans, while being 78% in Ja-orig. In this respect, the datasets are found to be significantly different ($X^2=51.92$, $df=1$, $p<.02$).

Figure 5-19 indicates a stronger preference for the Subject Theme to the non-Subject Participant (Complement) Theme in En-orig. Statistics show that Ja-trans is significantly different from En-orig on this regard ($X^2=314.613$, $df=1$, $p<.01$).

Ja-trans is significantly different from both En-orig and Ja-orig in terms of the choice between the Subject and non-Subject Participant as Theme. However, proportionally Ja-trans appears more similar to Ja-orig with respect to Subject and non-Subject Participant Theme, and hence it can be said to be moving away from En-orig and closer to the TL norm.

Focusing on the non-Subject Participant Themes in Japanese, roughly a quarter of them in Ja-trans are topicalised and marked by 'wa', while 21% of them are in Ja-orig. The rest and the majority of them are not topicalised but nonetheless realised as Theme by appearing in initial position due to Subject ellipsis.

No significant difference is found between Ja-trans and Ja-orig whether non-Subject Participant Themes are topicalised or not. This implies two things; Subject ellipsis is occurring in Ja-trans in the same way as in Ja-orig, and topicalisation of the non-Subject Participant is at work in Ja-trans in the same way as in Ja-orig. In short, Ja-trans follows the TL tendency in terms of how non-Subject Participants are realised as Theme.

5.3.3.2 Circumstance Theme

Regarding the overall realisation of Circumstance Theme, Ja-trans and Ja-orig are different, at the low value of $p<.10$. However, when it comes to what types of Circumstance are realised as Theme, the two datasets are significantly different with regards to five variables at different levels of significance.

Figure 5-20: Circumstance Theme type (Ja-trans)

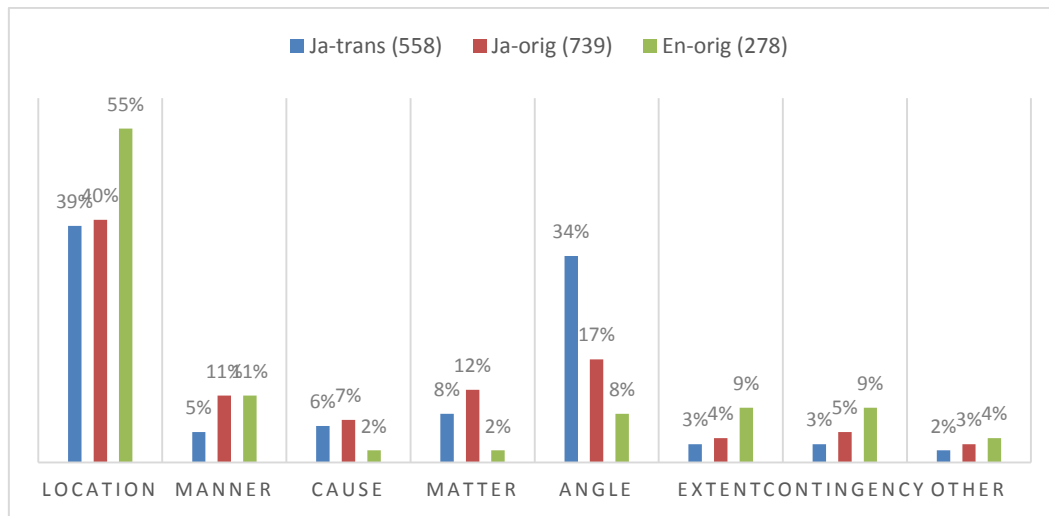


Table 5-4: Circumstance Theme types: significance (Ja-trans)

	Ja-orig vs. Ja-trans		Ja-trans vs. En-orig	
Location	=	ns	< En	+++
Manner	Orig >	+++	< En	+++
Cause	=	ns	Trans >	++
Matter	Orig >	+++	Trans >	+++
Angle	< Trans	+++	Trans >	+++
Extent	Orig >	+	< En	+++
Contingency	Orig >	+	< En	+++

In terms of the realisation of Location and Cause, Ja-trans and Ja-orig do not differ from each other. In this regard, however, Ja-trans and En-orig are significantly different, although the level of significance differs, at $p < .05$ for Cause and $p < .01$ for Location. These two variables are clear examples of Ja-trans moving towards TL norm.

With respect to Extent and Contingency too, Ja-trans is moving closer to the TL. The difference observed between Ja-trans and Ja-orig is smaller than that between Ja-trans and En-orig, i.e. the former pair is weakly different ($p < .10$), while the latter is significantly different ($p < .01$). In addition, the way Matter is realised suggests Ja-trans are moving closer to TL norm, although the gap between Ja-trans and Ja-orig in this respect is more significant ($p < .02$), stronger than for Extent and Contingency.

Angle is an interesting case as it is the only Circumstance type that appears more often in Ja-trans than in Ja-orig. It also appears in Ja-trans more frequently than En-orig. It is possible to interpret the realisational tendency of this variable as moving towards TL norm, however it is employed much more in Ja-trans than Ja-orig. This is likely to be caused by the particular expression in Japanese used to inform of the news source, 'によると *ni yoruto* [according to]'. The fact that Angle appears more in Ja-trans suggests that the expression is employed in Ja-trans even when the ST does not make use of the equivalent expression 'according to'. For example, verbal Processes with an unmarked Subject Theme in English can be translated into Japanese with a Circumstance Theme 'according to'.

With regard to Manner, Ja-trans is different from both Ja-orig and En-orig. Since there is no difference between Ja-orig and En-orig in this regard, its realisation is neither moving towards the TL nor remaining close to the ST.

Finally, Circumstance in Japanese has a category known as 'other', just as English does, which contains several examples that do not fit in any other categories. In fact, most examples in this category are similar in the meaning they create.

5-19 619 法人のうち 619 *hōjin no uchi* [among the 619 companies]

5-20 原告 99 人のうち *genkoku 99-nin no uchi* [among the 99 prosecutors]

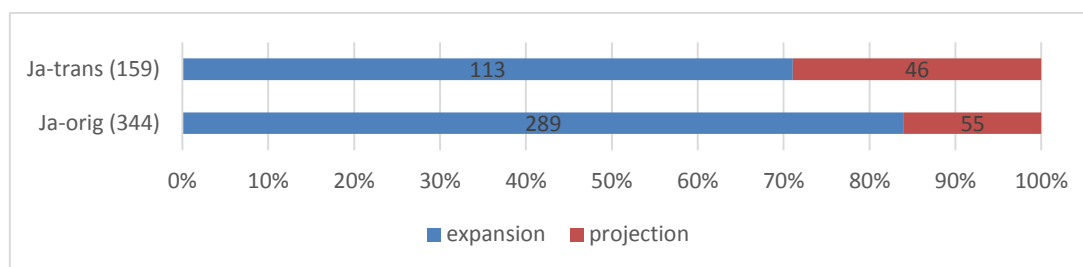
5-21 このうち *konouchi* [among these]

There are 18 instances of 'other' in Ja-orig and three in Ja-trans, 17 out of 21 are of this type. The raw count in Ja-orig is more than those of Role or Accompaniment, and they may merit a category of their own, which could be labelled for example as 'range'.

5.3.3.3 Logical Theme

A logical Theme in Japanese is either an expansion or a projection. In terms of the choice between them, Ja-trans and Ja-orig turn out to be significantly different ($X^2=11.79$, $df=1$, $p<.02$).

Figure 5-21: Logical Theme type (Ja-trans)



Although both datasets share the trend of making use of expansion more than projection, the proportion of projection accounts for more in Ja-trans (30%) than in Ja-orig (17%). This may be influenced by a combination of the frequent appearance of projection in English and the linguistic restriction of Japanese that the dependent β clause, including reported clauses if the Sayer/Sensor is ellipsed, are placed in initial position.

A typical example of the projected clause Theme of locution is offered in 5-22.

5-22

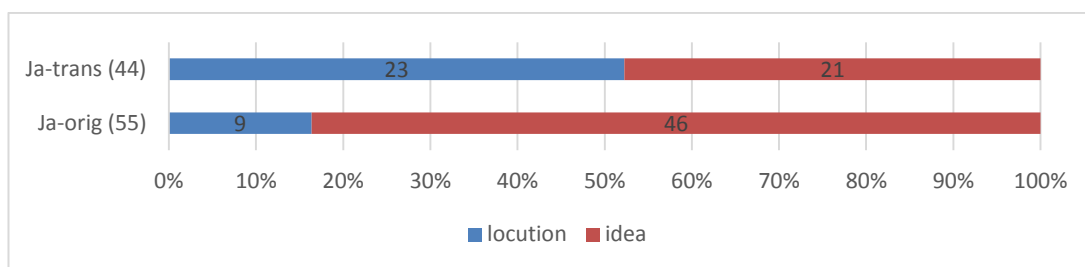
//学生寮の自室で	リシンを作ったと//	供述している。
//Gakuseiryō NO jishitsu DE	rishin O tsukutta TO//	kyōjutsushiteiru
//Student hall NO own room DE	ricin O make-pst-PROJ//	state-ASP
(He) states (he) generated ricin in (his) own room in the student hall.		
Projected: reported clause		Projecting
Theme		Rheme

Ja-trans-C50

The reported clause combines clauses as hypotaxis and this can be a Theme in Japanese (Chapter 4), unlike in English. Projected clauses can appear in initial position in English but they are direct quotations, which combine clauses as parataxis, and are also not Theme. This systemic difference, as well as the frequent Subject ellipsis such as in 5-22, is likely to have influenced projection Theme realisation in Ja-trans.

In fact, projected clauses of locution account for 50% in Ja-trans in the choice between locution and the idea, and only for 16% in Ja-orig. Ja-trans and Ja-orig were significantly different in this regard ($X^2=13.91$, $df=1$, $p<.02$).

Figure 5-22: Projection Theme type (Ja-trans)



Projected clauses are generally marked by the particle 'to' (Chapter 4) and combined with a verbal Process (locution) or a mental Process (idea). However, it is not impossible for them to be accompanied by material Processes and they cannot normally project locutions (Teruya 2006: 428). This occurs with ideas too and sometimes poses difficulties in analysis. The problem mainly relates to the boundary between Theme and Rheme in relation to the particle 'to'. The boundary in example 5-23-1 is after the particle 'to', taking the whole projection as Theme.

5-23-1

//新規受注は 60.6 と//	前月の 60.5 から	わずかに上昇した。
//Shinkijuchū WA 60.6 TO//	zengetsu NO 60.5 KARA	wazukani jōshōshita
//New order WA 60.6-PROJ//	previous month NO 60.5 KARA	slightly up-PST
The new orders index is 60.6, slightly up from 60.5 in the previous month.		
Projected clause		
Theme	Rheme	

Ja-trans-R25

On the other hand, the boundary can be situated after the Topic, e.g. the Theme conflates with the Topic in example 5-23-2.

5-23-2

新規受注は	60.6 と	前月の 60.5 から	わずかに上昇した。
Shinkijuchū WA	60.6 to	zengetsu-no 60.5 kara	wazukani jōshō-shita
New order WA	60.6 PROJ	previous month NO 60.5 KARA	slightly up-PST
The new orders index slightly went up to 60.6 from 60.5 in the previous month.			
Token(Medium)	Value	Circumstance	Process: material
Theme	Rheme		

Ja-trans-R25

This study employs the second analysis (5-23-2) taking ‘*shinkijuchū wa* [the new orders]’ as a co-referent within the T-unit to be consistent with the method for identifying Theme in hypotactically related clause complexes.

5.3.4 Patterns of Theme selection in Ja-trans

Figure 5-23 indicates where Ja-trans and Ja-orig are found to be significantly different in the system of THEME in Japanese. Realisation gaps between Ja-trans and Ja-orig are identified in the choices between univariate or multivariate Theme, between absolute and relative Theme, between experiential and logical Theme, and with Subject Theme and non-Subject Theme. Also significantly different are Circumstance Themes of Manner, Matter, Angle ($p < .02$), Extent and Contingency ($p < .10$), in the choice of expansion or projection as well as locution or idea.

Figure 5-23: Theme realisation gaps between Ja-trans and Ja-orig

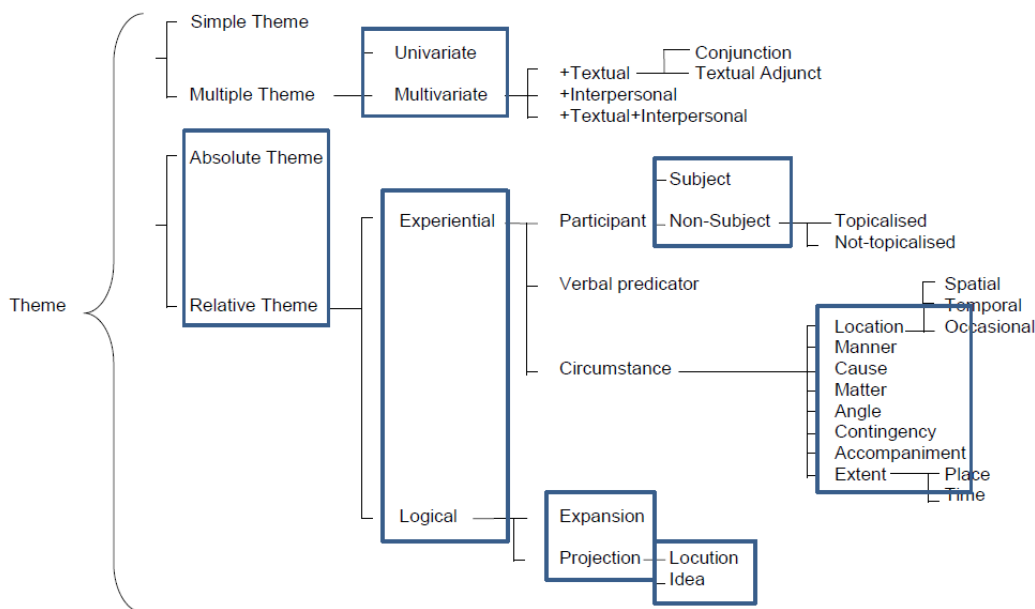


Table 5-5 summarises the orientation of Theme realisation in Ja-trans, whether it is similar to the TL (TL normalisation) or to the ST (ST influence). Variables included are mostly with those which Ja-trans and Ja-orig are significantly different but the table also includes Circumstances of Location and Cause with which the datasets do not differ but nonetheless a process of TL normalisation is observed.

Table 5-5: Orientation of Theme realisation (Ja-trans)

	Theme	Ja-trans	Level of trend
1	Univariate	TL normalisation	strong
	Multivariate	TL normalisation	strong
2	Interpersonal	TL normalisation	medium
3	Absolute	No direct comparison is made with En-orig but TL normalisation is inferred	
	Relative		
4	Experiential	No direct comparison made with En-orig	
	Logical	TL normalisation	strong
5	Subject	TL normalisation	strong
	Non-Subject	TL normalisation	strong
6	Expansion	No direct comparison is made with En-orig but ST influence is inferred	
	Projection		
7	Locution	No direct comparison is made with En-orig but ST influence is inferred	
	Idea		
8	Location	TL normalisation	strong
	Manner	Neither TL normalisation nor ST influence	
	Cause	TL normalisation	medium
	Matter	TL normalisation	strong
	Angle	TL normalisation	strong
	Extent	TL normalisation	medium
	Contingency	TL normalisation	medium

Ja-trans displays influences from the TL more than from the STs, i.e. TL 'normalisation' (Baker 1996, Teich 2003) or TL 'standardisation' (Toury 1995: 274-279) more than the SL 'shining through' (Teich 2003) or ST 'influence' (Toury 1995: 274-279).

On the basis of my analysis, TL normalisation in Ja-trans occurs for various reasons. It occurs where there is a system difference in Japanese and in English, for example in choices between absolute and relative Theme, and

between experiential and logical Theme. It also occurs where there is a difference in terms of the level of markedness, for example, in choices between univariate and multivariate Theme, or between the Subject and non-Subject Participant Theme. Circumstance Theme types experience a process of TL normalisation too, although the systems are similar with equivalent options.

In short, TL normalisation is observed both with the general and delicate choices of Theme unlike in En-trans, which indicate TL normalisation with general choices and ST influence with delicate choices.

Of particular interest among the instances of TL normalisation in Ja-trans are univariate Themes and Circumstance of Angle. They feature commonly in news articles in Japanese which are usefully utilised in translation.

Rare ST influence in Ja-trans is observed in the preference of Projection Theme over expansion, as well as the preference of the locution Theme instead of idea. This reflects the general frequency of the projected clause in the English news (Halliday & Matthiessen 2004). It also reflects the linguistic feature of Japanese that the dependent β clause, including reported clauses if the Sayer/Sensor is ellipsed, are placed in initial position, which English does not share.

Theme realisation in Ja-trans is clearly TL oriented. Processes of TL normalisation are observed from the general to more delicate choices in the system of Theme. ST influence also occurs in Ja-trans but it is limited to the choices of projection over expansion, and of locution over idea.

5.4 Shifts in En-trans and Ja-trans

This chapter has explored Theme realisation in En-trans and Ja-trans and patterns in terms of frequency have been discovered. Theme realisation in En-trans displays a mixture of TL normalisation and ST influence. While we would expect this to occur, I have found specifically where and how TL normalisation and ST influence occur. With regard to the general choices of Theme, i.e. the left-side of the system network of THEME, En-trans is similar to En-orig, suggesting shifts have occurred during translation. With regard to more delicate choices of Theme, i.e. the right-side of the system network, however, En-trans is close to Ja-trans, suggesting that either shifts have not occurred or that the movement towards the TL is not strong.

In contrast to En-trans, Theme realisation in Ja-trans has shown a strong tendency of TL normalisation. With respect to both general and delicate

choices of Theme, i.e. across the system network of THEME, Ja-trans is close to Ja-orig, indicating shifts have occurred during translation. It is clearly moving towards TL norm in terms of Theme realisation.

Chapter 6

Textual coherence of news and translation shift

This chapter concerns textuality of translated news from the viewpoint of genre and aims to answer research questions 3-5 and 6-8 which are set out in Chapter 1. Questions 3-5 concern translation shifts in terms of information content, which falls in the field of discourse in the stratified model of language. Questions 6-8 concern translation shifts in terms of information flow, focusing on thematic progression, the mode of discourse. The notion of text structure will be incorporated as an important aspect of 'texture' (Halliday 1978: 136), i.e. readability and acceptability of translated news.

Reflecting the two areas of investigation as described above, this chapter is divided into two sections. The first part presents the results of analysis of translation shifts to identify the ranks, types and locations of the information that are added, deleted or moved (6.1). The second part presents the results of analysis of translation shifts to identify what types of Theme are realised at specific locations within text, i.e. at the beginning of the text nucleus or satellites (6.2).

Section 6.2 firstly describes Theme realisation at the beginning of the text nucleus and satellites in En-orig and in Ja-orig (6.2.1) to identify contrasts between them as possible indicators of shifts. Then it continues to describe Theme realisation at the beginning of the text nucleus and satellites in En-trans and in Ja-trans in comparison to En-orig and to Ja-orig (6.2.2) to investigate where thematic shifts might occur and whether Theme realisation in translation adheres to the ST or moves closer to the TL convention (6.2.3).

In addition to the questions on translation shifts, this chapter also describes contrastive features of news articles written in Japanese (6.2.1), addressing research questions 9-10, set out in Chapter 1.

6.1. Text structure and field of discourse (information content)

The focus of this section is the relationship between the text structure and information that is added or deleted in news translation. Moves of information from one place in the ST to another in the TT are also dealt with here, bearing in mind that they are also relevant to information flow.

6.1.1. Frequency of additions, deletions and moves in translation

Additions, deletions and moves occur to almost all the translations in the corpus used here. Tables below illustrate the percentage and number of translated news articles out of 170 in En-trans and 170 in Ja-trans in which additions, deletions or moves occur (Table 6-1), and how likely they co-occur in one text (Table 6-2).

While both En-trans and Ja-trans display high occurrence rates for each translation shift, percentages for Ja-trans are slightly higher than those for En-trans.

Table 6-1: Number of texts involving additions, deletions or moves

	Additions		Deletions		Moves	
En-trans (170)	92% (157)		90% (153)		85% (144)	
Yomiuri/Mainichi	77	80	73	80	63	81
Ja-trans (170)	96% (163)		100% (170)		89% (152)	
CNN/Reuters	83	80	85	85	80	72

Additions, deletions and moves are often employed together. Table 6-2 shows that Ja-trans displays a higher percentage for deploying three of them together, compared to En-trans. However, when the percentages of deploying three types and deploying two types are combined, both En-trans and Ja-trans reach over 97%, with Ja-trans slightly higher at 99%.

Table 6-2: Number of texts in which additions, deletions or moves co-occur

	All 3 types		2 types		1 type		0	
En-tran (170)	71% (120)		26% (44)		3% (5)		0.5% (1)	
Yomiuri/Mainichi	47	73	34	10	3	2	1	0
Ja-tran (170)	87% (147)		12% (21)		1% (2)		0% (0)	
CNN/Reuter	78	69	7	14	0	2	0	0

Based on the analysis results, trends for frequently employing additions, deletions or moves each as a strategy in news translation, as well as employing several of them together will be examined. Following sections

examine additions, deletions and moves individually, before which, a note on the data that will be used for analysis in this section.

As mentioned in Chapter 4, there are articles that do not fit the description of hard news. As a genre, hard news articles are typically written adopting the generic structural convention of the said genre, conforming to the orbital structure consisting of a nucleus and a set of satellites (White 1998). Eleven articles in En-orig do not conform to it. Firstly, they do not have a text nucleus where the most newsworthy information is provided. Secondly, they narrate an event chronologically, instead of consisting of satellites that specify elements in the text nucleus. The major problem is that an absence of the nucleus and satellites structure means they will not be segmented into text units, which are needed to investigate if additions occur at the rank of text units, or what types of Theme are realised at the beginning of text units. Therefore, since the concept of generic structure is of essential importance for the purpose of investigation in this chapter, these articles are excluded from analysis here, both in 6.1 and 6.2 in order to maintain the reliability of analysis.

In addition, there are 16 articles in Ja-orig, all from *Mainichi Shimbun*, I have decided to exclude from analysis in this chapter. They do conform to the orbital structure of news but are followed by extra passages or a list below the name of the writer, i.e. after the news article itself is complete. Although these extra sections can be easily separated in the ST, information provided there is translated and appears in the TT. This potentially influences the results of data analysis, and as such, they are also excluded from analysis.

As a result, the texts used for analysis in this chapter were 154 texts in Ja-orig and En-trans, 85 from *Yomiuri* and 69 from *Mainichi*, and 159 texts in En-orig and Ja-trans (78 from CNN and 81 from Reuters).

6.1.2. Rank and type of additions, deletions and moves in translation

6.1.2.1. Additions

Adding important background information in news translation is a common approach (Bielsa & Bassnett 2009: 64) to make the translated news to suit the needs of the target readers. I will investigate this empirically, at scale and systematically to evidence and to add nuance to the claim cited above.

The table below shows the number of texts that involve additions according to the semantic rank.

Table 6-3: Number of texts involving additions according to the semantic rank

	Text unit	Sequence (T-unit)	Figure (Clause)	Element
En-trans-Yomiuri (85)	60% (51)	35% (30)	39% (33)	71% (60)
En-trans>Mainichi (69)	27% (16)	16% (11)	54% (37)	83% (57)
Ja-trans-CNN (78)	10% (8)	14% (11)	12% (9)	91% (71)
Ja-trans-Reuters (81)	11% (9)	16% (13)	17% (14)	88% (71)

At the rank of text unit En-trans displays more instances of additions compared to Ja-trans. The number of texts involving additions at the rank of text unit in En-trans-Yomiuri and En-trans>Mainichi differ considerably, which is caused by *Yomiuri* frequently adding one or more text units at the end of text to offer more explanations such as contextualisation or elaboration as well as occasional appraisal. A text unit or more are added to 51 articles from En-trans-Yomiuri at the end. *Mainichi* also makes use of the same strategy but not to the extent of *Yomiuri*.

A text unit can be added at the beginning of the text as the text nucleus, replacing the original nucleus. Both *Mainichi* and *Yomiuri* include three examples. These nuclei provide a general form of the news lead compared to the original's relatively specific and detailed ones. Example 6-1 presents a text nucleus which has replaced the original one, now functioning as a satellite.

6-1

Nucleus	<i>Trainee Buddhist monk beaten by supervisor at temple</i> <i>ZENTSUJI, Kagawa -- <u>A supervising Buddhist monk physically abused a trainee monk at a temple here in mid-October, it has been learned.</u></i>
Satellite	<u>The supervising monk in his 30s</u> hit and kicked a trainee monk in his 20s, who had come from Nara Prefecture to Zentsu-ji, the head temple of the Shingon sect's Zentsuji faction in Zentsuji, Kagawa Prefecture, for training. <u>The trainee monk</u> suffered injuries as a result. <u>Officials of the temple</u> acknowledged that the trainee monk was physically assaulted and offered an apology. <u>The trainee monk</u> , however, filed a damage report with Kagawa Prefectural Police.

The practice of adding a text unit to replace the original text nucleus in the ST is not observed in Ja-trans. However, it is worth mentioning that these do occur in the articles that have been excluded from analysis, due to the lack of the structure of a text nucleus and a set of satellites. An example of the opening paragraph of such an article is presented here.

6-2 Congratulations, the \$10 lottery ticket you just bought in our convenience store is worth \$1,000! Here's your money! En-orig-C15

This paragraph is deleted in the translated text and a more news-like lead is added, presenting a summary of the event, i.e. the shop owner and assistant are arraigned on grand larceny charges, accused of trying to cheat a lottery winner, whose ticket was worth \$1 million. Other than those articles excluded from analysis, there are no instances of these.

Additions at the rank of text units can occur in the middle of the text too. Examples of these show they present a situation following the main news event or comments from those involved in the event, providing more information that the translator considers useful for the target reader. Several examples of these include relatively specific information on the news topic, which suggests the TT is based on more than one ST.

Next, additions at the rank of sequence, which is the T-unit at the level of grammar occur throughout the text, functioning similarly to the added text units as explained above. Occasionally, a sequence is added in the text nucleus, following the first sequence and not at the opening but, in which case it often adds information to the first. For example, a sequence of 'she is the first female US ambassador to Japan.' is added in the nucleus in Ja-trans-C7 whose opening sentence (sequence) reports that Caroline Kennedy has been confirmed as a new US ambassador to Japan. It could have been a satellite but is placed in the nucleus, which reflects that the translator considers this piece of information important.

This article involves other added sequences in the middle of the text, presenting backgrounds of Kennedy and explaining the importance of good US-Japan relationship. These added sequences reflect the strong relevance of the event to the reader of translated news. The translator must have considered it useful to provide more information than is presented in the ST to maintain and strengthen the relevance of the news to the target readership.

Other than the matter of relevance, reasons for adding sequences include providing important background information that may not be known to the target reader. In a report that the majority of Americans favour legalising

marijuana (Ja-trans-R6), a sequence is added in translation which states the US federal government places marijuana illegal. Then the text continues to report that Washington and Colorado legalised it in 2012, which is in the ST. The function of the added sequence in this instance is to provide the background knowledge for the reader to understand the immediately surrounding co-text; legalisation of marijuana in Washington and Colorado.

This reasoning can also be applied to En-trans examples of added sequences. For example, En-trans-Y30 reports on the reform of English education, in which sequences are added to explain the background to it or ways of doing it to the target reader who are unlikely to be as familiar as the target readership of the original text about the issue. Additions of sequences occur at a place in the text where such information is considered necessary. Examples, in particular from En-trans show they appear throughout the text.

If additions at the rank of sequence are considered to take place to provide the readership with assistance to understand immediate surroundings of the added sequences, additions at the rank of figure, clause at the level of grammar, is assumed to do the same but with more limited surroundings; added figures help the reader to understand the information presented within the sequence in which the figure in question appears. Representative examples include:

- | | |
|---|--------------|
| 6-3 after leaving the taxi cab without paying | En-trans-M4 |
| 6-4 thereby helping to lessen the financial impact of... | En-trans-M11 |
| 6-5 as the raw data mostly consists of fragmental information | En-trans-Y2 |

Through these added figures, possible knowledge gap is filled or implicit information in the ST is made more explicit in the TT such as the steps of events or consequences of them.

The final rank is the element, phrase/group/word/morpheme at the level of grammar. Additions of this rank occur frequently both in En-trans and Ja-trans. Most examples are used to clarify nearby items or to help the reader understand the clause clearly, in which the added elements are present. These appear throughout the text. Typical examples from En-trans are 'in Okinawa', 'off Kanagawa', 'controversial (bill)' and from Ja-trans are 'in California', 'in the US', 'a private company', 'a research organisation' as well as an equivalent monetary worth in Japanese yen to the amount presented in US dollars, among others.

Other examples of additions include expressions of discourse such as ‘for instance’, ‘but’, ‘in comparison’ in En-trans and ‘with regard to’, ‘on (this matter)’ in Ja-trans. Elements relating to genre conventions are also common, especially in Ja-trans. A large number of the Ja-trans data have ‘米 [US, morpheme of 米国]’ at the beginning of the text, clarifying that the news is from the USA. Because of the almost fixed location, it is likely to be a genre convention. In addition, people’s names have affixes such as ‘さん or 氏 [Mr. Mrs. Ms]’ or ‘容疑者 [suspect]’ depending on the situation.

In summary, information is added to enable the translated texts function as a complete whole in the new context. It fills the knowledge gap of the reader, assists them to clearly understand the reported news, and adjusts the original text to create an acceptable discourse and to follow the conventions of the genre in the target culture. Additions at each rank provide assistance for the reader to understand the text locally from the rank of element and more globally towards the rank of text unit. While local assistance appears throughout the text, additions of text unit tend to appear towards the end of text as well as at the beginning to a certain extent.

As a final note to this section, a relatively high number of texts involving additions at higher ranks in En-trans compared to Ja-trans suggests two things. Firstly, adding information to components of news is chosen over deleting those, and secondly, more global assistance is required or preferred for texts in En-trans to be readable and/or acceptable to the target readership. Since main target readers of En-trans-*Yomiuri* and *Mainichi* are likely to be residents in Japan (*Yomiuri Shimbun* 2016: 8), the relevance of Japan’s national news to them is stronger than that of US’s national news to Japanese who are likely to be in Japan. This explains why additions are chosen over deletions; in order to translate the news of possibly unfamiliar and complex issues that may be relevant to the target reader, information at a higher semantic rank is necessary.

6.1.2.2. Deletions

Bielsa & Bassnett (2009:64) state ‘unnecessary information’ is eliminated in news translation, explaining that unnecessary means redundant, too detailed or too specific. In the same manner as before, this section discusses the phenomenon of deletion in news translation empirically, at scale and systematically.

The table below shows the number of texts that involve deletions according to the semantic rank. Ja-trans displays higher numbers of texts undergo

deletions than in En-trans. This is opposite from the trend for additions, which take place more frequently in Ja-trans.

Table 6-4: Number of texts involving deletions according to the semantic rank

	Text unit	Sequence (T-unit)	Figure (Clause)	Element
En-trans-Yomiuri (85)	22% (19)	24% (20)	20% (17)	75% (64)
En-trans>Mainichi (69)	22% (13)	32% (19)	38% (23)	87% (60)
Ja-trans-CNN (78)	62% (48)	77% (60)	78% (61)	95 % (74)
Ja-trans-Reuters (81)	90% (73)	74% (60)	73% (59)	95% (77)

Majority of the examples of deletion at the rank of text unit in the En-trans are located at the end of the text. Most of them function in the original to contextualise the event in news or present comments from people involved in the event. The tendency of deleting the final text units is shared in Ja-trans and in fact, there are more examples, one of which is presented here.

6-6 A president's standing in public opinion polling is considered a gauge of his clout. En-orig-C12

Deletions at the rank of sequence occur more frequently in Ja-trans than in En-trans. A typical example is projection.

6-7 The Tuesday shooting is being investigated, the sheriff's office said. En-orig-C5

Another example 6-8 presents a paragraph in which a sequence deletion occurs. The first sequence remains in translation but the second (underlined) is deleted.

6-8 CVS said it will lose about \$2 billion in annual sales and between 6 and 9 cents of profit per share this year. Analysts expect the company to report 2014 revenue of \$132.9 billion and a profit of \$4.47 per share, according to Thomson Reuters. En-orig-R32

Example 6-8 shows that a sequence deletion does not prevent the reader from understanding the news although it influences the level of understanding due to the deletion of detailed information. Deletions at the rank of sequence take place if the surrounding sequences provide enough information for the reader to understand the portion of the text without the deleted sequence.

Occasionally, a portion of text in which successions of sequences are involved is deleted. This suggests the text unit, in which deleted sequences exist, is judged understandable without the deleted information as 6-8 shows. Example 6-9 shows a more extensive version of 6-8. The first sequence remains in translation but the second and the third (underlined) are deleted.

6-9 The latest data breach comes after the FBI warned retailers last month to prepare for more cyberattacks after discovering about 20 hacking cases in the past year involving the same kind of malicious software used against Target Corp over the holiday shopping season.

The incident involving Target, the No. 3 U.S. retailer, was one of the biggest retail cyberattacks in history.

In a confidential, three-page report to retail companies the FBI described the risks posed by "memory-parsing" malware that infects point-of-sale (POS) systems, which include cash registers and credit-card swiping machines in checkout aisles.

En-orig-R33

Although White (2000) claims that each text unit is typically fairly short consisting of 1-3 sentences, the data in the current study include rather long text units and they are not always deleted as a whole but sequences within them are deleted, as 6-9 shows.

At the rank of figure, similar reasoning for deletions appears to be at work, i.e. deletions occur if they do not prevent the reader from understanding the sequence, in which the figure to be deleted is located. Many examples of deletion at this rank include projecting clauses. An example is provided in 6-10 in which underlined figure is deleted.

6-10 The Senate is expected to vote next week on Stanley Fischer's nomination for vice chairman of the Federal Reserve, securing his spot ahead of the May recess.

En-orig-C14

Several examples show that the remaining figure is connected to its surrounding co-text after a figure (underlined in 6-11) is deleted from the sequence they are placed. It means the contribution of figure deletions is not necessarily limited to the rank of sequence as below.

6-11 The hospital's president told reporters that the dedicated nurse died trying to save her patients.

'Nurses are protectors by nature. And Gail, she fit that profile,' said ...

En-orig-C14

At the rank of element, where deletions occur most frequently, proper nouns, time expressions such as 'for 16 days', 'earlier this year' and modifying elements such as 'potentially fatal', 'bemused' or '(data center) that will house banks of PCs' are the major candidates to be deleted. These are the same types that are likely to be deleted in Ja-trans as well. Ja-trans includes more instances than in En-trans. A possible reason for the numerous instances in Ja-trans is the frequent use of proper nouns in En-orig, especially people's names. Police officers, spokespersons, eyewitnesses are presented with their names in the English ST but in Japanese it is not as common as in English.

Deletions at the ranks of element and figure occur throughout the text and the same applies to sequence deletions to some extent. However, deletions at the rank of text unit tend to take place towards the end of the text. This is likely to be relevant to the notion of inverted pyramid (Bell 1991, 1998); the structure to arrange the information in order of importance in news production. Adjusting the length of news article is said to be relatively easily done by dealing with the information towards the end.

6.1.2.3. Moves

Unlike additions and deletions, moves are not only about information content but also relevant to information flow. This section focuses on the moves that take place across the boundary of the text nucleus and satellites to examine what kind of information is considered suitable to be included in the nucleus and what not in English and in Japanese, thus focusing on an aspect of moves that is more about content than flow.

Firstly, however, a general trend of move occurrence, i.e. in no relation to the boundary of the text nucleus and satellites, is presented. The table below shows the number of texts that involve moves according to the semantic rank. These ranks are assigned to the units to be moved while they are in the ST (Chapter 4 Section 4.1).

Table 6-5: Number of texts involving moves according to the semantic rank

	Text unit	Sequence (T-unit)	Figure (Clause)	Element
En-trans-Yomiuri (85)	6% (5)	26% (22)	20% (17)	62% (53)
En-trans-Mainichi (69)	9% (6)	70% (48)	23% (14)	68% (47)
Ja-trans-CNN (78)	31% (24)	54% (38)	35% (27)	79% (62)
Ja-trans-Reuters (81)	11% (9)	26% (21)	31% (25)	69% (56)

Moves across the boundary of the nucleus and satellites occur at all ranks. Generally, the element is the rank at which moves occur most frequently, although moves at the ranks of sequence and figure are also relatively frequent. Table 6-5 shows that Ja-trans employs moves more than En-trans in general, while moves at the rank of sequence in En-trans-Mainichi are prominent.

Now, turning to the moves that occur across the boundary of the text nucleus and satellites, Table 6-6 provides the number of news articles that include these shifts.

Table 6-6: Number of texts involving moves according to the semantic rank across the boundary of the text nucleus and satellites

	Text unit	Sequence (T-unit)	Figure (Clause)	Element
En-trans-Yomiuri (85)	2% (2)	12% (10)	1% (1)	41% (35)
En-trans-Mainichi (69)	4% (3)	30% (21)	9% (6)	48% (33)
Ja-trans-CNN (78)	0% (0)	10% (8)	6% (5)	23% (18)
Ja-trans-Reuters (81)	1% (1)	7% (6)	14% (11)	48% (39)

Moves across the boundary of the nucleus occur at all ranks and again, the element is the rank that these moves are most frequent. Moves at the rank of sequence are also relatively frequent in comparison to those at the rank of the text unit or the figure, possibly because they are more independent and more flexible to be moved around.

Bringing in the direction of moves, Table 6-7 presents the number of texts involving moves from the text nucleus towards satellites, and 6-8 involving moves from satellites towards the text nucleus.

Table 6-7: Number of texts involving moves from the text nucleus towards satellites according to the semantic rank

	Text unit		Sequence		Figure		Element	
En-trans (154)	3% (4)		14% (21)		6% (9)		42% (65)	
Yomiuri/Mainichi	1	3	7	14	5	4	33	32
Ja-trans (159)	0% (0)		1% (1)		3% (5)		11% (17)	
CNN/Reuters	0	0	1	0	0	5	5	12

Table 6-8: Number of texts involving moves towards the text nucleus from satellites according to the semantic rank

	Text unit		Sequence		Figure		Element	
En-trans (154)	1% (1)		7% (11)		3% (5)		7% (11)	
Yomiuri/Mainichi	1	0	4	7	3	2	7	4
Ja-trans (159)	1% (1)		5% (8)		4% (6)		22% (35)	
CNN/Reuters	0	1	1	7	0	6	5	30

Moves from the text nucleus towards satellites occur more frequently in En-trans than the other direction; moves towards the text nucleus. On the other hand, in Ja-trans, moves towards the text nucleus is slightly more frequent, although in Ja-trans moves that occur across the boundary of the text nucleus and satellite are infrequent except for those at the rank of element, which are prominent.

The type of information that is moved at each rank can be related to the type of information that is added or deleted. The main function of additions and deletions is assisting the reader's understanding of translated news by controlling information either locally or globally. The difference between moves and additions/deletions is that moves involve both additions and deletions in the process of translation, i.e. information is deleted from one location in the ST which is then added in the TT at a location that is different from the ST.

Hence, moves also occur in order to assist the reader in understanding the translated news. What kind of assistance is provided in terms of move may be explained by addressing the clear contrast found between the moves in En-trans at the rank of element in Table 6-7 and that in Ja-trans in Table 6-8. The difference between Table 6-7 and Table 6-8 is the direction of the moves; Table 6-7 presents moves from the text nucleus to satellites while Table 6-8 towards the text nucleus from satellites.

The contrast in question suggests that the text nucleus in news articles written in Japanese and in English are different in terms of the information included in it. Given that information is moved from the text nucleus down to satellites in En-trans, the assumption is that the text nucleus of news in Ja-orig contains information that the same text unit in English tends not to. This can be supported by that Ja-trans involves moves towards the text nucleus. Analysis finds that specific information is frequently moved to the text unit in Ja-trans.

For example, 6-12 presents the text nucleus of a news article written in English.

6-12 A video showing a blind man being repeatedly kicked and stomped by an unknown assailant was released by the Philadelphia Police Department Tuesday. En-orig-C1

When this piece is translated into Japanese, the age of the 'blind man' is provided in the TT's text nucleus, which is located in one of the satellites in the ST. Sion (2007: 140) states that news articles in Japanese provide all the detailed information about the criminal suspects in reporting a crime, such as name, age and occupation in the text nucleus. Although the current example is not about the suspect but the victim, the concept is the same. This type of moves occur typically at the rank of element.

An equivalent type of information 'in her 50s' is removed from the text nucleus of an ST and placed in a satellite in the TT (En-trans-M26). These are major moved elements and it can be said that the nucleus in Japanese news include more specific information which news articles in English do not normally place in the nucleus.

Additionally, the translation of the same text nucleus in 6-12 also includes moved information from a satellite about the victim's condition after the reported incident, 'the victim is left with injuries to his head and face'. This is at the rank of sequence and therefore does not occur as frequently as those at the rank of element. However, it shows an aspect that the text nucleus in Japanese tends to provide a summary of the incidents being reported (Thomson *et al.* 2008).

To summarise, information moved at the rank of element can be said to assist the reader's understanding of news, not only by what information is required but also where it is required. This is likely to apply to moved information at other ranks. Moreover, it is relevant to the conventional structure of news. Different type of information is placed in the text nucleus in English and in Japanese.

6.1.3. Patterns of additions, deletions and moves

Section 6.1 has examined the relationship between the text structure and information content. The general question is what kind of adjustments take place in terms of information content for the target reader to understand translated news texts. Patterns for translation shift are summarised below.

Both En-trans and Ja-trans indicate a high proportion for employing each of the additions, deletions or moves as well as deploying two or three of them together.

En-trans employs additions more than they do deletions in general, which is opposite from the tendency of Ja-trans. Noticeable features of additions in En-trans include that they occur at all ranks, assisting the reader's understanding from the local level to global although the rank at which they occur most frequently is the element, which is local. Additions at the rank of text unit in En-trans are significantly more than that of deletion, and compared to Ja-trans. En-trans-Yomiuri shows a high number of incidents at this rank, as they add many text units towards the end of the text.

Added information assists the reader to understand the translated news with three major aspects; by filling the knowledge gap, assisting them with the discourse and creating the text in a conventional way. These are relevant at all ranks of additions, but depending on the rank, the provided assistance can be local (element) or global (text unit). While local assistance appears throughout the text, additions of text unit tend to be either at the beginning or towards the end of the text.

In Ja-trans additions are fewer than deletions as well as compared to En-trans, although they do occur at the rank of element. On the other hand, Ja-trans involves deletions at all ranks. The rank at which deletions occur most frequently is the element but those at the rank of sequence are also common in Ja-trans. Being more independent than the figure and smaller than the text unit could be a possible reason for that. It is also likely to be relevant to that En-orig news in the corpus involve many instances of projections, which are often deleted in translation. Deletions can occur throughout the text, making contribution locally (element) but also globally (text rank), the latter tends to occur at the end of the text.

Moving to En-trans, 90% of the texts involve deletions, which is slightly less compared to Ja-trans. Among the four semantic ranks the element is where most deletions occur. En-trans-Mainichi has more examples of deletion, while En-trans-Yomiuri has more of addition.

Moves in En-trans occur at all semantic ranks with many incidents observed at the element and sequence ranks. Moves are discussed focusing on those taking place beyond the boundary of the nucleus and satellite. En-trans show more examples of moves from the nucleus. This suggests that the text nucleus

in the Japanese news include the kind of information that the text nucleus in English news does not, such as details of the criminal suspects.

Moves in Ja-trans also occur at all semantic ranks. To mirror the tendency of En-trans, Ja-trans show more examples of moves towards the nucleus. This too suggests that the text nucleus in Japanese includes the kind of information that English news does not in their nucleus. Coming from the both directions, this is likely to be one of the features of Japanese news.

In summary, based on the data analysis I have found additions, deletions and moves are differently deployed in translation depending on the translation direction. This can be said to reflect the target audience of the translated news. While En-trans of Japanese news are likely to be targeted towards speakers of English in Japan (*Yomiuri Shimbun* 2016: 8), Ja-trans of English news are targeted towards speakers of Japanese likely to be in Japan. Relevance of the news to the target reader is very different between the two. Geographical, political, cultural and social distance can also be different. This section has provided evidence and added nuance to the general consensus of additions, deletions and moves during news translation based on the empirical analysis. It has also demonstrated that the target readership influences how the strategies of additions, deletions and moves are deployed.

6.2 Text structure and mode of discourse (thematic progression)

Following the discussion of translation shifts in terms of information content, the current section examines these in terms of information flow, placing the focus on how Theme is realised at specific locations in the text nucleus and satellites within the text. Information is chosen and set in a text and its progression through Theme at the beginning of discourse unit is observed.

I examine whether or not Theme can be said to function as an orienter of the message at a level more global than the clause. These questions concern the internal thematic structure of the text, and they are set in attempt to establish a path to understanding the relationship between the Theme and functional segments of text in translation. In addition to understanding and establishing patterns of thematic shifts in translation, however, there is an aim of gaining a better understanding of the discourse structure of Japanese news articles as well as how Theme relates to it.

6.2.1 Thematic progression in En-orig and Ja-orig

This section describes Theme realisation at the beginning of the text nucleus and satellites in the En-orig and in the Ja-orig in order to observe contrasts and similarities between them. The contrastive account of them is then used as a basis to analyse Theme realisation at the same locations in En-trans and Ja-trans (6.2.2). Statistics provided in 6.2.1 are obtained using an online statistical analysis programme js-STAR, version 9.0.6j (Tanaka & Nakano 2018).

6.2.1.1 Theme at the beginning of the text nucleus

In this section Theme realisation at the beginning of the text nucleus is analysed. Although the text nucleus typically consists of the headline and what is called ‘the lead’ in journalism (Chapter 3 Section 2.3), my analysis focuses on the lead sentence, and not the headline. The reasons for this are the interdependence of the headline and the lead sentence(s) (White 1998: 266-267) as well as the different functions they serve in a text. Regarding the former as it appears in English news, White offers the explanation below:

The headline and opening sentence (or sentences) can be seen as representing a single unit or phase because, in the overwhelming majority of cases, the headline exactly repeats a sub-set of the informational content of the lead, serving simply to sign-post key meanings which will be presented more fully in the following sentence.

(White 1998: 266-267)

He goes on to clarify that this interdependence is due to the practice in the news production process, in which it is not the writer of the article but a subeditor who gives an article a headline after it is written, summing up the lead (*ibid.* 267). Although this aspect is beyond the scope of this project, the interdependence of the headline and the lead as a phenomenon is visible in the data as can be seen in the following example. The headline is indicated by bold letters and points of interdependence are underlined and indexed with subscript numerals.

6-13 Headline and lead in English news

¹Pilot ²killed ³in California ⁴air show ⁵crash

(CNN) -- ¹A pilot ²died Sunday when his small plane ⁵crashed at ⁴an air show at Travis Air Force Base ³in California, an official said.

En-orig-C71

This example clearly shows that the interdependence between the headline and the lead sentence is created by repeating the same core information content. The same phenomenon is also observed in the Japanese corpus used in this project. The Japanese headlines show a tendency of presenting key meanings as a list of lexis, not written in the form of clause. An example is provided here.

6-14 Headline in Japanese news

第一三共、	[[発疹など起こす]]化粧品	返金
<i>Daiichi Sankyō</i>	[[<i>hosshin nado okosu</i>]] <i>keshōhin</i>	<i>henkin</i>
Daiichi Sankyō	[[rashes, etc. cause]] skincare product	refund

Ja-orig-Y27

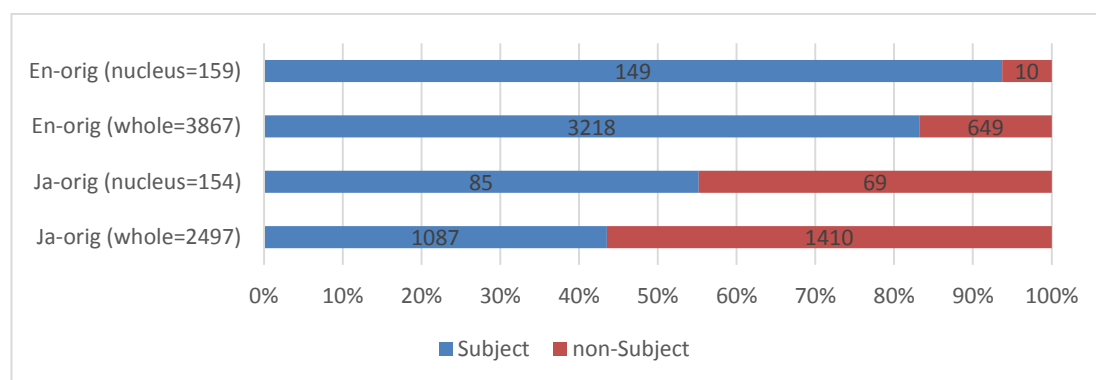
This headline consists of three noun groups with no grammatical relationship explicitly expressed between them although it is not difficult to recognise it. It is not difficult to identify the Theme either, which is the first noun, *Daiichi Sankyō*.

Despite the relative ease of identifying both the grammatical relation and Theme-Rheme structure, I favour analysing the lead sentence over the headline because it is written in the form of clause while the headline is not always so. Since Theme is a grammatical category, it helps to have data in the form of clause as that is where grammatical functions are realised and it allows me to observe data based on what is explicitly observable. Another reason to favour the lead sentence is related to one of the aims of my study, which is to explore how news articles develop as a text through the way in which Theme is realised at certain locations within the text.

Therefore, I look at the lead that is the first sentence within the discourse instead of the headline which is separated from the body of the text, and thus, strictly speaking, not part of the discourse. This point relates to the different functions they serve; the headline represents the key content of the text to inform the reader what the news is about, while the lead opens the news story to report it as a complete piece to fulfil its communicative purposes. For these reasons I will focus on the lead sentence to analyse the Theme realisation in the nucleus. In this chapter from this point on, in order to keep the terminology simple I will keep to the term ‘nucleus’ to address the lead sentence(s) of the text unless the distinction between the headline and the lead sentence(s) is necessary.

Concerning the beginning of the nucleus, the first discussion point is the choice between the Subject and non-Subject as Theme, beginning with the En-orig data. At the beginning of the nucleus the Subject is realised as Theme 94% of the time, which is higher compared to the whole En-orig corpus at 83%. This is confirmed as statistically significant ($X^2=11.530$, $df=1$, $p<.01$). In Ja-orig too, the Subject Theme appears significantly more at the beginning of the nucleus compared to the whole Ja-orig sub-corpus ($X^2=7.609$, $df=1$, $p<.01$).

Figure 6-1: Subject Theme or non-Subject Theme at the beginning of the text nucleus (En-orig & Ja-orig)



Even though the frequency differs, both En-orig and Ja-orig show a stronger tendency for the Subject to conflate with Theme at the beginning of the text nucleus than it generally does in the whole sub-corpus of En-orig and Ja-orig. This may be understood to suggest that the particular location at the beginning of the text nucleus is consequential in the choice of Subject Theme both in En-orig and Ja-orig.

However, I need to be cautious in proposing the Subject Theme to be a genre feature in news because it is not exclusively taking place at this location but simply a significant tendency. Given that the Subject Theme is an unmarked choice in English (Halliday & Mattiessen 2004) as well as in Japanese though the frequency differs, the Subject Theme can probably more safely be interpreted as a linguistic feature, which is intensified at the beginning of the text nucleus, and not as a defining feature of the text nucleus.

We have found that the same tendency of preferring the Subject Theme at the onset of the nucleus is perceived in both En-orig and Ja-orig. However, when we focus on the proportion of the Subject and non-Subject, there is a higher proportion for non-Subject Themes in Ja-orig (56%) than En-orig (17%). This inclination towards the non-Subject Theme in Ja-orig merits an explanation as this could be one of the areas where translation needs some adjustment in selecting the Theme at the beginning of the nucleus. But before moving on to

the discussion of what kind of non-Subject Themes are realised in Ja-orig, I will consider lexical items of the Subject Theme, especially the way they are marked by the different particles *ga*, the subject marker, or *wa*, the topic marker in Ja-orig. The choice of one over the other is sometimes restricted by grammar, sometimes due to genre convention and sometimes the translator's choice.

Regarding the types of Subject Theme in English, they can be either a clause or group but nearly all of them at the onset of the nucleus in my En-orig data are a nominal group (99%). Examples include 'a pilot', 'a woman', 'doctors', 'President Barack Obama', 'the US government' and 'Google Inc', etc. Depending on how easily the writer thinks they can be recovered in the mind of the reader, they are presented either as Given or New information (Halliday & Mattiessen 2004).

Similar distinction of Given and New can be seen in the Ja-orig data, which is indicated by the marking, either by *wa* or *ga*. This is not surprising as Subject Themes often conflates with the Topic but it can be a big issue in English into Japanese translation, especially if the translator is not a native speaker of Japanese. When grammar restricts a choice, e.g. the subject in a dependent clause are never marked by *wa*, it is not a problem to decide on one. If it is either a genre convention or a choice however, deciding on one is more difficult. This will be clarified below.

The Subject Theme was marked by *wa* in 47 instances and by *ga* in 39. Examples of the Themes marked by *wa* include '政府 *seifu* [the government]', '法務省 *hōmushō* [Ministry of Justice]' and '環境庁 *kankyōchō* [Environment Agency]' as well as committees within a ministry, prefectures, political parties or their leaders. Noda, who gives a detailed account on how the particles *wa* and *ga* are used from the point of view of Japanese as a foreign language, explains that the prime minister (of Japan), ministries or ministers are likely to be made the Topic of the first sentence in news reports as their activities are of interest to people, which makes them Given information (1996:156-157).

Other examples of the Subject Theme marked by *wa* include specific organisations such as '京都大病院 *Kyoto University Hospital*', '第一三共 *Daiichi Sankyo (Pharmacy)*', a political party leader '維新の会の石原氏 *Ishihara of Ishin-no-kaï*' or words such as '景気拡大 *keiki-kakudai* [economic expansion]', '中絶の多く *chūzetsu no ōku* [many cases of abortion]', or '大雪の影響 *ōyuki no eikyō* [the consequences of heavy snow]'. These are not inherently Given information and according to Noda (1996:157) unlikely to be made the Topic in the first sentence of news. However, they appear as the Topic with *wa* in

my data. A possible explanation is that they become Given information because they are introduced in the headline either directly or indirectly. In Example 6-15 below a direct lexical link is created by using the same word in the Theme of the nucleus and the headline, while the subsequent Example 6-16 shows an indirect but easily recognisable link between the hospital (Theme) and a transplant (headline).

6-15

Headline

第一三共、[[発疹など起こす]]化粧品 返金
<i>Daiichi Sankyō, [[hosshin-nado okosu]] keshōhin henkin</i>
Daiichi Sankyō, [[rashes ,etc. cause]] skincare product refund
Daiich Sankyō to refund skincare products that cause problems such as rashes

Theme

「第一三共ヘルスケア」(東京)は
<i>Daiichi Sankyō herusukea (Tokyo) WA</i>
<i>Daiichi Sankyō Healthcare (Tokyo) WA</i>
Daiichi Sankyō Healthcare
Theme: Subject (WA)

RHEME

5日、	通信販売限定の化粧品「ダーマエナジー」スキンケアシリーズについて、	//販売を中止し、返品・返金に応じると//	発表した。
5-ka	<i>tsūshinhanbai gentei NO keshōhin 'daama enajii' sukinkea shiriizu NI-tsuite</i>	<i>//hanbai O chūshishi, henpin henkin NI ōjiru TO//</i>	<i>happyōshita</i>
5 th	mail order only NO make-up 'derma energy' skincare series NI-tsuite	//sales O stop-SUSP, return, refund NI respond-PROJ//	announce-pst
announced that they will stop selling the mail order skincare series 'derma energy' and respond to return & refund			
Adjunct	Adjunct: Matter	Projected clause	Process

Ja-orig-Y27

6-16

Headline

妻の左肺に夫の右肺を移植...世界初の「左右逆」
<i>Tsuma NO hidarihai NI otto NO migihai O ishoku ... Sekaihatsu NO 'sayūgyaku'</i>
Wife's left lung NI husband NO right lung O transplant ... World 1st NO 'left-right reverse'
Husband's right lung transplanted to the left lung of wife... first left-right reverse in the world

Theme

京都大病院は
<i>Kyotodaibyoin WA</i>
Kyoto University Hospital WA
Theme: Subject / Topic

Rheme

14日、	//[[[[[肺の難病を発症した]]女性患者の左肺を切除し、代わりに夫の右肺の一部を移植する]]生体肺移植に成功したと//	発表した。
14 th	//[[[[[Hai NO nan'byō O hasshōshita]] josei kanja NO hidarihai O setsujoshi, kawarini otto NO migihai NO ichibu O ishokusuru]] seitai haiishoku NI seikōshita TO//	happyōshita
14 th	//[[[[[Lung NO incurable disease O develop-pst]] female patient NO left lung O remove-SUSP, instead husband NO right lung NO part O transplant]] live-donor lung transplant NI succeed-pst PROJ//	announce-pst
announced on the 14 th that they have been successful in the operation in which they removed the left lung of a female patient who had developed an incurable lung disease and transplanted to her the part of her husband's right lung		
Adjunct	Projected clause	Process

Ja-orig-Y72

These examples may give the impression that the beginning of the nucleus would normally start with a Theme that is also the Topic marked by *wa* if they are introduced in the headline. However, that is not always the case and there are 39 Subject Themes out of 85 at the onset of the nucleus which are not the Topic, thus, the particle *ga* is retained. An example of this kind is presented.

6-17

Headline

死亡時画像診断:NPOも 死因把握遺族の心のケアに
<i>Shibōji gazō shindan: NPO MO Shiinhaaku izoku NO kokoro NO kea NI</i>
Death-time image analysis: NPO too Cause of death identify bereaved family NO mind NO care NI

Theme

[[[[遺族の依頼を受けて遺体を一時的に預かる]]事業などを行っている]]] NPO法人「リシステム」(東京都千代田区)が
<i>[[[[Izoku NO irai O ukete itai O ichijitekini azukaru]] jigyō nado O okonatteiru]] NPO hōjin Risu Shisutemu (Tokyo-to Chiyoda-ku) GA</i>
[[[[Bereaved family NO request O accept-SUSP corpse O temporarily store]] service, etc. O do-ASP]] NPO organisation LiSS System (Tokyo, Chiyoda-ward) GA
An NPO Organisation LiSS System (Tokyo, Chiyoda-ward) which provides services such as temporarily storing corpses upon the order of the bereaved family
Theme: Subject (GA)

RHEME

今月から、	[[専用のCT(コンピューター断層撮影装置)で遺体を撮影する]] Ai(死亡時画像診断)を	始めた。
Kongetsu KARA	<i>[[sen'yō NO CT (compyūtaa dansō satsuei sōchi DE itai O satsueisuru]] Ai (shibōji gazō shindan) O</i>	<i>hajimeta</i>
This month KARA	[[designated CT DE corpus O photograph]] Ai O	start-pst
Started this month an Ai (service) to photograph corpses using a designated Computed Tomography		
Adjunct	Goal	Process

Ja-orig-M60

The Theme in the nucleus above could have been presented with *wa* without being unnatural as 'NPO (LiSS System)' has been mentioned in the headline albeit indirectly. This means the particle *ga* here is a result of the translator's choice. In other words, the translator has decided to describe an event's occurrence that is unknown to the reader using the 現象文 *genshōbun* (Noda 1996: 11-12), in which there is no Given information, instead of making use of the topic-comment structure, in which a Topic is established first and then a

Comment on it is supplied (Li & Thompson 1976, Noda 1996 among others). So the choice between *ga* and *wa* can be a choice of how the translator wishes to present an event. This explanation applies to the earlier examples in which *wa* has been chosen over *ga*. The choice between the *genshōbun* and the Topic-Comment structure is related to where information focus falls. In the Topic-Comment sentence, information focus falls on the comment part, while in the *genshōbun*, it falls on the entire sentence (Sunakawa 2005), while only the front receives the focus if the *genshōbun* sentence means that the nominal group marked by *ga* at the front has the meaning of ‘exclusively’.

Other than the information focus, the level of definiteness is also relevant. The difference between *ga* and *wa* in earlier examples is related to how easily the writer thinks a nominal group in question will be recovered by the reader. Between Kyoto University Hospital (Ja-orig-Y72) and an NPO organisation LiSS (Ja-orig-M60), a well-known university has a better chance to be established as a Topic that is inherently definite, while an individual or a group of people from the general public are less likely to, even though they are specified by embedded clauses to increase the definiteness as in the case of LiSS. This again goes back to the notion of Given and New.

Unlike the above examples, which suggest it is often a choice of the translator, there are cases in which the Subject Theme at the beginning of the nucleus needs to be marked by *ga*. It is when the nominalised noun ‘*koto* [the fact that]’ is the Subject. In this pattern all the newsworthy information is put together into the noun, creating a long Theme followed by a short Rheme, which is often ‘*wakatta* [became known]’ or a similar Process.

6-18

柏連続通り魔:容疑者宅から大麻押収
<i>Kashiwa renzoku tōrima: yōgisha-taku KARA taima ōshū</i>
Kashiwa assaults: suspect's home KARA cannabis seize

Theme

[[[千葉県柏市の連続通り魔事件で、強盗殺人容疑で逮捕された]]自称無職、竹井聖寿容疑者(24)の自宅から、大麻が押収されていた]]	ことが
[[[[Chiba-ken Kashiwa-shi NO renzoku tōrima jiken DE, gōtōsatsujin-yōgi DE taihosareta]] jishō mushoku, Takei Seiju yōgisha (24) NO jitaku KARA taima GA ōshū sareteita]]	koto GA
[[[[Chiba-prefecture Kashiwa-city NO series of assaults DE, robbery-murder suspect DE arrest-psv-pst] self-claiming unemployed Takei Seiju suspect (24) NO home KARA, cannabis GA seize-psv-ASP-pst]]	the fact that GA
The fact that cannabis has been seized from the home of Seiju Takei (24) who was arrested on suspicion of robbing and murdering relating to a series of assaults in Kashiwa-city, Chiba	
Modifier	Head
Theme: Subject (GA) Medium	

RHEME

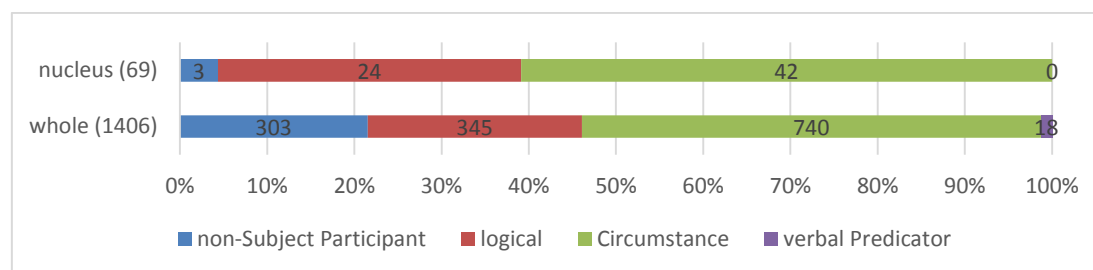
捜査関係者への取材で	分かった。
<i>Sōsakankeisha E NO shuzai DE</i>	<i>wakatta</i>
Investigation related people E NO interview DE	reveal-psv-pst
has been revealed in an interview to investigators	
Adjunct	Process

Ja-orig-M52

The Subject Theme in this example is loaded with information. This sentence pattern is called 同定文 *dōteibun*, identifying sentence (Noda 1996, Sunakawa 2005) in Japanese linguistics, and the Theme and Rheme can be inverted to reorder the Given and New information to mean ‘what has become known is the fact ...’. In the reordered sentence *wa* is used to mark ‘what has become known’ as that is Given in this context. Grammatically speaking, both orders are possible, but in news reports, especially at the beginning of the nucleus, it is always the order of New^Given, which is the conventional way of breaking news in Japanese. In terms of information focus it is the front part of the sentence that receives focus since the latter part of the sentence that is Given information is out of focus. This structure occurs 15 times out of 85 Subject Themes in Ja-orig, enough to be conspicuous, thus worth the translator’s being aware of as a possible translation device.

Having discussed the Subject Theme so far, I move on to consider non-Subject Themes at the onset of the text nucleus in Ja-orig. The focus is placed on Ja-orig as there are few instances (6%) of non-Subject Themes at the beginning of the text nucleus in En-orig; one 'there' from the existential sentence and six Circumstance Themes. Compared to this low occurrence rate in the En-orig data, Ja-orig include 69 instances (46%) of non-Subject Theme at the beginning of the text nucleus. Examples include Participant, logical, and Circumstance Themes. This is another area where translation shifts potentially occur. The figure below summarises the proportion of these three non-Subject Themes.

Figure 6-2: Non-Subject Theme type at the beginning of the text nucleus (Ja-orig)



The frequency of these Themes differs between the sub-corpora of Ja-orig nucleus and Ja-orig whole ($X^2=13.867$, $df=3$, $p<.01$). The statistics indicate that in terms of non-Subject Participant Theme and logical Theme, the two datasets are considered to be significantly different at $p<.01$, while not significant in terms of Circumstance Theme and verbal Predicator Theme. This suggests that Participant Themes are less likely to occur at the beginning of the text nucleus and logical Themes are likely to increase, which can be interpreted as relevant to the location at the beginning of the text nucleus. The choice of Circumstance Themes, however, is confirmed as not significant, i.e. unrelated to the location, thus remaining the general trend of the register feature. Examples for each category are below.

Non-Subject Participant Themes are rare at this location and none of them is topicalised. That is, the word order is changed to front the Participant in the Theme position but it is not made the Topic, thus not marked by *wa*. All three non-Subject Participant Themes are immediately followed by the Subject that is marked by *ga*. In 6-19 below the Subject and Actor, the University of Kyoto, which undertakes a study of iPS to treat cancer, appears immediately after the long Theme in the text nucleus, in the order of Scope-Actor-Process.

6-19

Headline

iPS細胞で免疫細胞、がん攻撃…治療法開発へ
<i>iPS saibō DE men'ekisaibō gan kōgeki ... chiryōhō kaihatsu E</i>
iPS cells DE immune cell, cancer attack ... treatment method development E
Immune cell by iPS cells, attack cancer...to develop treatment method

Theme

[[[がんを攻撃する]]免疫細胞をiPS細胞(人工多能性幹細胞)から大量に作り、がん患者に投与する]]]]	治療法の開発に、
[[[[Gan O kōgekisuru]] men'ekisaibō O iPS saibō KARA tairyō NI tsukuri, gankanja NI tōyosuru]]	<i>chiryōhō NO kaihatsu NI</i>
[[[[Cancer O attack]] immune cells O iPS cells KARA large amount create-SUSP, cancer patient NI give]]	treatment method NO development NI
Development of a new treatment method that creates large amount of immune cells form iPS cells that attack cancer to inject to cancer patients	
Modifier: embedded clause	Head
Theme: Scope	

RHEME in the nucleus

河本宏・京都大再生医科学研究所教授らのチームが	乗り出す。
<i>Kawamoto Hiroshi Kyoto-dai saisei ikagaku kenkyūjo kyōju-ra NO chiimu GA</i>	<i>noridasu</i>
Kawamoto Hiroshi Kyoto University Regenerative Medicine Institute professors team GA	endeavour
Actor	Process
the team led by Professor Hiroshi Kawamoto from Regenerative Medicine Institute, Kyoto University will start	

Ja-orig-Y44

Grammatically speaking, it is possible to topicalise the Theme above by adding *wa* to it. However, at the beginning of the text nucleus in a news article, in particular, in this article, it is unlikely to occur because the information included in the Theme is too complex for the writer to present it as Given information, even though the keyword ‘治療法 [method of treatment]’ is

introduced in the headline. In fact, it appears that this particular word order and un-topicalised Theme is the best amongst the four possibilities set out below to start this news article, excluding the one I have just rejected.

- 1) Scope – Actor (*ga*) – Process (current order and the best option)
- 2) Actor (*ga*) – Scope – Process (less effective than the above)
- 3) Actor (*wa*) – Scope – Process (not impossible but a little awkward)
- 4) Scope – Actor (*wa*) – Process (not impossible but awkward)

I will briefly explain why number 1) is the best. Number 4) is the least favoured amongst the four. If an Actor is the topic of a sentence, it will be natural, though not obligatory, to front it. With the current example it is difficult to see why not to front the Actor if the word order in number 4) were to be taken.

Number 3) is not a good option either because the Actor is not the university, which is well-known and can be presented as Given easily, but an individual who belongs to it. Unless the person in question is generally recognised by many and is easily recoverable as information, it is difficult to treat them as Given information, which is similar to the already rejected word order that is not included in the list above.

Number 2) is not a bad option as it is the normal word order, which leads us to think there is a reason for the writer to avoid it and take the one selected. I determine that the current word order is the best as it starts with the most newsworthy information, which is, however, not completely new since several related keywords are introduced in the headline. Therefore, instead of treating it as truly new information and place it towards the end of the sentence, which is the unmarked position for New information, the writer has fronted it as Theme, taking care not to topicalise it because of the low level of recoverability. It is understood to be the writer's conscientious choice based on discourse reasons, and not because of grammatical or generic restrictions.

Next, examples of non-Subject Theme are logical Themes (24 instances). Compared to the Ja-org whole data, the occurrence rate for Logical Themes at the beginning of the text nucleus is high.

6-20 Enhancement

Theme

デジタル教材の活用を	目指し、
<i>Dejitaru kyōzai NO katsuyō O</i>	<i>mezashi</i>
digital teaching material NO use O	aim-SUSP
Aiming to make use of digital teaching materials	
Scope	Process
Theme: logical	

RHEME

東京都荒川区は	6日、//4月から区立の全34小中学校にタブレット端末約9200台(先行配布分を含む)を配布すると//発表し、2014年度区予算案に事業費約8億円を計上した。
<i>Tokyo-to Arakawa-ku WA</i>	<i>6-ka, //shigatsu KARA kuritsu NO zen 34 shōchūgakkō NI taburetto tanmatsu yaku 9200-dai (senkō haifubun O fukumu) O haifusuru TO// happyōshi, 2014-nen'do kuyosan'an NI jigyōhi yaku 8-oku en O keijōshita</i>
Tokyo Arakawa-ward WA	6 th , //April KARA public NO all 34 elementary and middle school NI tablet terminal about 9200 (advance distribution O include) O distribute –PROJ// announce-SUSP, 2014 fiscal year budget NI project grant about 8 billion yen yen O allocate-pst
announced on 6 th that they would distribute about 9200 tablet terminals to all 34 public elementary and middle schools (including the ones already given away) from April, and allocated 8 billion yen to budget for the fiscal year of 2014	

6-21 Projection

Theme

小学生に	公立高校を	アピールしよう	と、
<i>Shōgakusei NI</i>	<i>kōritsu kōkō O</i>	<i>apīrushiyō</i>	<i>TO</i>
elementary school pupil NI	public high school O	Promote-vol	PROJ
To promote public high schools to elementary school pupils			
Participant		Process	
Theme: projected clause			

Rheme

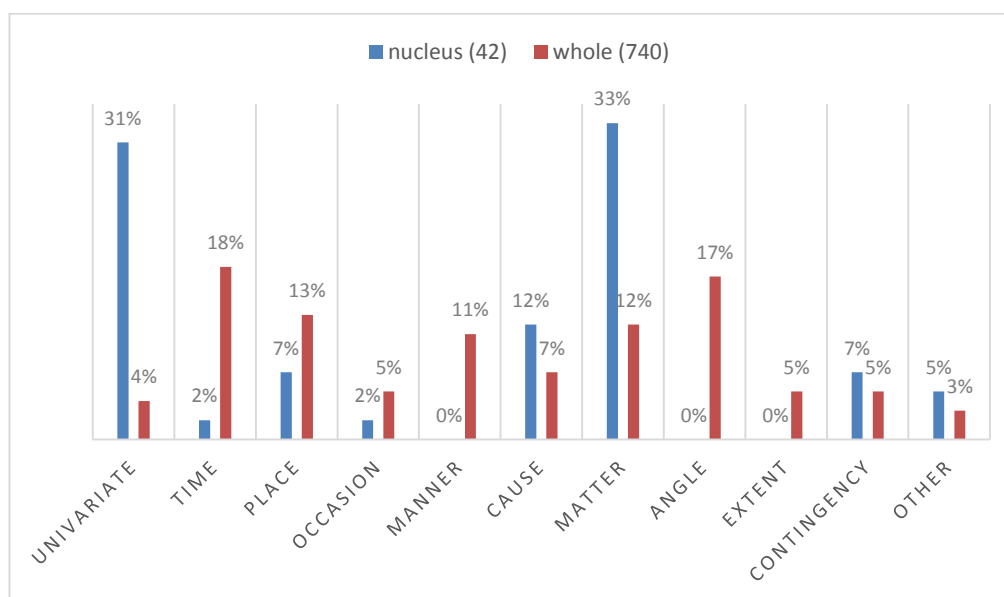
首都圏で、	[[小学生向けに高校の説明会を開く] 公立高や教育委員会が	目立ち始めた。
<i>Shutoken DE</i>	[[<i>shōgakusei muke NI kōkō NO setsumeikai O hiraku</i>]] <i>kōritsu-kō YA kyōikuiinkai GA</i>	<i>medachi hajimeta</i>
Metropolitan area DE	[[elementary school pupils towards NI high school NO briefing O hold]] public high schools and boards of education GA	show-start-pst
public high schools and boards of education that hold briefing meetings for elementary school pupils started to show		
Adjunct	Medium	Process

Ja-orig-Y8

When dependent clauses are realised as Theme at the beginning of the text nucleus, they are never topicalised in Ja-orig. It is not impossible to do so but there are restrictions and indeed the two examples here cannot be topicalised; the first one is purely formal in that the suspensive form cannot be topicalised and the second a combination of form and function. Projected clauses cannot be topicalised when they function like a Circumstance.

The final point of discussion with respect to non-Subject Themes is Circumstance Themes. Although frequency of choosing any Circumstance Theme at the onset of the nucleus does not differ significantly from that of the whole Ja-orig, there are particular Circumstances that are likely to be realised as Theme at this location. The figure below displays a distribution of different Circumstance Themes at the beginning of the text nucleus in comparison to the whole Ja-orig.

Figure 6-3: Circumstance Theme type at the beginning of the text nucleus (Ja-orig)



The frequency of Circumstance Themes differs between the sub-corpora of Ja-orig nucleus and Ja-orig whole ($X^2=93.994$, $df=10$, $p<.01$). The statistics indicate that the two sub-corpora are considered to be significantly different at $p<.01$ in terms of univariate Theme, temporal Location (Time), Matter and Angle as Theme. They are also different in terms of Manner but at a lower significance level of $p<.05$. The practical interpretation of this result is that univariate Circumstance and Matter are likely to occur at the beginning of the text nucleus while Angle is hardly ever realised as Theme at the location. Time and Manner are also unlikely to appear as Theme here. These can be interpreted as relevant to the location at the beginning of the text nucleus.

Let us observe some examples of these.

6-22 Univariate 1

Theme

13日午後8時40分ごろ、	大阪市都島区内代町1の市道で、
<i>13-nichi gogo 8:40 goro</i>	<i>Osaka-shi Miyakojima-ku Uchishiro-cho 1 NO shi-dō DE</i>
13 th pm 8:40 about	Osaka-city Miyakojima-ward Uchishiro-town 1 NO public road DE
About 8:40 p.m. on the 13 th , on the public road at Osaka-city Miyakojima-ward Uchishiro-town 1	
Adjunct	Adjunct
Theme: univariate	

Rheme

[[横断歩道を歩いていた]] 高齢男性が	大阪府警城東署のパトカーに	はねられ、	死亡した。
[[[Ōdanhodō O aruiteita]]] kōrei dansei GA	Osakafukei Jōtōsho NO patokaa NI	hanerare	shibōshita
[[zebra crossing O walk- ASP]] elderly man GA	Osaka Prefecture Jōtō Police NO patrol car NI	hit-psv-SUSP	die-pst
An elderly man who was crossing at a zebra crossing was hit by a patrol car belonging to the Osaka Prefecture Jōtō Police and died			
Medium	Agent	Process	Process

Ja-orig-M63

Majority of the univariate Themes position the temporal Location before spatial and they all report a relatively specific time as above. Few examples that put the spatial Location first followed by the temporal Location show a tendency of reporting the time of event less specifically as below.

6-23 Univariate 2

Theme

横浜市神奈川区のマンションの1室で	12日夜、
Yokohama-shi Kanagawa-ku NO mansion NO 1-shitsu DE	12-nichi yoru
Yokohama-city Kanagawa-ward NO apartment NO room DE	12 th evening
In an apartment in Yokohama-city Kanagawa-ward on 12 th evening	
Adjunct	Adjunct
Theme: univariate	

Rheme

[[[[住人とみられる]]女性が倒れている]]のが	見つかり、
[[[[Jūnin to mirareru]] josei GA taoreteiru]] no GA	mitsukari
[[[[resident TO see-psv]] josei GA ly-ASP]] NOM GA	find-psv-SUSP
Medium	Process
A female who is considered to be the resident was found lying (on the floor)	

Ja-orig-Y48

The final example of Circumstance Theme types is Matter.

6-24 Matter

Theme

[[北区赤羽北1で見つかった]]不発弾について、
[[<i>Kita-ku Akabane-kita 1 DE mitsukatta</i>]] <i>fuhatsu-dan NI-tsuite</i>
[[<i>Kita-ward Akabane-kita 1 DE find-psv-pst</i>]] unexploded bomb NI-tsuite
Theme: Adjunct
Concerning the unexploded bomb found in Kita-ku Akabane-kita

Rheme

陸上自衛隊は	17日、	[[信管と呼ばれる]]起爆装置を	取り除き、
<i>Rikujōjietai WA</i>	<i>17-nichi</i>	[[<i>shinkan TO yobareru</i>]] <i>kibakusōchi O</i>	<i>torinozoki</i>
Ground Self-Defence Forces WA	17 th	[(detonation) fuse TO call-psv] detonating device O	remove-SUSP
Ground Self-Defence Forces on 17 th removed the detonating device called (detonation) fuse			
Actor	Adjunct	Goal	Process

Ja-orig-M16

I have identified that certain types of Circumstances, such as univariate Circumstance and Matter, are more likely to be realised at the beginning of the text nucleus, while others such as Angle are unlikely to appear at the same location. When a more general level of Circumstance, the same level as logical and Participant Theme, is a variable, the frequency of Circumstance Theme is not statistically significant to separate Ja-orig nucleus from Ja-orig whole. However, when a more delicate level of Circumstance, i.e. types of Circumstance, is a variable, some of them are able to indicate the two datasets in question are different. (cf. Teich 2003)

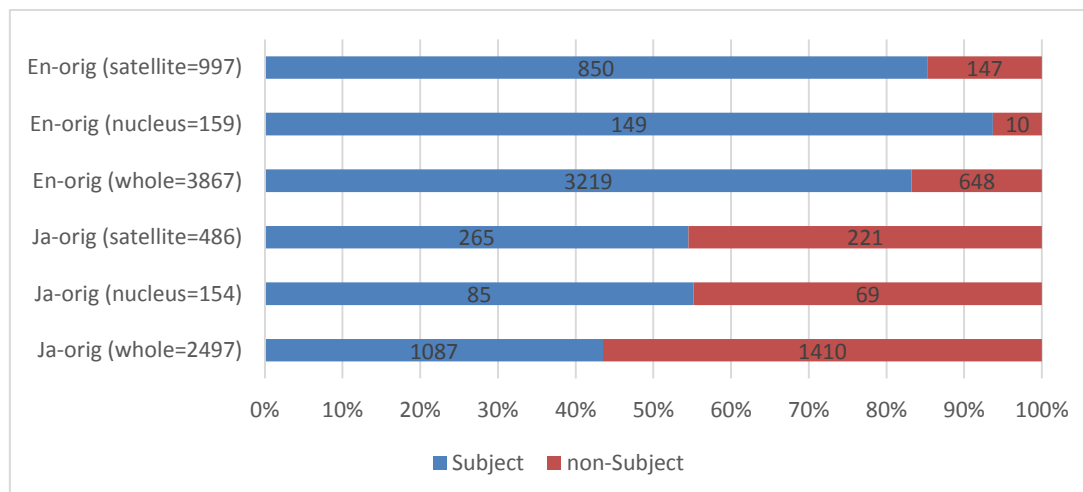
Before ending this section, I will briefly report that non-Subject Themes at the beginning of the text nucleus are rarely marked by the topic marker *wa* though it is possible to do so (Chapter 4 Section 3.3.3). As far as the Theme at the beginning of the text nucleus is concerned, *wa* is reserved to mark the Subject Theme exclusively and only one Circumstance Theme (spatial Location) marked by *wa* is found in the Ja-orig data (Ja-orig-Y40).

6.2.1.2 Theme at the beginning of satellite

This section concerns the beginning of satellites and explores Theme realisation at that location to find out if it differs from the beginning of the text nucleus. The first point of discussion is the choice between the Subject and non-Subject Theme, followed by the choice among non-Subject Themes. Under each point of discussion I will examine En-orig first then move on to Ja-orig.

Figure 6-4 below summarises the proportion of Subject and non-Subject Theme at the beginning of satellites in En-orig and Ja-orig, compared with the beginning of the text nucleus as well as throughout the sub-corpus in each language.

Figure 6-4: Subject Theme or non-Subject Theme at the beginning of satellite (En-orig & Ja-orig)



At the beginning of satellites in En-orig the Subject is chosen as Theme less frequently than it is at the beginning of the text nucleus. This difference is confirmed to be statistically significant ($X^2=7.647$, $df=1$, $p<.01$). Compared with the whole En-orig data, the frequency of the Subject Theme at the onset of satellites appears similar to the whole En-orig, which is statistically confirmed to be not significant ($X^2=2.204$, $df=1$, ns). These two comparisons show that the Theme choice at the onset of satellites is unlikely to be influenced by the specific location at the beginning of satellites and reflects the general register feature of preferring the Subject Theme. Hence, in terms of the choice between the Subject and non-Subject in En-orig, the location of the onset of the text nucleus is meaningful, while the onset of satellites is not.

In contrast to the En-orig sub-corpus, which shows a clear inclination for the Subject to be chosen as Theme in general, in the Ja-orig sub-corpus both Subject and non-Subject are chosen as Theme with a similar frequency.

However, as Figure 6-4 above shows, the Subject Theme appears more at the beginning of satellites compared to the whole Ja-orig sub-corpus and it is statistically significant ($X^2=72.355$, $df=1$, $p<.01$). When compared with the beginning of the text nucleus, in terms of the frequency of the Subject Theme, Ja-orig nucleus and Ja-orig satellites are found not different from each other ($X^2=0.003$, $df=1$, ns). Thus, in the Ja-orig sub-corpus the locations of the beginning of the text nucleus as well as that of satellites are not meaningful, whereby they do not play a role in the choice of Subject or non-Subject Theme.

To sum up the examination of Theme choice at the beginning of satellite, it can be said that the Subject is chosen more frequently at this location both in En-orig and Ja-orig, although the degree of preference of the Subject is stronger in En-orig than in Ja-orig. Before moving on to discussing the types of non-Subject Theme, I will note several findings as points of interest, which are not directly relevant to the research questions but insightful nonetheless. The first issue is the multiple Theme, especially textual, followed by the likelihood of the Theme being marked by *wa* depending on the type of Theme, and the location they appear in the text.

While the multiple Themes realised at the beginning of the text nucleus are limited to univariate multiple Themes, 14 multivariate multiple Themes are found at the beginning of satellites. There is one interpersonal Theme '*gutaitekini WA* [specifically]' but others are textual Themes such as '*mata* [also]', or '*ippō* [on the other hand]'. Given that the total number of multiple Theme instances is 167 in the Ja-orig sub-corpus, the probability for them to appear at the beginning of satellites (8%) is not overly high but indicative.

Textual Themes such as '*shikashi* [but]', '*sarani* [in addition to]' or '*tada* [however]' commonly relate the clause they are realised in to the immediately preceding clause/unit, and so they act as connectors to put the neighbouring clauses/units together rather than to separate them. '*Mata* [also]' or '*ippō* [on the other hand]' can be used to bring in additional or alternative information, which is a reason why they are realised at the beginning of the satellites and never at the beginning of the text nucleus.

Now turning to the likelihood of the Theme being marked by *wa*, an interesting point about Subject Themes that appear at the beginning of satellites is that the overwhelming majority of them, more than 89% (235 out of 265), are marked by *wa* compared to 55% of the Subject Theme at the beginning of the text nucleus. Non-Subject Themes are also marked by *wa* although less frequently than the Subject Themes. Still, given that the non-Subject Themes

that are realised at the beginning of the text nucleus hardly ever appear with *wa*, this can be a thematic feature helpful for translators.

The example below illustrates different types of Theme that are realised at the beginning of satellites and how *wa* is attached to many of them. Only the Themes in the satellites are presented here and points of reference are underlined and indexed with subscript numerals.

6-25

<p>₁浦和・差別行為「繰り返せば資格剥奪も」₂村井氏</p> <p>₁<u>Urawa</u> sabetsu kōi 'kurikaeseba shikaku hakudatsu MO' ₂<u>Murai-shi</u></p> <p>₁<u>Urawa</u> racial discrimination 'if repeated, barred from games' ₂<u>Mr Murai</u></p> <p>[[[[今月8日にJ1リーグ・浦和—鳥栖戦が行われた]]埼玉スタジアムで、人種差別とも受け取れる垂れ幕が掲げられた]]問題について、Jリーグは13日、₁浦和に対し、[[₃無観客試合の開催という]]リーグ史上最も厳しい処分を下した。</p> <p>[[[[Kongetsu 8-ka NI J1-līgu Urawa-Tosu sen GA okonawareta]] Saitama sutajiamu DE, jinshu sabetsu TO MO uketoreru taremaku GA agerareta]] mondai NI-tsuite, J-līgu WA 13-nichi, ₁<u>Urawa</u> NI-taishi, [[₃mukankyakyaku jiai NO kaisai To-iu]] līgu-shijō mottomo kibishii shobun O kudasita.</p> <p>Regarding the incident in which a banner that could be perceived as racist is displayed at Saitama Stadium on March 8 where a match between Urawa and Tosu was held, J. League on 13th ordered the ₁<u>Urawa</u> to play a home ₃<u>game without spectators</u>, the heaviest punishment it had imposed in the history of J. League.</p>	
--	--

Satellite 1	<p>[[[[1月末に就任した]]₂村井満チェアマンが打ち出す]]フェアプレー精神の徹底を</p> <p>[[[[1-gatsu matsu NI shūninshita]] <u>Murai Mitsuru cheaman</u> GA uchidasu]] feapurē seishin NO tettei O</p> <p>Endorsement of the notion of fair play that ₂<u>Chairman Mitsuru Murai</u> who took the post in January O (Non-Subject Participant)</p>
Satellite 2	<p>₂村井チェアマンは、</p> <p><u>Murai cheaman</u> WA</p> <p><u>Chairman Murai</u> WA (Subject)</p>

Satellite 3	[[9種類の処分の中から ₃ 無観客試合を選択した]]理由について、 [[9-shurui NO shobun NO naka KARA <u>mukankyaku jiai O sentakushita</u>]] riyū NI-tsuite about the reason why (he) selected <u>the game with no spectators</u> from the 9 types of punishment (Matter)
Satellite 4	₂ 村井チェアマンは、 <u>Murai cheaman</u> WA <u>Chairman Murai</u> WA (Subject)
Satellite 5	[[₂ チェアマンが特に問題視した]]のは、 [[<u>Cheaman GA tokuni mondaishi-shita</u>]] no WA What <u>the Chairman</u> took especially seriously (Subject)
Satellite 6	₁ 浦和は <u>Urawa</u> WA <u>Urawa</u> WA (Subject)
Satellite 7	₁ 浦和サポーターによる相次ぐトラブルに対しては、 <u>Urawa sapōtā NI-yoru aitsugu toraburu NI-taishite</u> WA with regard to recurring troubles caused by <u>Urawa</u> supporters (Matter)

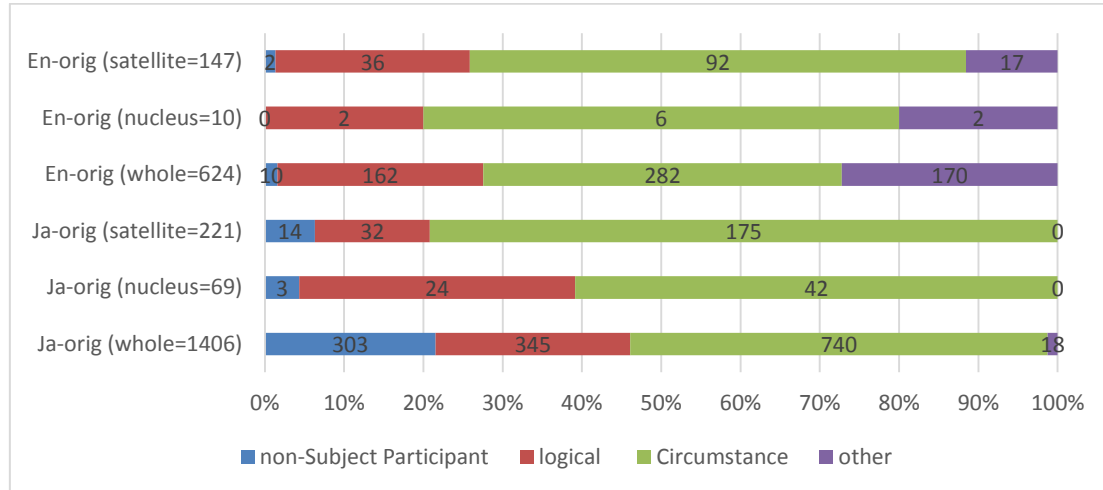
Ja-orig-Y49

It is clear that Themes marked by *wa* relate back to the elements introduced in the text nucleus. A satellite tends to start with Given information, creating a link to the text nucleus from where each Given information comes, and provides more information on the Given element by elaborating it, contextualising it, etc. However, not all the components in a Theme marked by *wa* are present in the text nucleus (e.g. satellites 5 & 7 in Example 6-15). Also not all the Themes realised at the beginning of satellites are marked by *wa* (satellites 1 & 3 in 6-15). At any rate, *wa* marking to some extent helps distinguish satellites from the lead in the text nucleus, especially when used with non-Subject Themes, as explained above.

I now move on to examining the choice amongst non-Subject Theme types at the onset of satellites. Using the figure below, I will observe if non-Subject Theme choice at the onset of satellites is different from that of the text nucleus and the whole sub-corpus in both En-orig and Ja-orig. The category 'other' in En-orig in the figure includes enhanced Themes (93 instances) and verbal

Predicator Theme (77 instances) (Chapter 4 Section 3), while the same category used for Ja-orig contains several unclear items.

Figure 6-5: Non-Subject Theme type at the beginning of satellite (En-orig & Ja-orig)



In the figure above the proportion of non-Subject Theme types looks rather dissimilar between En-orig (satellite), En-orig (nucleus) and En-orig (whole). However, Chi-square test conducted with a pair of En-orig (satellite) and En-orig (nucleus) shows that they are not significantly different from each other ($X^2=0.775$, $df=3$, ns), which may have been due to variable numbers being small in En-orig (nucleus).

On the other hand En-orig (satellite) and En-orig (whole) are found significantly different ($X^2=19.623$, $df=3$, $p<.01$). Statistics indicate that the two datasets are significantly different in terms of logical Theme and other Themes ($p<.01$). It is not difficult to understand why enhanced Themes and verbal Predicator Themes are not realised at the beginning of satellites or the text nucleus. Starting a new text unit, whether it is a text nucleus or a satellite, with a verbal Predicator Theme is almost impossible. Example 6-26 below shows how a verbal Predicator Theme (underlined) is realised.

6-26 “We’re hoping to eradicate this ...” said Vicki Kramer,... En-orig-R-1

It is not easy to start a new text unit with an enhanced Theme either as the main function of these sentence structures is to emphasise an element within a given clause, thus it makes more sense to use them in the middle of the discourse, not at the beginning. The existential ‘there’ is one exception.

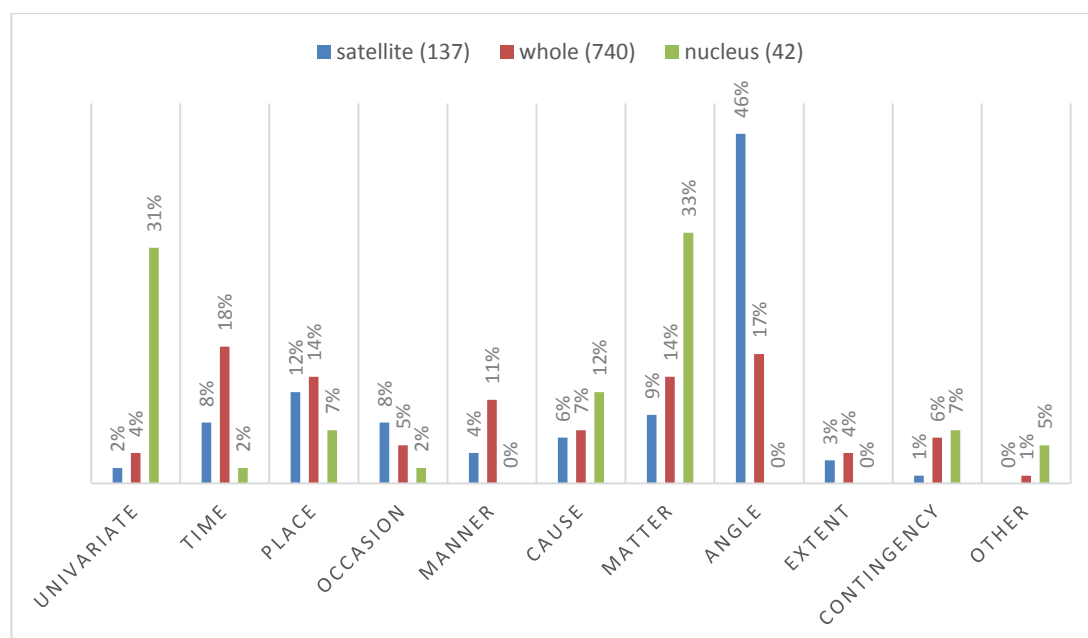
Thus, with respect to the possible genre influence on the non-Subject Theme choice at the onset of satellites, the temporary interpretation of the statistical

tests is that it is probably influenced by the location of Theme where it is realised, which is a genre feature.

Turning now to Ja-orig, but keeping to the beginning of satellites, Ja-orig (satellite) is compared with Ja-orig (nucleus) and found not significantly different from it ($X^2=\text{NaN}.000$, $df=3$, ns). However, a comparison with Ja-orig (whole) confirms that the two datasets differ significantly ($X^2=56.870$, $df=3$, $p<.01$). Statistics show that they are significantly different in terms of Participant, logical and Circumstance Themes ($p<.01$).

Amongst the three different non-Subject Theme types I will focus on Circumstance Themes to observe if they show signs to indicate satellite as opposed to the nucleus.

Figure 6-6: Circumstance Theme type at the beginning of satellite (Ja-orig)



Circumstance Theme of Accompaniment and Role are put together as 'other' since neither of them are realised as Theme at the beginning of satellites in the sub-corpus of Ja-orig.

The frequency of the Circumstance Themes significantly differ between Ja-orig (satellite) and Ja-orig (nucleus) ($X^2=89.991$, $df=10$, $p<.01$). Statistics show univariate Theme, Matter and Angle are accountable for the difference. Similarly, Ja-orig (satellite) and Ja-orig (whole) are found to be significantly different ($X^2=93.994$, $df=10$, $p<.01$). The statistics indicate that the two datasets are considered to be significantly different at $p<.01$ in terms of univariate Theme, Time, Matter and Angle as Theme. These results mean that Time and Angle are highly likely to occur at the beginning of satellites while

univariate Theme and Matter are unlikely. These can be interpreted as relevant to the location in the text where Theme is realised.

Univariate multiple Themes at the beginning of satellites are fewer compared to the beginning of the nucleus. This is one of the thematic features that can be employed to distinguish satellites from the nucleus. Satellites make more use of simple Locations rather than univariate multiple Locations.

Angle appears frequently at the start of satellites and I can say confidently it can be a thematic feature to distinguish satellites from the nucleus.

6.2.1.2 Comparison of En-orig and Ja-orig

Analysis of En-orig and Ja-orig has revealed several interesting results in terms of the way Theme is realised at the beginning of different discourse segments, at the very beginning of the news article, i.e. the text nucleus, and at the beginning of satellites, which function to specify elements introduced in the text nucleus (Chapter 3 Section 2.3). As well as presenting main similarities and differences between the two datasets to be used as a basis to analyse Theme in translation, this section provides a contrastive description of news articles written in Japanese to answer research questions 9-10, which are set to discover linguistic features of them. Focusing on the two locations in text: at the beginning of text nucleus and at the beginning of satellites, I will summarise the discussions above with respect to Theme realisation in En-orig and in Ja-orig.

At the start of the text nucleus, the Subject Theme is the majority in En-orig, accounting for 94%, which is significantly high compared to 83% in En-orig (whole)($p < .01$). In Ja-orig in contrast, the Subject Theme accounts for 56% at the beginning of the text nucleus, higher than the whole Ja-orig (43%). There is a clear difference between En-orig and Ja-orig in terms of how frequently the Subject Theme is employed at the onset of the text nucleus. However, they are similar in that the Subject is more frequently employed at the onset of the text nucleus, compared to the whole sub-corpora in En-orig and Ja-orig.

Focusing on how Subject Themes are realised in Ja-orig, in relation to the particle used, they are marked both by the Topic marker *wa*, and the Subject marker *ga*, as expected. *Wa* is employed 47 times while *ga* 38 at the beginning of the text nucleus, where *wa* marking is almost exclusively for the Subject, and non-Subject Themes are hardly ever marked by *wa*. In addition to the choice of particles, another interesting type of the Subject Theme appears in 同定文 *dōteibun*, identifying sentence, which manipulates the word order to achieve an order of New^Given to mean 'what has become known is the fact

...'. Its use is not abundant (15 times out of 85 articles), nevertheless, it is worth recognising this structure as a conventional way of breaking news in Japanese, which potentially causes a shift in translation.

Moving on to non-Subject Themes, in particular relation to Ja-orig since they are not common in En-orig, I have found Circumstance Themes are the majority at this location while logical Themes are also often employed. Among the types of Circumstance Theme, univariate multiple Theme (31%) and Matter (33%) are frequently realised at the beginning of the text nucleus. Because of its frequency in Ja-orig as well as its absence in En-orig, the former is likely to be a feature of news written in Japanese, which may be a cause for shifts.

Now I turn to the beginning of satellites. In En-orig, the Subject Theme is the majority (85%) at this location. However, compared to the onset of the text nucleus (94%), it is less frequent and its realisation is closer to the whole En-orig. Thus, the location at the beginning of satellites is not relevant to how the Subject Theme is realised in that position.

With respect to Ja-orig, the Subject is more frequently realised as Theme (67%) at the beginning of satellites compared to the onset of the text nucleus (56%). These Subject Themes are regularly marked by *wa* (89%), while at the beginning of the text nucleus, Subject Themes marked by *wa* account for just 55%, and in that location *wa* exclusively marks the Subject. At the beginning of satellites, however, *wa* marks other Themes such as a Circumstance, though it is less frequent.

Particular type of Circumstances are realised in specific locations in text. Firstly, univariate multiple Themes appear overwhelmingly frequently at the beginning of the text nucleus but very rarely at the beginning of satellites. Secondly, Matter is also often realised at the beginning of the text nucleus. They can be used to signal the start of the nucleus but unlikely to do so in the case of satellites. Thirdly, Angle is realised significantly often at the beginning of satellites but very rarely at the beginning of the text nucleus. As such, it can be used as an indicator to signal the beginning of a new satellite within the whole text.

6.2.2 Thematic progression in translation

Having explored Theme realisation in the non-translation sub-corpora of En-orig and Ja-orig, the current section examines the same topic in translation, again focusing on the beginning of the text nucleus and the beginning of satellites. Regarding these locations Theme realisations in translation are

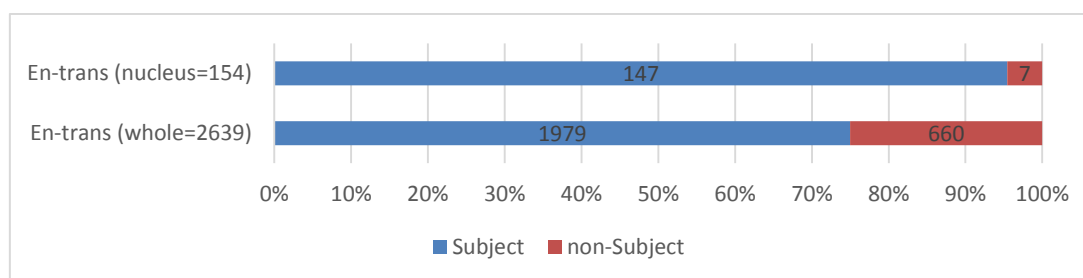
compared with different datasets to observe how Theme is realised at particular locations in the text and whether shift of Theme occurs during translation.

6.2.2.1 Theme at the beginning of text nucleus

In this section Theme realisation in translation at the beginning of the text nucleus is examined. Starting with English, I will address the question of the choice between the Subject and non-Subject as Theme in En-trans, followed by Ja-trans. Then I move on to discuss the options of non-subject Themes in Ja-trans.

Figure 6-7 below illustrates the proportion of the Subject Theme and non-Subject Theme in En-trans at the beginning of the text nucleus in comparison to the whole sub-corpus of En-trans.

Figure 6-7: Subject Theme or non-Subject Theme at the beginning of the text nucleus (En-trans)

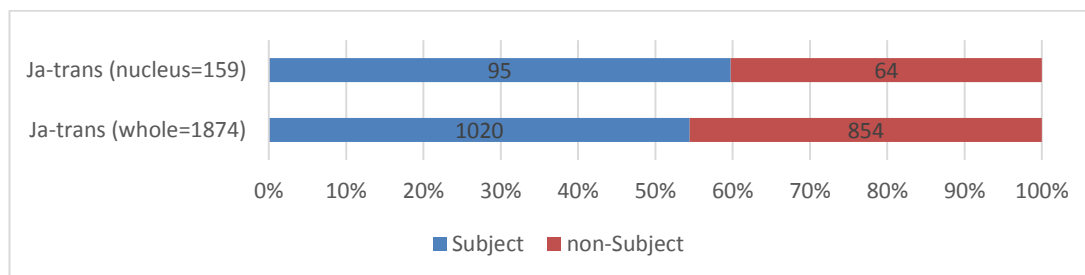


In En-trans at the beginning of the text nucleus the Subject is chosen as Theme more frequently than it generally is throughout the sub-corpus of En-trans. The percentage for the Subject Theme at the onset of the text nucleus is 94%, more than in En-trans (whole) at 75%. Regarding the choice of Subject or non-Subject, the two datasets of En-trans (nucleus) and En-trans (whole) are confirmed to be significantly different ($X^2=32.405$, $df=1$, $p<.01$).

This suggests that the location in the text at the beginning of the text nucleus is associated with what is chosen as Theme in En-trans. Selecting the Subject as Theme is nothing special as it is an unmarked choice in English in general (Halliday & Mattiessen 2004: 73). Nevertheless, my data indicate a stronger preference of Subject Theme at the beginning of the text nucleus in En-trans than the general tendency observed in En-trans (whole), that is, in no relation to a specific location in the text.

Moving on to Ja-trans, the proportion of the Subject and non-Subject Themes at the beginning of the nucleus is shown in Figure 6-8 below and it looks rather different from Figure 6-7 for En-trans above.

Figure 6-8: Subject Theme or non-Subject Theme at the beginning of the text nucleus (Ja-trans)



In Ja-trans at the beginning of the text nucleus, Subject Themes are realised slightly more often than non-Subject Themes but both are chosen as Theme at more or less the same frequency. The ratio of the Subject and non-Subject Themes is almost one to one not only at the beginning of the text nucleus but also throughout the sub-corpus of Ja-trans. This means that Subject or non-Subject Themes are realised in the same way regardless of the location in the text where they are realised. Hence, the two datasets, Ja-trans (nucleus) and Ja-trans (whole), are confirmed not statistically different ($X^2=1.467$, $df=1$, ns). In short, in Ja-trans, the choice of Subject or non-Subject as Theme at the beginning of the text nucleus is unlikely to be relevant to the location in the text where Theme is realised.

Hence, at the beginning of the nucleus, the Subject is employed as Theme more often than non-Subject in both En-trans and Ja-trans. However, the inclination towards the Subject Theme in En-trans is stronger than it is in Ja-trans. In addition, in terms of the frequency of the Subject Theme, En-trans (nucleus) is significantly different from En-trans (whole), while Ja-trans (nucleus) is not different from Ja-trans (whole). Thus, the location at the beginning of the text nucleus is significant for the Theme choice in En-trans but not in Ja-trans.

Before moving on to the question of what kinds of non-Subject Themes are realised at the beginning of the text nucleus, I will present and discuss two issues of interest that are relevant to the Subject Theme at the onset of the text nucleus in Ja-trans; long Themes and the use of the particle *wa*.

Long Subject Themes have been discussed earlier (6.2.1.1) as having an abstract noun of '*koto* [the fact that]' as the head noun with a long embedded clause to modify it, which is typically followed by a short Rheme. In Ja-orig, they are realised exclusively at the beginning of the nucleus and never at the onset of the satellite. In Ja-trans too these long Themes are realised at the beginning of the nucleus only. An example of these is presented below in combination with the original ST text nucleus in full, which illustrates that the

long Subject Theme in Ja-trans includes almost all the information presented in the original text nucleus.

6-27 Long Subject Theme in Ja-trans

Theme

[[オバマ米政権の上層部が、	[[国際テロ組織アルカイダの関係者で米国への攻撃を計画中とされる]]米国人の殺害作戦を	検討している]]	ことが
<i>[[Obama bei seiken NO jōsōbu GA</i>	<i>[[kokusai tero soshiki arukaida NO kankeisha de beikoku E NO kōgeki O keikakuchū TO sareru]] beikokujin NO satsugaisakusen O</i>	<i>kentō shiteiru]]</i>	<i>koto GA</i>
Obama US administration NO upper section GA	[[international terrorist organisation al Qaeda NO related person and US E NO attack O planning PROJ consider-psv]] American NO killing operation O	consider-ASP	the fact GA
The fact that the senior members of Obama administration are considering an operation to kill an American citizen who is suspected to be related to the international terrorist organisation al Qaeda and to have been scheming an attack against the US			
Actor	Phenomenon	Process	
Modifier: embedded clause			Head
Theme: Value			

Rheme

11日までに	分かった。
<i>11-nichi made ni</i>	<i>wakatta</i>
11 th by	(become-pst) know
Became known by 11 th	
Adjunct	Process

Ja-trans-C37

The text nucleus of the original ST for the above translation is presented here for comparison. Theme is underlined.

6-28 The Obama administration is in high-level discussions about staging an operation to kill an American citizen involved with al Qaeda and

suspected of plotting attacks against the United States, a senior U.S. official tells CNN. En-orig-C37

Long Subject Themes have a structure of '*koto* [the fact that] *ga wakatta* [became known]/*akirakani natta* [revealed]'. I have noted that it is the choice of the translator to employ this structure and it can be a useful device to start a news article although the number of occurrence is not abundant (15 times out of 85 Subject Themes). In Ja-trans it occurs eight times out of 154, considerably fewer than Ja-orig. This is possibly because the structure is complex and there are simpler, more unmarked structures of Subject+Predicate which can be used. Hence, it is an option which the translator could employ if desired.

The next issue to mention here is the use of the particle *wa*. I have established *wa* as marking something that helps distinguish satellites from the nucleus to some extent (6.2.1.3). We have seen that in Ja-orig the Subject Theme at the beginning of the satellite is marked by *wa* in 82% of cases, which is a lot more often than at the beginning of the nucleus at 58%. At the beginning of the text nucleus in Ja-trans the percentage of the Subject Theme marked by *wa* is 64%.

Examples found in Ja-trans include 'the White House', 'the US government', 'Toyota', 'NASA' amongst others. They are presented as Given information with *wa*; the translator must have judged them as recoverable from the reader's real world knowledge, which is reasonable. There were several instances whereby identification was not easily recoverable, such as 'a treasury official' or 'The *Council on Tall Buildings and Urban Habitat*'. However, it is the speaker, or the translator in this case, who decides whether to present something as Given or New information (Halliday & Matthiessen 2004). In any case, these slightly doubtful presentations of Given information do not interfere with the discourse flow of the translated news articles which read as well as those that employ undoubtedly Given information.

Except for the Subject Theme, non-Subject Themes are seldom marked by *wa*, and a rare example of them is presented here.

6-29 non-Subject Theme marked by *wa*

Theme

地滑りの現場では
<i>Jisuberi NO genba DE WA</i>
Landslide's site DE WA
At the landslide's site
Adjunct

Rheme

25日、	10人の遺体が	新たに	発見され、
<i>25-nichi</i>	<i>10-nin NO itai GA</i>	<i>aratani</i>	<i>hakkensare</i>
25 th	10 people NO body GA	additionally	discover-psv
10 bodies are additionally discovered on 25 th			
Adjunct	Medium	Adjunct	Process

Ja-trans-C56

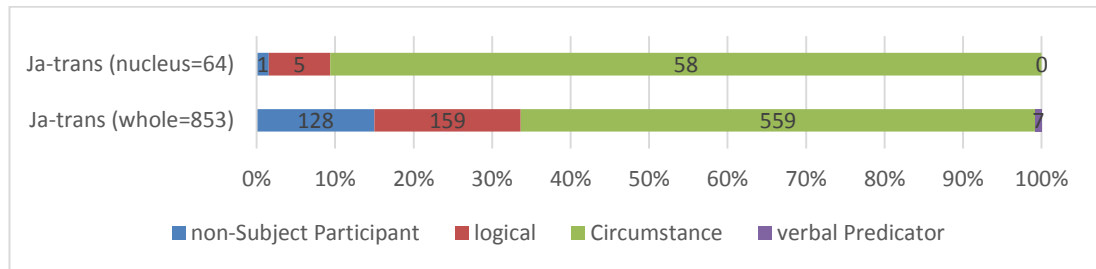
In Example 6-29, the particle *wa* is used to present the landslide site as Given information. This indicates that this is not the first time the landslide is reported and the previous article has been published not long ago. The translator probably expects the reader to remember the event, thus, *wa* is used. The sentence is perfectly appropriate without *wa* as well. Hence, the choice was made to employ *wa* here but this was not essential.

Having discussed the Subject Theme at the beginning of the text nucleus in both En-trans and Ja-trans, I will now examine non-Subject Themes that appear at the beginning of the text nucleus in Ja-trans. Non-Subject Themes occur frequently in Japanese, so examining them here is useful for upcoming comparisons of Ja-trans and other datasets later in the chapter.

I will first look at what kinds of non-Subject Themes are realised at the beginning of the text nucleus in comparison to the whole of the Ja-trans data. Unlike Ja-trans, very few non-Subject Themes are realised at this location in En-trans, so an explicit examination of the same question is not conducted here.

Figure 6-9 below summarises the number of instances of non-Subject Theme in Ja-trans. The 'other' category includes the absolute Theme and five unclassified Themes.

Figure 6-9: Non-Subject Theme type at the beginning of the text nucleus (Ja-trans)



What is clearly seen in Figure 6-9 above is that at the onset of the text nucleus in Ja-trans, Circumstance is chosen as Theme most of the time (90%). Also in Ja-trans (whole), Circumstance is the most frequently realised Theme but the percentage is 65% and lower than Ja-trans (nucleus). Also clearly visible is the difference in the proportion of the non-Subject Participant Theme between the two datasets. It appears just once (1%) in Ja-trans (nucleus) while throughout Ja-trans (whole) it is 15%.

According to the Chi-square test Ja-trans (nucleus) and Ja-trans (whole) are significantly different ($X^2=17.646$, $df=3$, $p<.01$) in terms of how non-Subject Themes are realised. Statistics show that the frequency of logical Theme is accountable for the two datasets to be different but its significance level is lower ($p<.05$). These results suggest that non-Subject Theme realisation at the beginning of text nucleus is different from the general tendency.

Now to a very rare case of non-Subject Participant Theme at the onset of the text nucleus in Ja-trans, as well as in Ja-orig. The only example from Ja-trans (nucleus) follows.

6-30

Theme

[[アルツハイマー病は米国で[[心疾患とがんに次ぐ]]第3位の死因に浮上しているとする]]	実態調査の結果を
[[Arutsuhaimaabyō WA beikoku DE [[shinshikkan TO gan NI-tsugu]] dai-3-i NO shiin NI fujyōshiteiru TO suru]]	jittaichōsa NO kekka O
[[Alzheimer's disease WA USA DE [[heart disease TO cancer NI follow]] the 3 rd cause of death NI come up PROJ take]]	survey NO results O
Results of survey that show Alzheimer's disease is the 3 rd common cause of death in the US following heart disease and cancer	
Modifier: embedded clause	Head
Theme: Complement	

Rheme

米国の研究チームが	まとめ、	5日の神経学会誌に	発表した。
<i>Beikoku NO kenkyū chiimu GA</i>	<i>matome</i>	<i>5-ka NO shinkeigakkaishi NI</i>	<i>happyō shita</i>
USA's research team	collate-SUSP	5 th NO neurology journal NI	publish-pst
A US research team has collated and published (it) in the neurology journal (that is published) on the 5 th .			
Actor	Process 1	Adjunct	Process 2

Ja-trans-C45

In this example the non-Subject Participant is chosen as Theme instead of the Subject. The general word order is Subject-Participant-Verb so the changed order is understood to be the translator's choice. Although it is slightly marked, the current word order is probably better than placing the Subject (Actor) as Theme. By placing one of the most important key words 'Alzheimer's disease' at the very beginning of the long Theme, it gives a signal to the reader what the news is about. The way the English ST opens is likely to have influenced the translator's choice of doing so.

6-31 Alzheimer's disease ravages the brain, robbing its victims not only of their memories but often their ability to do things as basic as swallowing.

Now, a study of aging patients suggests its true toll may top half a million lives a year -- a figure that would put Alzheimer's just below heart disease and cancer on the list of America's top killers. En-orig-C45

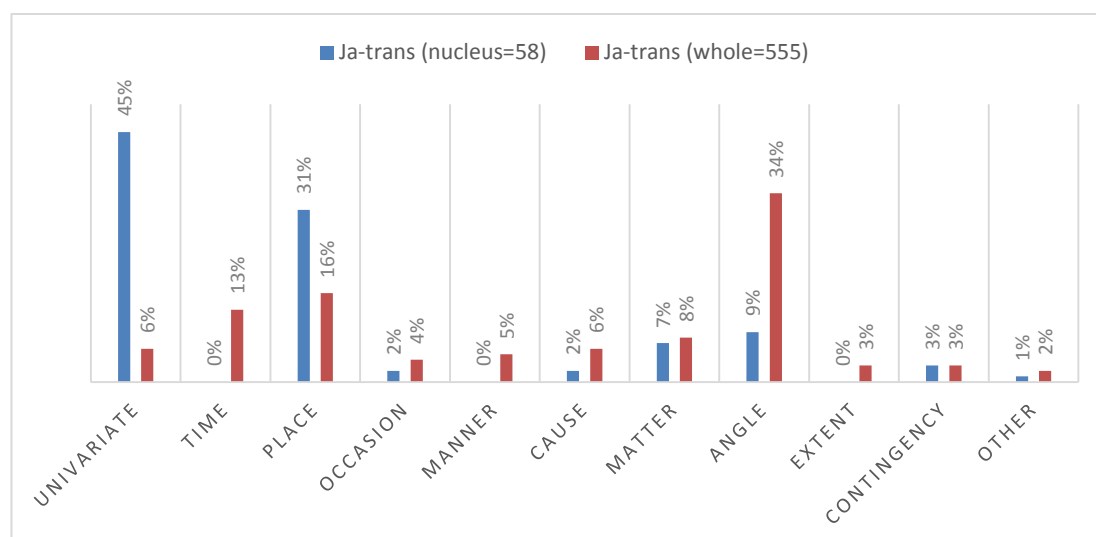
Unlike the typical one, this article delays the most important news point till the second paragraph, not placing it right at the beginning. Evans (2000) explains in his practical account of journalistic writing that this is a story-telling technique applied to news in order to give some structural variation in the newspaper.

Going back to the Japanese translation, as we have seen, it rightly takes the second paragraph as its text nucleus, but does not start with the Subject Theme that is employed in the English ST. Instead, 'Alzheimer's disease' which comes at the beginning of the ST is picked to be used in Japanese. Although it is not the Theme itself but only a part of it, 'Alzheimer's disease' is made the Topic in an effective way.

Incidentally, current non-Subject Participant Theme cannot be made the Topic in this case as it is not Given information, that is, not easily recoverable. Alternatively, the Subject 'US research team' could have been Theme and Topic, conflating two functions. However, this would reduce the impact of 'Alzheimer's disease' that the current deployment has. This example also shows it is a choice of the translator; changing the word order as well as making use of the particle *wa*, similar to the example we considered earlier.

Now I go back to the main discussion of Theme choice at the beginning of the text nucleus in Ja-trans. The next point is Circumstance Theme. I will first describe main types of Circumstance Themes chosen at the onset of the text nucleus in Ja-trans compared with the whole Ja-trans sub-corpus. I will also mention the possibility of Circumstance Themes at this location being marked by *wa*, to contrast with Subject Theme marked by it at the same location.

Figure 6-10: Circumstance Theme type at the beginning of the text nucleus (Ja-trans)



Several striking differences are visible in Figure 6-10 above as to what Circumstance is selected as Theme, at the beginning of the text nucleus in Ja-trans. The two types often chosen as Theme at this location are univariate Themes and Place. In contrast, an employment of Angle is considerably fewer compared to its generally frequent use. Also, Time is never realised as Theme at the onset of the text nucleus in comparison to the relative frequency in general. The two datasets Ja-trans (nucleus) and Ja-trans (whole) are confirmed to be significantly different ($X^2=101.844$, $df=10$, $p<.01$) in terms of the frequency of eleven types of Theme.

According to statistics, univariate Circumstance and Place are likely to occur at the beginning of the text nucleus ($p<.01$), while Time is never realised as

Theme at the same location and Angle are also unlikely to be ($p < .01$). This suggests that univariate Circumstance and Place as Theme in Ja-trans are relevant to the specific location in the text at the beginning of the text nucleus, while Time and Angle are not.

It is useful to identify that univariate Circumstances are realised as Theme almost exclusively at the beginning of the text nucleus, while Time never is and Angle very rarely. These differences could either be due to the location in the text where Theme is realised, TL influence or ST influence. I will examine these possibilities later in the chapter.

Before finishing this section, I will briefly describe the trend of using *wa* marking at the beginning of the text nucleus in Ja-trans. To put it simply, *wa* marking is almost exclusively for the Subject Theme. Out of 159 Themes at the onset of text nucleus in Ja-trans, 95 are Subject Themes, out of which 63 are marked by *wa*. Other than these Subject Themes, there is only one Circumstance Theme which is marked by *wa*.

Following the discussion of Theme choice at the beginning of the text nucleus in translation, the next section investigates Theme choice at the beginning of satellites, also in translation.

6.2.2.2 Theme at the beginning of satellite

Now the focus of location in the text is moved from the text nucleus to satellites and this section examines Theme realisation in translation at the beginning of satellites. As before, I will first address the question of the choice of Subject or non-Subject as Theme in En-trans, followed by Ja-trans. Then I will move on to discuss the options of non-subject Themes both in En-trans and Ja-trans.

Figure 6-11 below summarises the proportion of the Subject and non-Subject Themes in En-trans (satellites) along with the whole sub-corpora of En-trans as well as En-trans (nucleus). The frequency of the Subject and non-Subject Themes in En-trans at the beginning of satellites is compared with En-trans (whole) and En-trans (nucleus) to address two questions; 1) whether Theme realisation at the beginning of satellites is associated with the location in the text where it is realised, and 2) whether it is different from how Theme is realised at the beginning of the text nucleus.

Figure 6-11: Subject Theme or non-Subject Theme at the beginning of satellite (En-trans)

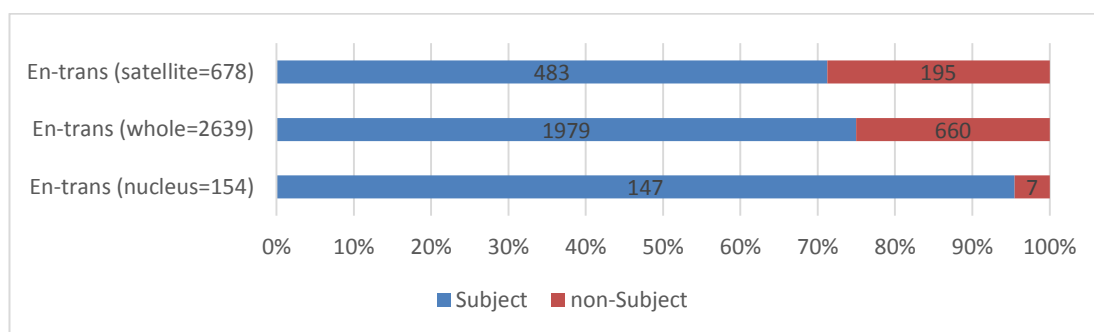
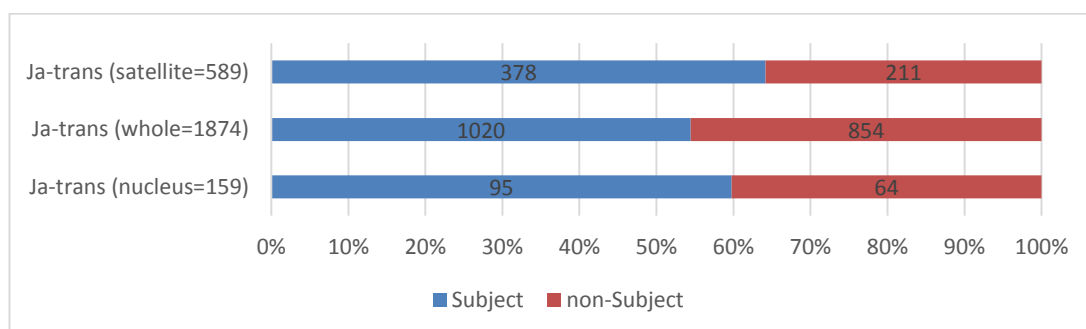


Figure 6-11 above illustrates that the proportion of the Subject Theme at the beginning of satellites is similar to that of En-trans (whole) but different from En-trans (nucleus). Statistics confirm that En-trans (satellites) and En-trans (nucleus) are significantly different ($X^2=38.722$, $df=1$, $p<.01$). Similarly, En-trans (satellites) and En-trans (whole) are confirmed to be different ($X^2=57.421$, $df=1$, $p<.01$). Theme realisation in both En-trans (satellites) and En-trans (nucleus) is likely to be related to the location in the text², at the onset of satellites in our case.

Let us now turn to Ja-trans to examine the choice between the Subject and non-Subject as Theme, keeping the location at the beginning of satellites.

Figure 6-12: Subject theme or non-Subject Theme at the beginning of satellite (Ja-trans)



Regarding the frequency of the Subject and non-Subject Themes at the beginning of satellites in Ja-trans, a trend different from En-trans emerges. The proportion of the Subject and non-Subject Themes in Ja-trans (satellites), Ja-trans (whole) and Ja-trans (nucleus) appears relatively similar to one another. Statistics show that Ja-trans (satellites) is significantly different from

² Note the location of a satellite (text unit) within text; towards the beginning, middle or towards the end, is not relevant. The focus here is the beginning of each satellite where Theme is realised.

Ja-trans (whole) ($X^2=16.954$, $df=1$, $p<.01$) but not different from Ja-trans (nucleus) ($X^2=0.874$, $df=1$, ns).

There are two implications to draw from the above results; 1) a significant difference confirmed between Ja-trans (satellites) and Ja-trans (whole) is likely to indicate that the location at the beginning of satellites is meaningful for Theme realisation, and 2) Theme realisation in Ja-trans (satellites) and Ja-trans (nucleus) are not significantly different, thus specific locations such as text nucleus and satellites may not be relevant to the way the Subject or non-Subject are realised as Theme. However, a specification of a more general level such as beginning of the text unit, instead of more specific text nucleus and satellites is probably useful.

The observations above put together, regarding the choice of Subject or non-Subject Themes at the beginning of satellites, Theme realisation at the onset of satellites is different from that at the onset of the text nucleus in En-trans. Meanwhile, in Ja-trans Theme realisation at the onset of satellites is not different from that at the onset of the text nucleus. However, both Ja-trans (satellites) and Ja-trans (nucleus) are significantly different from the whole sub-corpus in terms of the frequency of the Subject and non-Subject Themes, which leads to the assumption that the less specific location in the text as the beginning of text unit is relevant but not the distinction between satellites or text nucleus.

Before concluding this section I will introduce a nominaliser as a linguistic feature related to Theme realisation. The nominaliser is 'no' and the structure of our interest is 分裂文 *bunretsubun* [cleft sentence] 'no (wa) – da' (Noda 1996: 64-74; Sunakawa 2005: 211-213). This is similar to what we have previously seen with 'koto (ga) –' (6.2.1.1). The detailed linguistic analysis is beyond the scope of this study but what is relevant to us is that there is a clear distinction where in the news text each structure is realised. In short, 'koto (ga)' always appears in the text nucleus and 'no (wa)' always in the satellite. The nominaliser 'no' picks up information introduced earlier, typically from the previous Rheme, to set it up as a Topic. Because of this function, it is always marked by *wa* as Given and appears in the satellite to pick up information from the Rheme in the text nucleus. The example below includes such a Theme in a satellite; 'no' picks up 'rescued' in the text nucleus. As each satellite is dependent on the text nucleus but independent from other satellites, employing a linking device such as 'no wa –' is a beneficial way of starting a satellite.

6-32: Text nucleus

米ニューヨーク市のマンハッタン南部で3日夕、	[[男性がビルとビルの間狭いすき間にはさまって動けなくなっている]]のが	見つかり、
<i>Bei Nyūyōku-shi NO Manhattan nambu DE 3-ka yū</i>	<i>dansei GA biru TO biru NO aida NO semai sukima NI hasamatte ugokenaku natteiru NO GA</i>	<i>mitsukari</i>
US NY city NO Manhattan south DE 3 rd evening	man GA building TO building NO between NO narrow gap NI stick-ASP NOM GA	find-psv-pst
In the south of Manhattan in NYC, US in the evening of the 3 rd a man was found in the narrow gap between buildings and unable to move,		
Adjunct	Medium	Process
Theme	Rheme	

救急隊が	ビルの外壁の一部を	取り崩して	救助した。
<i>Kyūkyūtai GA</i>	<i>biru NO gaiheki NO ichibu O</i>	<i>torikuzushite</i>	<i>kyūjo shita</i>
Ambulance GA	building NO outside wall NO part O	break-SUSP	rescue-pst
An ambulance team broke a part of the outside wall and rescued (him).			
Actor	Goal	Process	Process
Rheme (continuing from above)			

Satellite

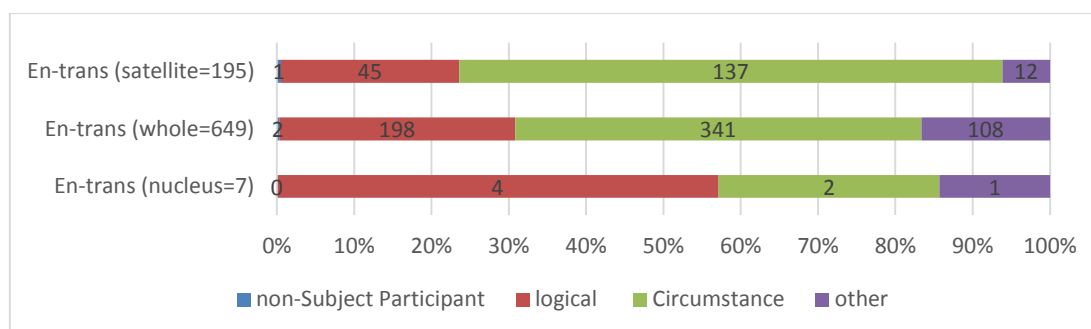
[[助け出された]]のは	19歳の男性	で、
<i>[Tasukedasareta] no WA</i>	<i>19-sai NO dansei</i>	<i>de</i>
Rescue-psv NOM WA	19 year-old man	be-SUSP
The (person) that was rescued is a 19 year old man		
Value	Token	Process
Theme	Rheme	

Although ‘no wa – da’ and ‘koto ga –’ are similar in terms of the structure, they operate differently in discourse. Regarding their functions in discourse, whether they introduce the main news or pick up a topic, and whereabouts in

text they appear, is different. They are linked to particular locations in the text and as such can be taken as a genre feature in the news genre.

I now move on to examine the choice amongst non-Subject Theme types at the onset of satellites. The focus is to observe whether non-Subject Theme realisation at the beginning of satellites is different from that at the text nucleus or in the whole sub-corpus of translation. I will start the examination with En-trans. The 'other' category in Figure 6-13 includes enhanced Theme types and verbal Predicator Themes.

Figure 6-13: Non-Subject Theme type at the beginning of satellite (En-trans)



At the beginning of satellites in En-trans, non-Subject Themes are realised differently from the whole En-trans data. In the figure above, how the frequency of the four non-Subject Theme types in En-trans (satellite) compares with En-trans (whole) is presented. Statistics inform that the two datasets are significantly different from each other ($X^2=22.960$, $df=3$, $p<.01$). The difference lies in the frequency of logical ($p<.05$), Circumstance and 'other' Themes ($p<.01$). As is clearly seen in the figure, Circumstance Themes are realised significantly more ($p<.01$) at the beginning of satellites while logical Themes are fewer ($p<.05$) as well as 'other' ($p<.01$) than En-trans (whole). This suggests that non-Subject Theme realisation at the beginning of satellites is associated with the location in the text where Theme is realised.

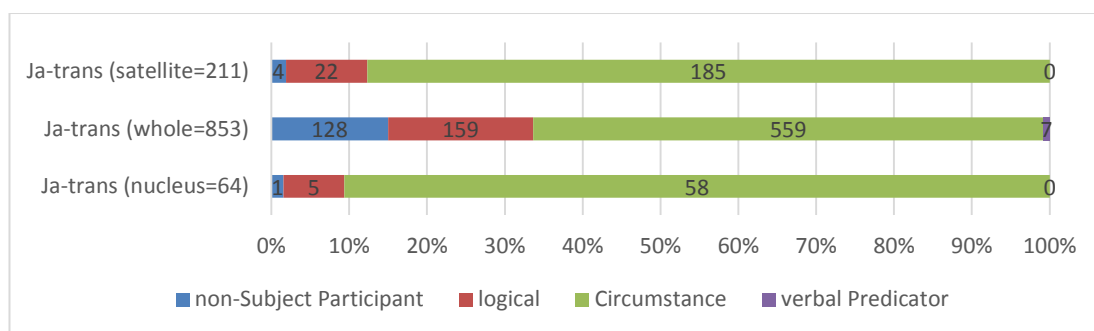
Continuing from above, I now examine if non-Subject Themes are realised differently at the beginning of satellite and the beginning of text nucleus. From Figure 6-13 above they appear to be different, which the Chi-square test does not confirm ($X^2=5.669$, $df=3$, ns). This may be due to that the variables in En-trans (nucleus) are too small.

The comparisons above show that Subject Theme realisation in En-trans at the beginning of satellites is different from the general tendency in the whole sub-corpus of En-trans as well as at the beginning of text nucleus in En-trans. The differences can arguably be attributed to the difference in location of text.

I will investigate in later sections whether or not the associations between Theme realisation and location can be established as a genre feature.

For now I turn to Ja-trans to examine non-Subject Theme realisation at the beginning of satellites. The 'other' category in Figure 6-14 below includes absolute Theme and unclassified Themes (5).

Figure 6-14: Non-Subject Theme type at the beginning of satellite (Ja-trans)

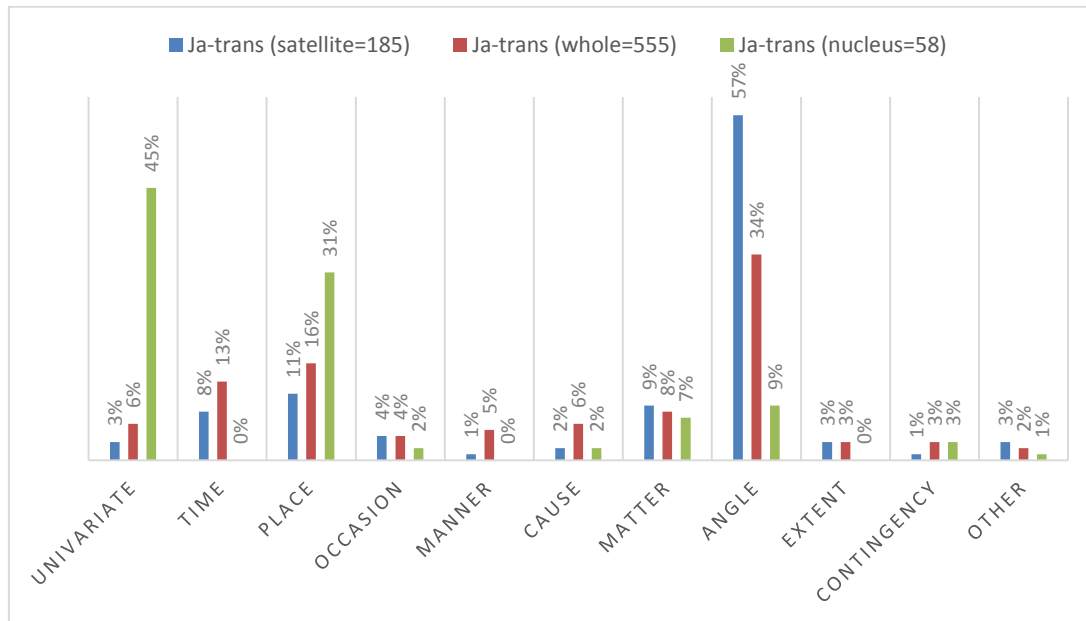


Distribution of the four types of non-Subject Themes at the beginning of satellites in Ja-trans appears somewhat different from the whole sub-corpus of Ja-trans. The two datasets of Ja-trans (satellite) and Ja-trans (whole) are confirmed to be significantly different from each other ($X^2=43.738$, $df=3$, $p<.01$). This suggests that non-Subject Theme realisation at the onset of satellites is associated with its location in the text. Statistics show the difference lies in the frequency of Participant, logical and Circumstance Themes ($p<.01$).

Meanwhile, the results of the comparison of Ja-trans (satellite) and Ja-trans (nucleus) in terms of the frequency of non-Subject Themes show that they are confirmed not significantly different from each other ($X^2=NaN00000$, $df=3$, ns). It means that the way non-Subject Themes are realised at the beginning of satellites is the same as that at the beginning of the text nucleus in Ja-trans. Hence, the distinction between the text nucleus and a satellite is not significant. We have observed the same pattern with respect to the choice between Subject and non-Subject Themes at the beginning of satellites. However, a significant difference is found between Ja-trans (satellite) and Ja-trans (whole) in terms of how non-Subject Themes are realised ($p<.01$), which means again that the location of Theme at the beginning of a text unit is meaningful but the distinction of the type of text unit, i.e. the text nucleus or satellite, does not seem to be relevant.

I now move on to examine the choice amongst Circumstance Theme types at the onset of satellites in Ja-trans.

Figure 6-15: Circumstance Theme type at the beginning of satellite (Ja-trans)



Realisation of Circumstance Themes at the beginning of satellites, which is Ja-trans (Satellite) in the figure above, shows interesting contrasts with the general tendency of their realisation, Ja-trans (whole). The most noticeable contrast occurs in the type of Angle, which at the onset of satellites is realised 57% of the time but 34% of the time in Ja-trans (whole). This is one of the Circumstance types that accounts for the difference between Ja-trans (satellites) and Ja-trans (whole). In addition to Angle, univariate Themes and Manner are also realised differently in the two datasets although the level of significance is lower than Angle at $p < .05$; so are Time, Place and Cause also have a lower level of significance ($p < .10$). In all, Ja-trans (satellites) and Ja-trans (whole) are confirmed statistically different ($X^2=51.164$, $df=10$, $p < .01$).

Practically, Angle is very likely to be realised at the onset of satellites in Ja-trans, which in turn suggests that a given text unit which starts with it is a satellite. On the other hand, univariate Themes and Manner and also to some extent Time and Cause as Theme are unlikely to be at the onset of satellites, which in turn suggests that a given text unit of which they are realised at the beginning is not a satellite. To examine these suggestions I will now turn to look at how Ja-trans (satellites) and Ja-trans (nucleus) compare with each other in terms of Circumstance Themes.

The frequency of Circumstance Themes differs between Ja-trans (satellite) and Ja-trans (nucleus) ($X^2=110.515$, $df=10$, $p < .01$). Statistics indicate that in terms of univariate Theme, Place and Angle, two datasets are found to be significantly different ($p < .01$), as they also are in terms of Time but at a lower

level of significance ($p < .05$). The practical interpretation of these results is that univariate Circumstance and Place are likely to occur at the beginning of the text nucleus while Angle and Time hardly ever are.

Putting the two comparisons above together, I have found that Circumstance Theme realisation at the onset of satellites in Ja-trans differs from the whole sub-corpus of Ja-trans as well as that at the beginning of the nucleus. It is very likely to be associated with the specific locations where these Themes are realised and they can potentially be indicators for either the text nucleus or satellites by their prominence or absence in them.

6.2.2.3 Comparison of En-trans and Ja-trans

This section summarises the main findings regarding the way Theme is realised in translation at specific locations in the text; at the beginning of the text nucleus and at the beginning of satellites. Focusing on the text nucleus first then moving to satellites, findings relating to each location gained in En-trans and Ja-trans are summarised to make a comparison. This section also aims to answer one of the research questions; Are Themes realised at the beginning of the text nucleus distinguishable from those at the beginning of the satellite?

I will start with the location at the beginning of the nucleus. Regarding the choice of Subject or non-Subject Theme, En-trans indicate a strong preference towards the Subject Theme. Subject as Theme is an unmarked and natural option in English. However, the inclination for the Subject Theme here is stronger than the general trend of choosing it and it is possible the move towards Subject as Theme is intensified by the location at the beginning of the text nucleus.

In contrast to En-trans, Theme choice of Subject or non-Subject at this location in Ja-trans does not differ from the general trend observed without relation to a specific location in the text. However, in terms of non-Subject types, Theme realisation at the onset of the text nucleus indicates a difference from the way these Themes are realised throughout the Ja-trans data.

Moving on to the location at the beginning of satellites, regarding the frequency of the Subject and non-Subject Theme, En-trans (satellites) are significantly different from En-trans (nucleus). Also in terms of non-Subject Themes these two datasets are statistically different. In all, in En-trans, Theme realisation at the beginning of satellites is different from that of En-trans (nucleus). However, it will be difficult to distinguish a Theme in a satellite from another in the text nucleus because the choices of Theme are limited to either

the Subject or non-Subject, and, for instance, Subject can easily be Theme in either of them.

Turning to Japanese, at this location on the choice of Subject or non-Subject Ja-trans (satellites) are marginally different from Ja-trans (nucleus). However, with regards to the choice of non-Subject, the two datasets are significantly different. With these Themes however, it will be difficult to distinguish a Theme in the text nucleus from another in a satellite. Meanwhile, with Circumstance Themes, it will be possible to tell the text nucleus from a satellite if univariate Theme appears. Angle as Theme in a text unit is very likely to indicate a satellite. Place often appears as Theme in the text nucleus but it is not a strong indicator as it also appears in a satellite or any other location in the text.

Finally, to answer the research question placed at the beginning of this section, in general it is very difficult to distinguish a Theme as realised at the beginning of the nucleus or satellite with Subject, non-Subject, non-Subject types and so on. However, certain types of Circumstances in Japanese are identified exclusively appearing either in the text nucleus or satellite. Univariate Themes almost exclusively appear at the beginning of the text nucleus while Angle is typically realised at the beginning of satellites.

6.2.3 Translation shift in thematic progression

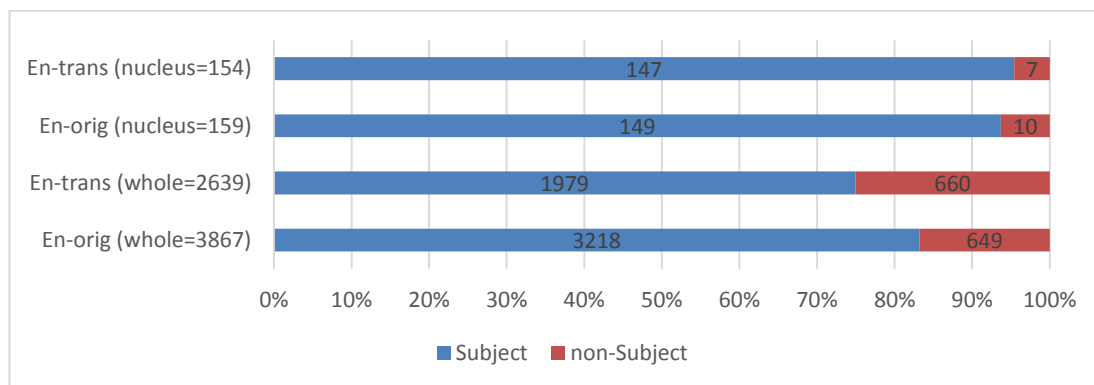
So far in the current chapter I have examined Theme realisation in non-translated news and translated news focusing on different locations in the text. This section will examine whether shifts occur in translation in terms of how Theme is realised at specific locations in the text. If shifts occur, I will seek to identify patterns of these. For these purposes I will investigate; 1) how Theme realisation in translation compares with that observed in the target language (TL) original corpus, 2) how Theme realisation in translation compares with that observed in the source text (ST) corpus. The Theme realisations in question are at the beginning of the text nucleus and satellites.

6.2.3.1 Comparison of translation and target language original

The main aim of this section is to observe major differences between translations and TL originals in terms of Theme realisation. Focusing on the beginning of the text nucleus and the beginning of satellites, different Theme realisations in translations and in TL originals are observed to find possible environments for translation shifts.

I will start looking at Theme at the beginning of the text nucleus in En-tran in comparison with En-orig, and move on to Ja-trans to compare with Ja-orig.

Figure 6-16: Subject Theme or non-Subject Theme at the beginning of the text nucleus (En-trans & En-orig)



Earlier in the chapter (6.2.2.1) we have seen that at the beginning of the text nucleus in En-trans the Subject Theme is chosen significantly more often than it generally is in the whole sub-corpus, which shows Theme realisation regardless of the location in the text. I now compare En-trans with En-orig to observe the choice between the Subject and non-Subject Themes at the beginning of the text nucleus. The proportion of the Subject and non-Subject Themes appears similar for En-trans (nucleus) and En-orig (nucleus) in Figure 6-16, both showing that the Subject is chosen in around 95% of cases. The two datasets are confirmed as not different ($X^2=0.013$, $df=1$, ns). This means the En-trans data is considered the same as En-orig, showing no trace of translation, as far as Theme choice between the Subject and non-Subject at the beginning of the text nucleus is concerned.

Hence, En-trans and En-orig share the same trend in which the Subject Theme is realised at the beginning of the text nucleus. I have recognised that the preference of the Subject Theme in En-orig is strengthened at the beginning of the text nucleus (6.2.1.2). This may be a form of 'standardization' (Toury 1995: 274-279) but also be due to the location in the text. Looking at the relationship of En-trans (nucleus) and En-trans (whole), and that of En-orig (nucleus) and En-orig (whole) in Figure 6-16 above, I notice a similarity with respect to the frequency of Subject and non-Subject Themes. It is possible that En-trans takes the TL norm with regard to Subject Theme at the beginning of the text nucleus.

Let us now look at the Japanese data to see how Ja-trans compares with Ja-orig in terms of the choice between the Subject and non-Subject Themes at the beginning of the text nucleus.

Figure 6-17: Subject or non-Subject Theme at the beginning of the text nucleus (Ja-trans & Ja-orig)

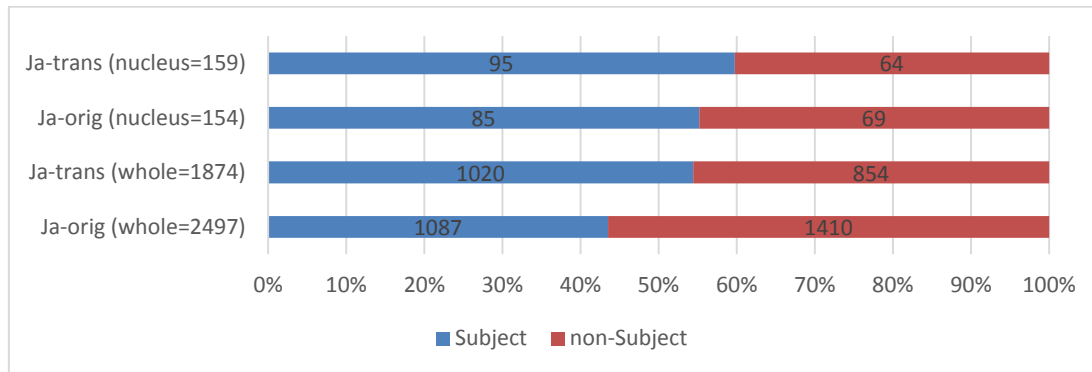


Figure 6-17 above indicates that Ja-trans (nucleus) displays a similar proportion for the Subject Theme and non-Subject Theme as Ja-orig (nucleus). The two datasets are not significantly different from each other ($X^2=1.081$, $df=1$, ns). In short, in Ja-trans Theme choice between the Subject and non-Subject at the beginning of the text nucleus is unlikely to be influenced by the location where it is realised in the text. So Ja-trans and Ja-orig share the same tendency in terms of Theme realisation at the beginning of the text nucleus, which is to select both Subject and non-Subject as Theme more or less equally.

Following the distinction of Subject and non-Subject Themes, I now discuss non-Subject Themes in Ja-trans, keeping the location at the beginning of the text nucleus.

Figure 6-18: Non-Subject Theme type at the beginning of text nucleus (Ja-trans & Ja-orig)

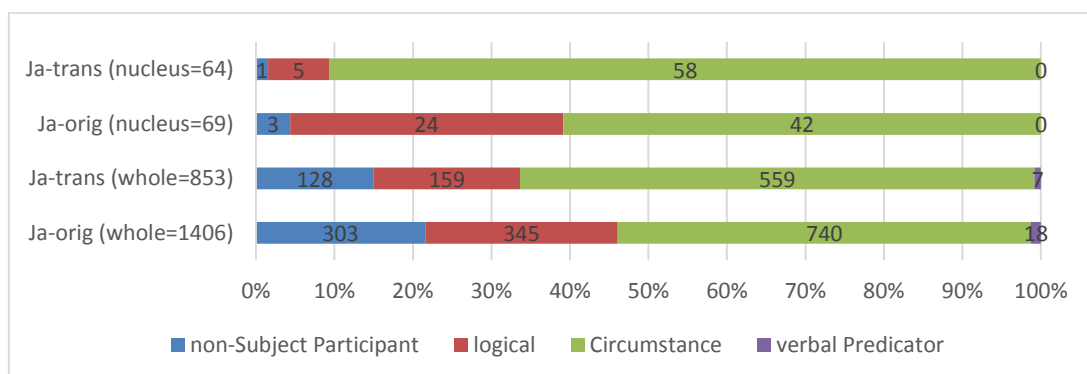
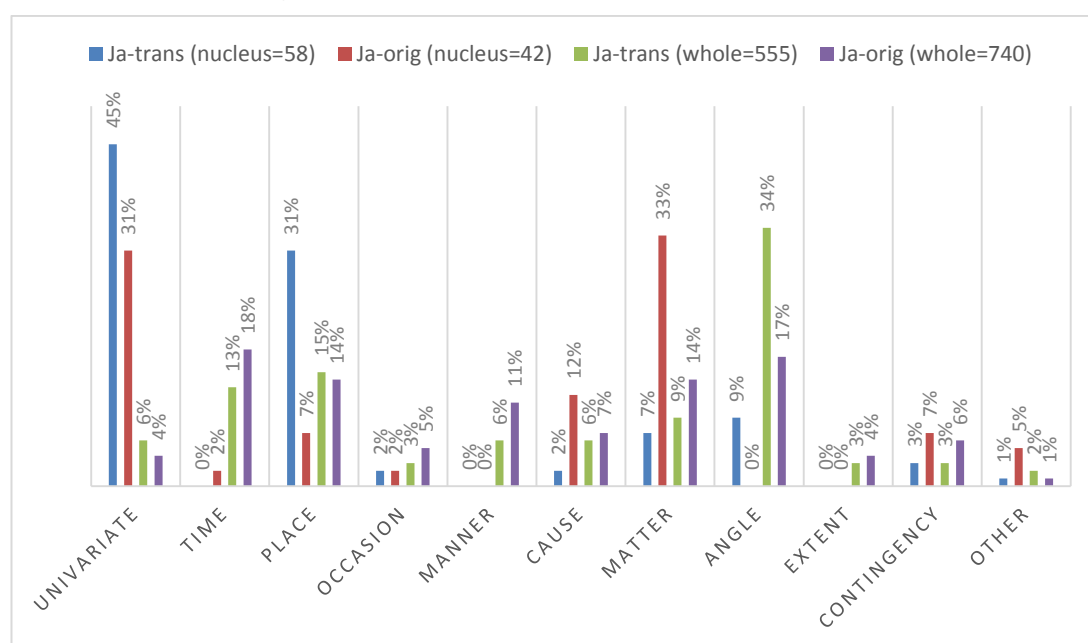


Figure 6-18 above shows that the proportion of non-Subject Theme types in Ja-trans (nucleus) is dissimilar to Ja-orig (nucleus). The Chi-square test confirms Ja-trans (nucleus) and Ja-orig (nucleus) are significantly different ($X^2=15.843$, $df=2$, $p<.01$). According to statistics their difference is in terms of the frequency of logical and Circumstance Themes ($p<.01$), which is clearly visible in Figure 6-18 above. A large proportion of non-Subject Themes at the

beginning of the text nucleus in Ja-trans is Circumstances. Types of Circumstance Theme will be discussed in the section below. For now I suggest that the realisation of non-Subject Themes in Ja-trans is linked to the location where they appear in the text, i.e. at the beginning of text nucleus.

Let us now turn to Circumstance Theme types to observe if they display a similar tendency in the way they are realised at the beginning of the text nucleus in Ja-trans (nucleus) & Ja-orig (nucleus). I have already discussed what types of Circumstance Theme are likely or unlikely to be realised at the onset of the text nucleus in Ja-orig (6.2.1) and Ja-trans (6.2.2) separately. Figure 6-19 below combines the results from those.

Figure 6-19: Circumstance Theme type at the beginning of text nucleus (Ja-trans & Ja-orig)



There are two noticeable features that take place in both Ja-orig and Ja-trans at the beginning of the text nucleus. Firstly, univariate Themes seem prominent, showing frequent appearances both in Ja-trans and in Ja-orig. On the other hand, Angle never appears at the onset of the text nucleus in either of the datasets. Their prominence or absence at the beginning of the text nucleus suggests that this location at the beginning of the text nucleus is meaningful. I can confidently say that univariate Themes are likely to indicate the text nucleus while Angle means a satellite. Given that both Ja-orig and Ja-trans share the same tendency here the assumption is that translation follows the TL norm to some extent.

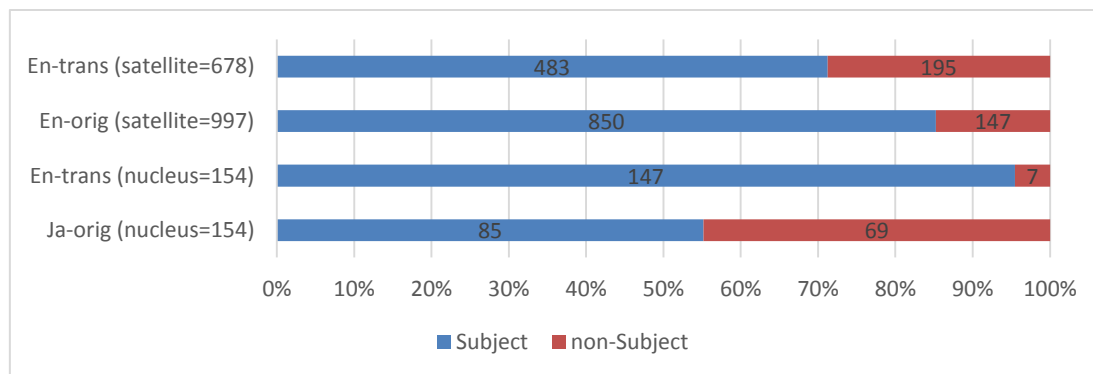
Incidentally, there are differences in how the same Circumstance types are realised in Ja-orig and Ja-trans. One is the way Matter is realised and the

other Place. Matter is realised as Theme in Ja-orig more often than in Ja-trans, while Place appears in Ja-trans more often than in Ja-orig.

Speculative reasons for Place Themes to appear as such may be relevant to the timelag between the publication of the original news and its translation. Not all the translations are produced immediately after the original news either because of the possibility or necessity. Either way the relevance of the time of the news event is reduced or not very high from the beginning in the translation.

So far this section has focused on the beginning of the text nucleus to observe Theme realisation in order to find similarities and differences between the translation and TL original. The following part will discuss Theme realisation at the beginning of satellites, still with the aim of finding similarities and differences between translations and TL originals.

Figure 6-20: Subject or non-Subject Theme at the beginning of satellite (En-trans & En-orig)



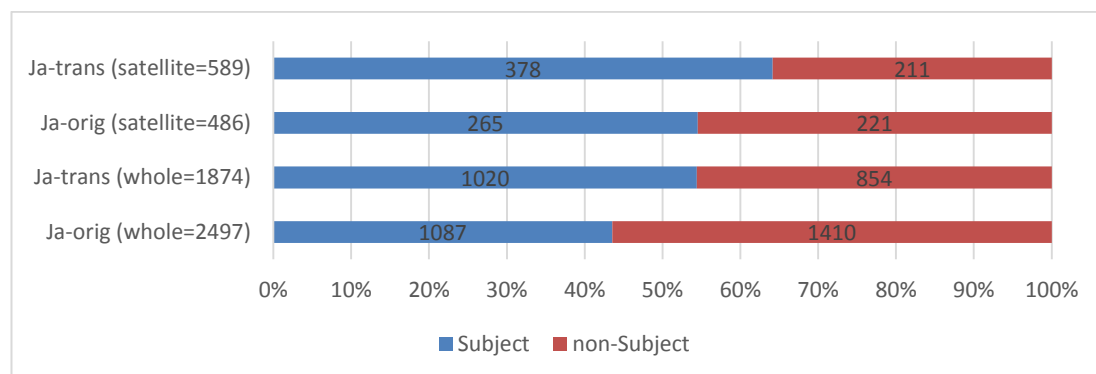
Earlier I examined Theme realisation at the onset of satellites in En-trans and also in En-orig separately to find out if it is different from that at the onset of the text nucleus in both cases. Now I compare En-trans (satellites) and En-orig (satellites) to see if they share the same tendency of Theme realisation at this location. If they do, the two datasets are considered the same with regards to how the Subject and non-Subject Themes are realised at this location, which in turn suggests that En-trans takes the TL norm, a form of 'standardization' (Touy 1995: 274-279). Statistics show that they are significantly different ($X^2=33.275$, $df=1$, $p<.01$). A possible reason for the relatively high frequency of non-Subject Themes in En-trans is the ST influence.

Theme realisation in En-trans at the beginning of satellites is different from the beginning of the text nucleus, which is likely to mean that the location in the

text is meaningful to the way Theme is realised. However, in this case it is not the same trend as the TL shows.

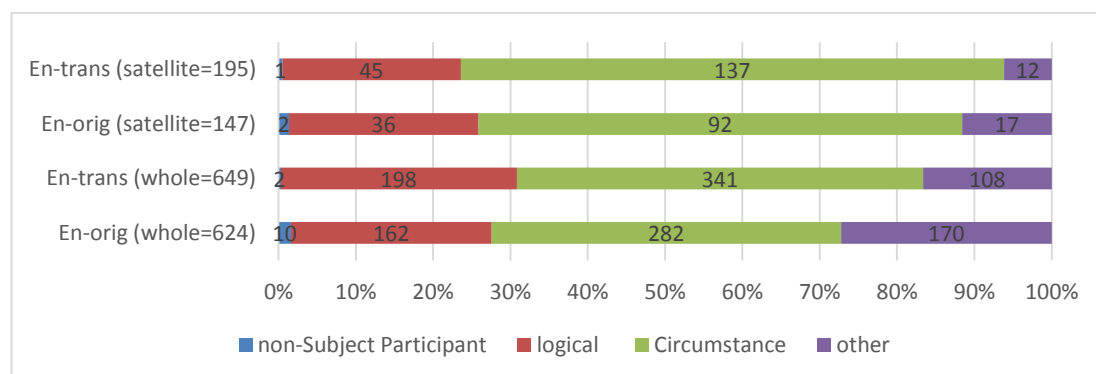
I now look at the Japanese data and observe the relationship between Ja-trans (satellites) and Ja-orig (satellites) in terms of the frequency of the Subject and non-Subject Themes.

Figure 6-21: Subject or non-Subject Themes at the beginning of satellite (Ja-trans & Ja-orig)



Although Ja-trans (satellites) and Ja-orig (satellites) seem to display a similar proportion of Subject and non-Subject Theme at the onset of satellite in Figure 6-21 above, they are found statistically different ($X^2=14.043$, $df=1$, $p<.01$). Statistics indicate that a non-Subject is chosen as Theme more frequently in Ja-orig (satellites) ($p<.05$) while in En-trans it is the Subject that is more often chosen as Theme ($p<.05$). That the Subject Theme is preferred in Ja-trans is likely to be a result of English ST influence. Translation shifts may have occurred here and this will be examined later in the chapter. For now let us move on to discuss non-Subject Theme choice at the beginning of satellites, first in En-trans, which is followed by Ja-trans.

Figure 6-22: Non-Subject Theme type at the beginning of satellite (En-trans & En-orig)

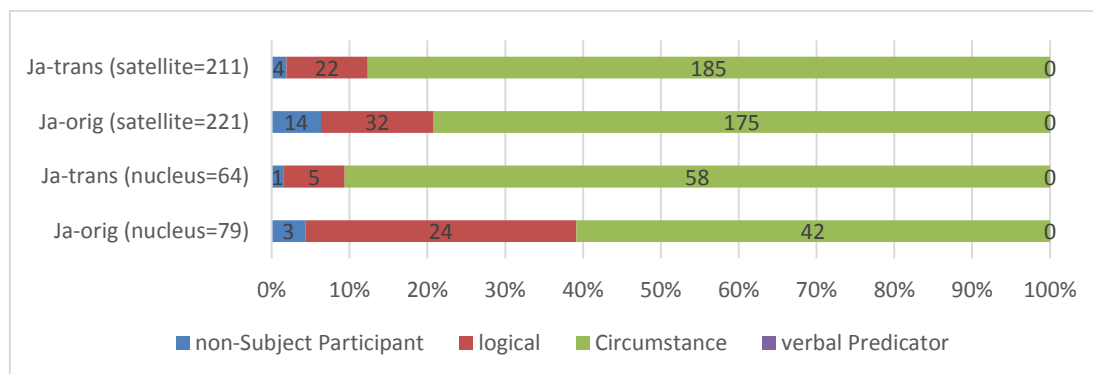


With respect to non-Subject Theme realisation, both En-trans (satellites) and En-orig (satellites) have been identified to be significantly different from the

corresponding sub-corpora (whole) (6.2.2.1 & 6.2.2.2). On the same question I will examine whether En-trans (satellites) is similar or different from En-orig (satellites) in the way how non-Subject Theme is realised. The Chi-square test finds them to be not different ($X^2=4.388$, $df=3$, $p<.01$). En-trans shows more Circumstance and 'other' Themes ($p<.05$) and fewer logical Themes than En-orig ($p<.05$). Hence, even though non-Subject Theme realisation in En-trans seems to be influenced by the location at the beginning of satellites, it does not share the same tendency with the TL, which consequently means a process of standardization (Toury 1995: 274-279) is not occurring.

We will now look at Japanese, addressing the same question.

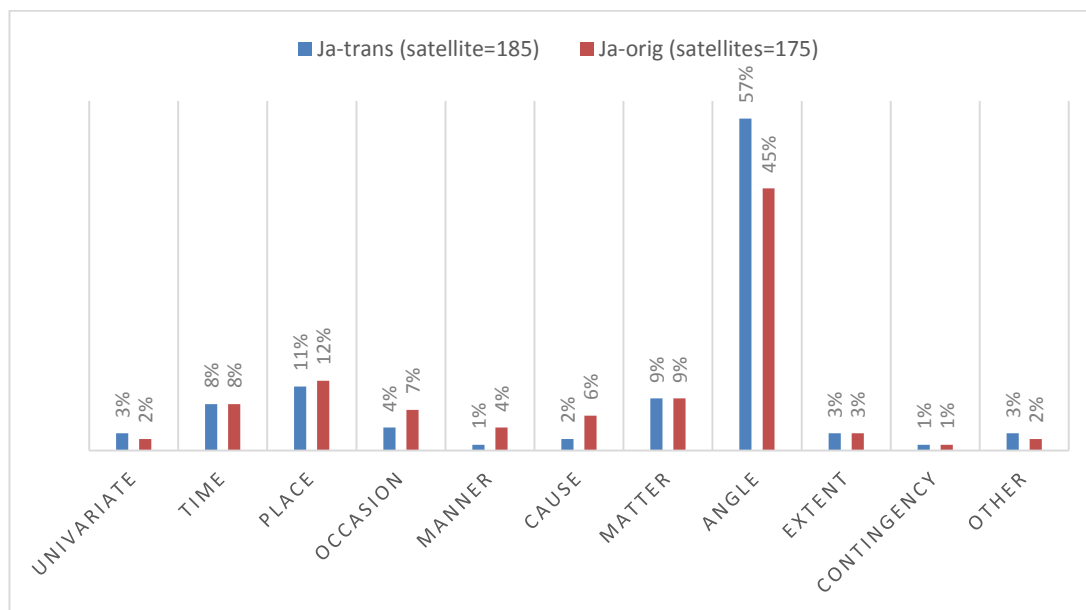
Figure 6-23: Non-Subject Theme type at the beginning of satellite (Ja-trans & Ja-orig)



Ja-trans (satellite) and Ja-orig (satellite) are found to be statistically different ($X^2=7.458$, $df=2$, $p<.05$). The difference is relevant to how Participant and Circumstance Themes are realised ($p<.05$). In practical terms, this means Circumstance Themes are more in use in Ja-trans while Participant Themes are less ($p<.05$). So in terms of non-Subject Theme realisation at the onset of satellites, Ja-trans and Ja-orig do not share the same patterns. Both datasets show an indication that the location in the text is associated with the way non-Subject Themes are realised. However, in our case here each sub-corpus has a different manner for non-Subject Themes to be realised at the beginning of satellites. Neither process of 'standardization' nor 'interference' (Toury 1995: 274-279) are recognised.

Before moving on to the following section to compare translation with the ST, I have one more set of Theme types of which the realisation is of interest in this section, namely Circumstance Theme.

Figure 6-24: Circumstance Theme type at the beginning of satellite (Ja-trans & Ja-orig)



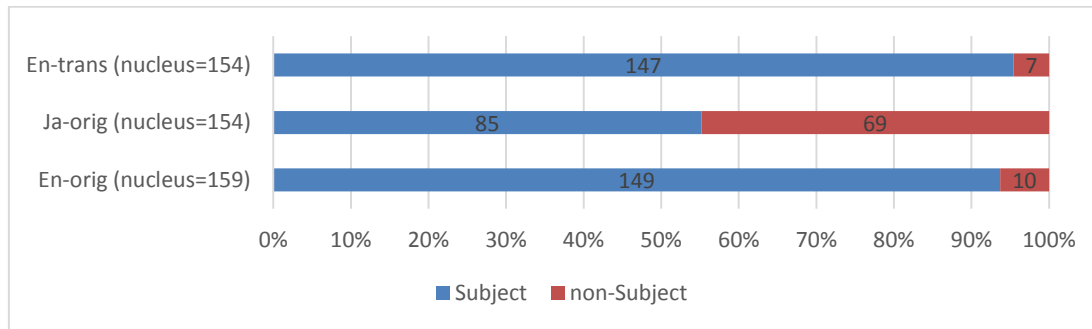
With respect to Circumstance Themes, Ja-trans (satellite) and Ja-orig (satellites) are confirmed to be statistically different ($X^2=16.953$, $df=10$, $.05 < p < .10$). In Figure 6-24 there are several interesting issues. In both Ja-trans and Ja-orig, Angle is realised as Theme prominently, while univariate Themes are hardly realised at this location. This appears to be reverse from the way they are realised as Theme at the beginning of the text nucleus (6.2.1.1 & 6.2.2.1). Angle and univariate Themes have been discussed earlier as possible text unit indicators. Current examination confirms that univariate Theme indicates the text nucleus and Angle indicates a satellite.

6.2.3.2 Comparison of translation and source text

The main aim of this section is to observe major differences between the translation and the ST in terms of Theme realisation. Focusing on the beginning of the text nucleus and then the beginning of satellites, different Theme realisations in translation and in the STs are observed to find possible environments for translation shifts.

Firstly, I look at the relationship of En-trans (nucleus) and Ja-orig (nucleus). En-orig (nucleus) is included in the following Figure 6-25 for reference.

Figure 6-25: Subject Theme or non-Subject Theme at the beginning of the text nucleus (En-trans & Ja-orig)



With regards to how Subject or non-Subject Themes are realised at the beginning of the text nucleus, Figure 6-25 above shows that En-trans (nucleus) and Ja-orig (nucleus) are dissimilar. The two datasets are indeed confirmed to be statistically different ($X^2=74.880$, $df=1$, $p<.01$). The difference is clearly seen in the figure; non-Subject Themes are employed 46% of the time in Ja-orig, while only 5% in En-trans. This means that non-Subject Themes in Ja-orig are not reserved in En-trans, instead Subject Themes are chosen; a form of translation shift takes place. Since Subject Theme is an unmarked choice in English, this indicates that a process of ‘standardization’ (Toury 1995: 274-279) or ‘normalization’ (Baker 1996) occurs, reflecting the norm of the TL.

Most of the non-Subject Themes in Ja-orig are replaced by the Subject in En-trans but several remain. In the following part, in an attempt to ascertain whether the choice of non-Subject Themes in En-trans is a result of translation strategy or not, I will consider possible reasons for these based on the examples obtained from the data. The number of occurrences is small (8) and there may be differences between the practices of *Yomiuri* and *Mainichi*; seven articles are from *Yomiuri* and one from *Mainichi*. Hence, the observation is speculative and cannot be over-generalised.

Between the non-Subject Theme in En-trans and the Theme in the original ST three types of relationship are identified from the examples. The most understandable is an equivalent non-Subject pair. The second is also a non-Subject Theme pair which do not correspond with each other. In the third the TT employs a non-Subject Theme while a Subject Theme is found in the ST. We will see examples for each category, followed by brief explanations.

6-33 non-Subject Theme in En-trans (equivalent to ST)

Although rapid economic growth slowed further in the July-September quarter, the nation’s real gross domestic product, excluding the effects of price fluctuations, recorded growth of 0.5 percent, marking positive

growth for four quarters in a row, the Cabinet Office said Thursday in a preliminary report. En-trans-Y23

Although translating the non-Subject Theme to the equivalent non-Subject Theme in translation is the most straightforward approach, there are only two examples of this kind including 6-33 above. On the one hand it is strange not to keep to the original non-Subject Theme if a non-Subject Theme is employed in En-trans. On the other hand it suggests that the translator avoids translating non-Subject Themes mechanically into English being aware of its preference of the Subject.

The second relationship is the not-equivalent non-Subject Themes in the TT and ST. An example is presented below, followed by the original ST.

6-34 non-Subject Theme in En-trans (not equivalent to ST)

With an International Olympic Committee advance team currently visiting Tokyo, the Tokyo Organizing Committee of the Olympic and Paralympic^{sic} Games shifted into top gear with their exchange of opinions about how to coordinate preparations Thursday.

En-trans-Y56

2020年東京五輪・パラリンピックの準備状況などについて、国際オリンピック委員会 (IOC) が日本側に確認する本格折衝が3日、東京都内で行われた。

About the preparation situations for the 2020 Tokyo Olympics and Paralympics, a full-scale negotiation was held in Tokyo on the 3rd, in which an International Olympic Committee (IOC) asked Japan side questions

Ja-orig-Y56

The non-Subject Theme in the original ST is translated into English employing a non-Subject Theme, but not an equivalent one. En-trans-Y56, one of the two examples of this type in the corpus, starts the nucleus with a clause Theme but the Theme of the nucleus in Ja-orig-Y56 is a Circumstance Matter. Although the number of instances are too small to generalise, these examples show that even though it seems natural to keep the non-Subject Theme that appears in the ST in translation, that option is not often taken.

Meanwhile, a Subject Theme in the original Japanese ST is not kept when translated into English. There are four examples including the one below, which is followed by its original ST.

6-35 non-Subject Theme in En-trans (Subject Theme in ST)

Aiming to develop an artificial intelligence robot smart enough to pass a University of Tokyo entrance exam, a group of researchers has been

working on a project to find out its standard score and ascertain its chances of entering the nation's top university. En-trans-Y14

国立情報学研究所(東京都千代田区)などが開発した人工知能が先月から、大手予備校が作った東京大入試などの模擬試験に取り組んでいる。

An artificial intelligence that the National Institute of Informatics (Tokyo Chiyoda-ward) and others developed has been taking mock examinations. Ja-orig-Y14

The Subject Theme at the beginning of the text nucleus in Ja-orig-Y14 (underlined) is not kept in En-trans-Y14, which starts the nucleus with a clause Theme (underlined). In fact, the text nucleus of En-trans-Y14 presented above has been added during translation, rather than the translation from the ST text nucleus. Example 6-36 below is the text unit that follows the text nucleus, i.e. the first satellite, which is in fact the ST's text nucleus which also includes information moved from elsewhere (underlined) in the text.

6-36 Under the project titled "Can a Robot Pass the University of Tokyo (Today) Entrance Exam?" led by Prof. Noriko Arai at the National Institute of Informatics, the robot developed by the institute and other entities has been taking mock entrance examinations since last month, including those for the University of Tokyo.

The original Subject Theme could have been kept in translation but that option has not been chosen. Possible reasons for this include the influence of the fixed β^{α} order in Japanese (Teruya 2006) or the translator's intention to place New information at the unmarked position towards the end of the clause (Fries 1994: 233, Halliday & Matthiessen 2004: 89).

Summing up the above, examples of non-Subject Themes in En-trans at the beginning of the text nucleus do not seem to offer any clear reason for why they remain in translation rather than being translated utilising less marked Theme. It is particularly difficult to explain why a Subject Theme in ST is not used in En-trans, which is a natural course. What can be maintained at least is that examples from En-trans show a resemblance in function to non-Subject Themes that occur in En-orig as below.

6-37 non-Subject Themes in En-orig

Amid a tidal wave of negative publicity, a Colorado school system has let a 6-year-old boy return to school and said it won't classify his kissing a girl on the hand as sexual harassment. En-orig-C30

Despite budget uncertainties, NASA on Tuesday issued a solicitation for a commercially operated space taxi to ferry astronauts to the International Space Station, an attempt to break Russia's monopoly on crew transport by 2017. En-orig-R22

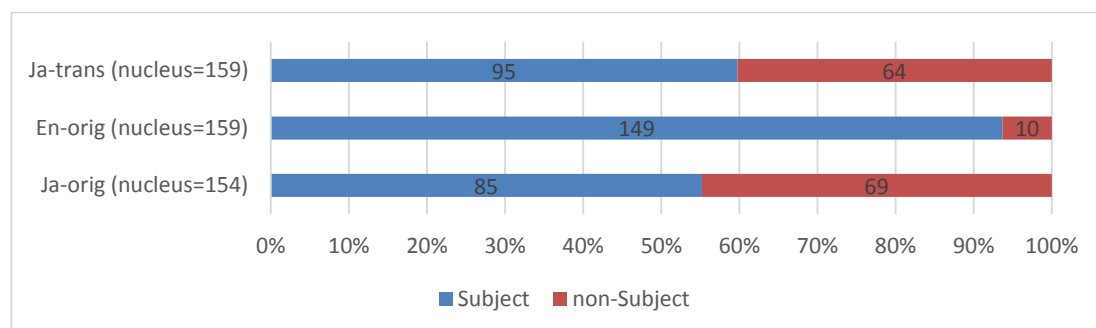
In a major milestone for gay rights, the U.S. government on Monday expanded recognition of same-sex marriages in federal legal matters, including bankruptcies, prison visits and survivor benefits. En-orig-C33

After more than three years of jousting with the White House press corps, President Barack Obama's chief spokesman, Jay Carney, will step down next month to pursue yet-to-be-named opportunities, Obama announced on Friday. En-orig-R80

Although the examples taken from En-orig are more concise than those from the En-trans data, these Themes are very similar in terms of their function. Fries (1994: 232) defines Theme as 'providing framework for interpretation' and although it applies to Theme in general and is not limited to non-Subject Themes under discussion here, the explanation fits well to describe the Themes above. Being chosen as Theme over the unmarked Subject, non-Subject Themes effectively assist the reader to understand the information presented in Rheme.

Let us now move on to compare Ja-trans with En-orig, which is the sub-corpus of STs of Ja-trans. Figure 6-26 below shows a clear difference between the two datasets in terms of the proportion of the Subject and non-Subject Themes.

Figure 6-26: Subject Theme or non-Subject Theme at the beginning of the text nucleus (Ja-trans & En-orig)

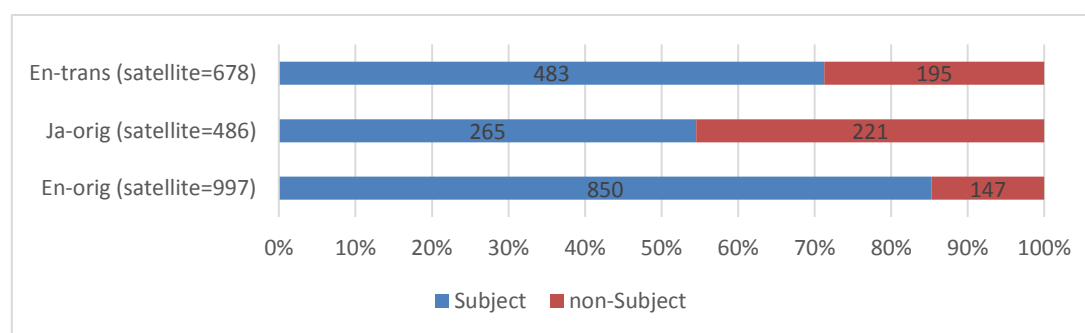


When Ja-trans and En-orig are compared with each other focusing on the Theme realisation at the beginning of the text nucleus, they are found to be significantly different ($X^2=49.472$, $df=1$, $p<.01$). In Ja-trans the Subject and non-Subject are realised as Theme more or less equally frequently while in

En-orig a strong preference for the Subject Theme (94%) is observed over non-Subject Theme (6%). This suggests that in Ja-trans a non-Subject Theme is chosen even when the original ST has the Subject Theme. This is a clear sign of shifts. Shifts here refers to a process of ‘standardization’ (Touy 1995: 274-279) as what occurs in translation is to employ non-Subject Themes even when the ST starts the text nucleus with a Subject Theme. Ja-trans and En-orig are found to be not significantly different ($X^2=0.491$, $df=1$, ns).

Moving the focus of location from the text nucleus to satellites, En-trans are compared with Ja-orig, their parallel ST corpus. The aim of this is to examine whether ST influences can be observed in the way the Subject and non-Subject Themes are realised in En-trans. I will first discuss En-trans, and then turn to Ja-trans.

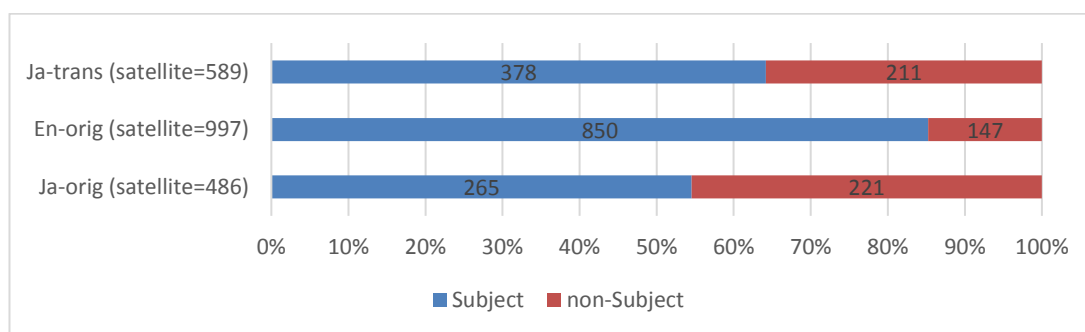
Figure 6-27: Subject Theme or non-Subject Theme at the beginning of satellite (En-trans & Ja-orig)



Statistics confirm En-trans (satellites) and Ja-orig (satellites) are significantly different ($X^2=33.702$, $df=1$, $p<.01$). They also indicate that non-Subject Themes are realised significantly more in Ja-orig, compared to En-trans ($p<.01$). Considering that the Subject is an unmarked Theme, the figure suggests a process of ‘standardization’ (Touy 1995: 274-279) is at work. At the same time, however, it is possible to observe the ST’s ‘interference’ (*ibid.* 267-274) or the ‘ST shining through’ (Teich 2003: 209) because non-Subject Theme is chosen more frequently in En-trans than in En-orig ($p<.05$), and more in Ja-orig than in En-trans, indicated as En-orig > En-trans > Ja-orig.

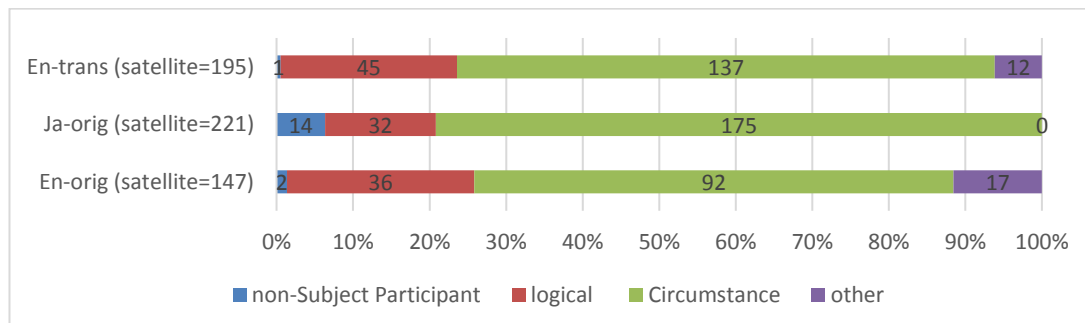
Now, let us compare Ja-trans and En-orig in terms of the choice of Subject or non-Subject Theme at the beginning of satellites to see if standardization or interference can be observed similarly to En-trans satellites.

Figure 6-28: Subject Theme or non-Subject Theme at the beginning of satellite (Ja-trans & En-orig)



Ja-trans and En-orig are confirmed to be significantly different ($X^2=92.930$, $df=1$, $p<.01$). Statistics also show that non-Subject Themes are realised significantly more in Ja-trans, compared to En-orig ($p<.01$). This suggests that the ST interferes in translation (Toury 1995: 274-279). Just as with what has been seen with En-trans, it can also be said that the process of 'standardization' (*ibid.* 267-274) is at work here too as the Subject is chosen as Theme more frequently in Ja-trans than Ja-orig ($p<.05$), and more in En-orig than En-trans.

Figure 6-29: Non-Subject Theme type at the beginning of satellite (En-trans & Ja-orig)



We saw earlier that non-Subject Themes in En-trans (satellites) are realised differently from En-orig (satellites). Here I will observe whether the difference between En-trans (satellites) and En-orig (satellites) can be related to the ST of En-trans, which is Ja-orig. The Chi-square test shows that En-trans (satellites) and Ja-orig (satellites) are significantly different ($X^2=28.576$, $df=3$, $p<.01$). The difference lies in how Participant Themes and 'other' Themes are realised ($p<.01$), while no difference is found in terms of Circumstance Themes (ns).

Of these, 'other' Themes in En-trans, which are enhanced Themes and verbal Predicator Themes, occur due to the linguistic difference between English and Japanese. There are no clause patterns comparable to those in Japanese whereby enhanced Themes are realised in English. The frequency of

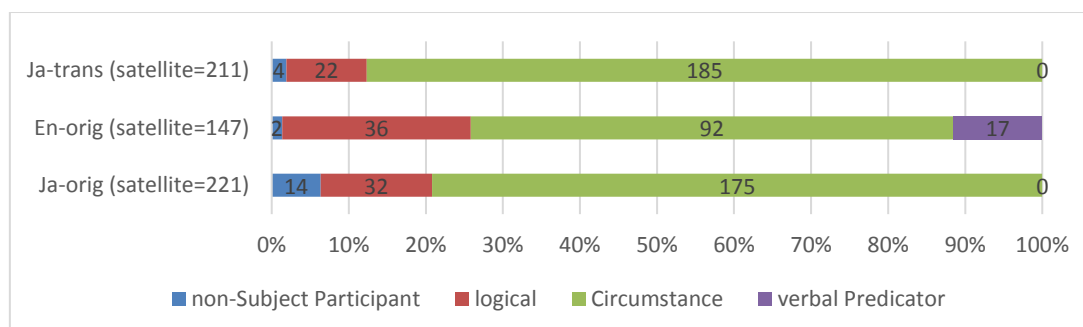
Participant Theme realisation is different because of the different level of markedness of the type; this is not uncommon in Japanese but it is the 'most' marked Theme in English (Halliday & Matthiessen 2004: 73). It is very likely that the translator strategically avoids Participant Themes in translation.

Circumstance Themes are also marked in English while they are very common in Japanese. Therefore, it is surprising no significant difference is found in the way they appear between En-trans and Ja-orig. Again this may be because of the level of markedness. Circumstance Themes are marked in English but less so compared to Participant Themes.

In summary, non-Subject Theme realisation in En-trans at the beginning of satellite is different from that of Ja-orig, partly because of the difference in the linguistic system and partly because of the level of markedness of a particular choice.

Before ending this section I will look at the relationship between Ja-trans (satellites) and Ja-orig (satellites) with respect to non-Subject Theme types.

Figure 6-30: Non-Subject Theme type at the beginning of satellite (Ja-trans & En-orig)



The two datasets are found to be statistically different ($X^2=42.176$, $df=2$, $p<.01$). The difference is relevant to how Participant and Circumstance Themes are realised ($p<.01$). In practical terms, this means Circumstance Themes are significantly more in use in Ja-trans while Participant Themes are significantly less ($p<.05$).

6.2.4 Patterns and motivations of shifts in thematic progression

Previous sections have been examining and discussing Theme realisation at specific locations in the text, during which indications for shifts have appeared. In this section I will report several possible patterns of translation shift in thematic progression.

The first situation in which translation shifts are likely to occur is where the two languages involved have Theme systems with similar options available for a particular choice but their usage, value or markedness are different. Subject,

non-Subject Participant and Circumstance Themes appear as examples of these in my data.

On the question of Subject or non-Subject Theme, both En-trans and Ja-trans indicate significant difference from their respective ST sub-corpora, which means a shift has taken place. Both the Subject and non-Subject Theme types are available within the systems of English and Japanese. However, although the Subject is an unmarked Theme in both languages, it is not the only unmarked Theme in Japanese as Circumstances are also unmarked (Teruya 2006). Therefore, an option that may appear to be equivalent can have different usage, value or markedness. In such cases, the translator is likely to opt for a shift, where a process of 'standardization' (Toury 1995: 274-279) or 'normalization' (Baker 1996) occurs.

Non-Subject Participant and Circumstance Themes also belong to this category. However, they are slightly different from the Subject and non-Subject Theme choice. The choice of Subject or non-Subject takes place both at the beginning of text nucleus and satellites. Non-Subject Participants, however, are unlikely to appear at the beginning of the text nucleus. On the other hand, Circumstances are realised as Theme at both locations. The difference is that certain types of Circumstances appear only at the beginning of the text nucleus and not satellite, and others appear only at the beginning of satellite but not the text nucleus. Earlier on, I established these specific types as text unit indicators. Since these types of Circumstances are strongly associated with particular text units, if shifts occur in them, they will take place at the genre level, giving consideration to locations in the text as well as linguistic, cultural, and other concerns.

Chapter 7

Conclusions and further research

The present study has investigated shifts in news translation from English into Japanese and from Japanese into English. The main aims were to discover patterns of translation shifts and factors influencing them. Two types of shifts were in focus; information content and information flow. In addition, the study explored non-translated news articles in Japanese to discover linguistic and textual features of them in contrast to the news written in English.

Chapter 1 outlined motivations for the study and Chapter 2 introduced Systemic Functional Linguistics (SFL) as the framework of the study. Chapter 3 provided the definition of translation shifts employed in this study, as well as discussing the conventional structure of news articles; text nucleus and a set of satellites. Chapter 4 described the research design, and provided a discussion of the corpus and analytical approaches. Chapters 5 and 6 presented the results of analysis and discussions of them. Chapter 5 focused on the general frequency of Theme realisation in translation, while Chapter 6 concerned the textuality of the translated news. Linking linguistic features with the conventional structure of text, Chapter 6 discussed additions, deletions and moves of information as well as the thematic development.

7.1 Main findings and contributions

7.1.1 Translation shifts in Theme selection (general frequency)

The first set of research questions concerned Theme selection in terms of frequency. The main questions were **‘what are the patterns of Theme realisation in translation?’** and **‘is the Theme realisation in translation ST oriented or TL oriented?’** In order to answer the former, based on which the latter question is answered, I will firstly address the sub-question to the former question, starting with En-trans and then continuing to Ja-trans.

- What elements are frequently realised as Theme in translation into English and into Japanese?

Themes can be simple or multiple, and the former were the majority (86%) in En-trans. When multiple Themes were realised, the textual Theme was frequently combined with an obligatory experiential Theme (87%), rather than the interpersonal Theme. Textual Themes were realised by Conjunctions more often (58%) than textual Adjuncts (42%).

Next, as for ideational Theme, the most common was the unmarked Subject Theme (75%). If a marked Theme was chosen, the most frequently realised was Adjunct Theme (60%). Among the different types of Adjunct, Location (36%) and Angle (28%) were the most frequently realised.

Turning to Ja-trans, between simple and multiple Theme, the former was realised more frequently (93%). When multiple, the normal multivariate multiple Theme was the majority (70%) but univariate multiple Themes were also realised 30% of the time. The latter multiple Theme, a combination of Circumstances of temporal and spatial Location, was a particular type of Circumstance in the register of news written in Japanese.

Multivariate multiple Themes were realised by a textual Theme (90%) combined with an experiential Theme. Textual Themes were realised by textual Adjuncts more frequently (73%) than Conjunctions (27%), which showed a clear difference from En-trans.

Between experiential and logical Theme, experiential ones constituted the majority (92%). As for the experiential, non-Subject Participant Themes accounted for the majority of the experiential Themes (66%), followed by Circumstances (33%). Among the Circumstance Themes, Location (39%) and Angle (34%) were realised frequently, which was the same in En-trans. When a clause was realised as a Logical Theme, which can be expansion or projection, expansion was more frequent (70%).

Based on the above findings, I shall now address the second main question, i.e. **'is Theme realisation in translation ST or TL oriented?'** by answering the two sub-questions below. Given that in these conclusions, I am seeking to bring the comparison together, I will answer them both together.

- How does Theme realisation in translation compare with that of the comparable corpus in either English or Japanese as the TL?
- How does Theme realisation in translation compare with that of the parallel corpus in either English or Japanese as the TL?

To discover patterns, I used a triangular comparison method and observed Theme realisation in translation with the comparable corpus of non-translated texts in the TL, as well as the parallel corpus of the STs. I will begin with En-trans, continuing to Ja-trans. With respect to En-trans, I found Theme realisation in En-trans displayed a mixture of 'standardization' (Toury 1995) or TL normalisation (Baker 1996) and ST 'influence' (Toury 1995) or ST shining through (Teich 2003: 209), as expected. Particularly significant as a

contribution to the field is my discovery of specifically where and how TL normalisation and ST influence occurred.

Strong TL normalisation was observed in En-trans in terms of the choices between Conjunction and textual Adjunct Theme, and also between unmarked Subject Theme and marked. This normalisation was also observed with respect to Complement Theme. On the other hand, strong ST influence was observed in terms of Circumstances of Location, Angle and Extent and these are briefly described below.

The first example of TL normalisation is the choice between Conjunctions and textual Adjuncts. Conjunctions were employed more frequently (58%) than textual Adjuncts (42%). In En-orig, Conjunctions were employed more frequently (87%) than textual Adjuncts, while in Ja-orig it was not Conjunctions (26%) that were frequent. En-trans showed a clear tendency of moving towards TL norms. This was partly due to the systemic difference between English and Japanese, in particular how Conjunctions and Mood are utilised towards the end of clauses in Japanese.

The second example of TL normalisation is the choice between the unmarked Subject Theme and marked Theme. The former was a common choice for Theme in En-trans (75%). This was the same with En-orig (84%), but different from Ja-orig (54%). En-trans tends to move closer to the TL. Both languages have the category of Subject and other comparable options but the level of markedness of the Subject is different in English and in Japanese in that the Subject is not a prominent unmarked Theme in Japanese unlike in English. Thus, TL normalisation occurred due to choice, rather than necessity.

The final example of TL normalisation concerns Complement Theme, which is another representative example of TL normalisation due to its different degree of markedness. A Complement Theme in English is highly marked (Thompson 2004: 146) and was realised in En-orig 2% of the time. In Japanese on the other hand, it is not as marked as in English since this often occurs due to Subject ellipsis (22% in Ja-orig).

There were also examples of ST influence, which was observed with Adjunct Themes, in particular Location (36%) and Angle (28%). In terms of Location, En-trans was closer to Ja-orig (40%) than to En-orig (55%), indicating a strong tendency of ST influence. As for Angle, En-trans employed it more frequently than it occurred in En-orig (8%) as well as in Ja-orig (17%). Similar to the choice of the unmarked and marked Theme, these Adjuncts are available in both languages. The difference is that the degree of markedness is not as

great as in the case of marked Theme. Comparable elements were readily accessible for Theme in translation, leading to significant ST influence.

In summary, TL normalisation in En-trans was observed with regard to the general choices of Theme, e.g. marked or unmarked Theme. On the other hand, with respect to more delicate choices of Theme, e.g. types of textual Theme or Adjunct, ST influence was observed in En-trans.

In contrast to En-trans, Theme realisation in Ja-trans showed a strong tendency of TL normalisation. Representative examples of TL normalisation in Ja-trans were found in the employing of univariate multiple Theme, choices between the Subject and non-Subject Participant, as well as in terms of logical Theme and Circumstances of Location, Matter and Angle.

As an example of TL normalisation, univariate multiple Themes are of particular interest here as they suggest a link between genre and translation practice. This particular multiple Theme, consisting of two Circumstances of temporal and spatial Locations, were realised in Ja-trans (30%), which was more frequent than in Ja-orig (18%). With a near absence of these in En-orig (0.4%), this was a clear example of TL normalisation.

Between the choice of Conjunctions and textual Adjuncts, the latter was more frequent in Ja-trans (73%), similar to Ja-orig (74%). It was different from En-orig (13%), which showed a clear preference for Conjunctions (87%). This is another case of a clear TL normalisation, again, partly due to the systemic difference in terms of how clauses are linked in the clause complex.

As the experiential Theme, Subject accounted for 59% in Ja-trans, which was closer to Ja-orig (50%) than to En-orig (90%), again evidencing TL normalisation. The reason for Ja-trans moving closer to TL norm was not systemic restriction but choice. This was due to the different level of markedness of Subject and Circumstance as Theme.

Among Circumstance Themes, frequently realised in Ja-trans were Location (39%) and Angle (34%). Ja-trans was not different from Ja-orig (40%) in terms of Location, but different from En-orig (55%). In terms of Angle, Ja-trans exceeded the proportion of it in Ja-orig (17%) as well as in En-orig (17%), constituting a very clear process of TL normalisation.

In summary, Ja-trans was clearly TL oriented, With respect to both general and delicate choices of Theme, Ja-trans is close to Ja-orig as a result of shifts.

7.1.2 Translation shifts in Theme selection (at generic stages)

The next set of research questions concerned the relationship between the Theme selection and the text structure. The main question was **'are there any patterns of Theme realisation in news translation at specific locations in text; at the beginning of the text nucleus and the beginning of each satellite?'** To discover patterns of shift, I took the same method of triangular comparison as in 7.1.1 and observed Theme realisation at the beginning of the nucleus and satellites.

There were sub-questions to the above main question;

- Are Themes realised at the beginning of a specific structure component distinguishable from those at the beginning of different types of structure component?

In the case of En-trans, the Subject Theme was more frequently realised at the onset of the text nucleus (94%) than of satellites (85%), which was potentially caused by the location in the text; the difference between the beginning of the text nucleus or that of satellites. However, it will be difficult to distinguish the Subject of the text nucleus from that of satellites because they do not differ from each other.

In contrast, with regard to Ja-trans, I have identified three types of Circumstances which could indicate for either the text nucleus or satellites. Firstly, univariate Theme (45%) and Place (31%) were prominently realised at the onset of the text nucleus. On the other hand, Angle (57%) was likely to indicate a satellite by its presence.

- What are the factors influencing translation shifts, if they occur?

There are three major factors contributing to shifts; locations in the text where Theme is realised, TL normalisation/standardization and ST influence. Answers to the following sub-question will provide examples.

- How does Theme realisation at the beginning of different structure components in translation compare with the comparable TL corpus in English or Japanese?

Theme realisation at the beginning of the text nucleus and satellites indicated translation shifts. In En-trans, both ST influence and TL normalisation was observed in Theme realisation. Location in the text also played a part in how theme was realised. At the onset of the text nucleus, the Subject was the preferred choice both in En-trans (95%) and in En-orig (85%). This could be interpreted either as TL normalisation or due to the location in which Theme

is realised. In contrast, at the onset of satellite, the Subject Theme in En-trans was realised less frequently (72%) than in En-orig (85%), which was a result of ST influence.

Ja-trans also displayed processes of TL normalisation and ST influence. Location in the text played a stronger part in Theme realisation compared to En-trans. An example of TL normalisation was observed at the onset of the text nucleus, where the Subject and non-Subject Theme were realised similarly as in Ja-orig (no significance). On the other hand, ST influence was observed at the onset of satellites. Ja-trans was found to be significantly different from Ja-orig in the way the Subject was realised as Theme (60%) compared to En-orig (55%), influenced by the ST. Realisations of Circumstance Themes were associated with specific locations in the text. With respect to the way univariate Themes were realised, Ja-trans showed a similarity with Ja-orig by employing them prominently at the onset of the text nucleus (45% & 31%). In contrast, Ja-trans showed a difference from Ja-orig in the way Angle was made use of at the onset of satellites (0% & 9%), which suggested the locations where Theme is realised were meaningful, and at the same time TL normalisation was at work. On this account, no difference was found between the datasets.

- How does Theme realisation at the beginning of different structure components in translation compare with the parallel ST corpus in English or Japanese?

A comparison of translation with the parallel ST corpus confirmed both TL normalisation and ST influence. In addition, it also revealed another factor contributing to adopt translation shifts, namely levels of markedness of Themes.

As for TL normalisation, it occurred both in En-trans and Ja-trans at the onset of the text nucleus in the choice of Subject or non-Subject. In En-trans, the Subject Theme accounted for 95%, while it accounted for 54% in Ja-orig. Given that the percentage for the said Theme at the same location in En-orig was 94%, Theme realisation in En-trans was a clear case of TL normalisation. Similarly in Ja-trans, the Subject Theme accounted for 60%, while 94% in En-orig. Ja-orig included 54% of the Subject Theme. Ja-trans too moved closer to TL norms.

An example of ST influence was observed in En-trans in terms of the choice of Subject or non-Subject Theme at the onset of satellites. Non-Subject Themes accounted for 29% in En-trans, while they accounted for 56% in Ja-

orig and 15% in En-orig. Although TL normalisation was also at work, En-trans was also shown to be influenced by Ja-orig.

Finally, an example of translation shift due to the different degree of markedness was observed in En-trans at the onset of satellites, in particular in terms of non-Subject Participant Theme (0.5% compared to 6% in Ja-orig). In addition, shifts caused by the same reason; the different degree of markedness were observed in Ja-trans too with respect to non-Subject and Circumstance Theme realisation.

7.1.3 Translation shifts in additions, deletions and moves

There are three main questions on translation shifts in terms of information content, the first being **'how frequently do additions, deletions or moves occur in news translation?'**, the second being **'are there any patterns of additions, deletions or moves?'** and the third **'do English translations and Japanese translations have similar patterns for additions, deletions or moves?'**

Nearly 100% of the texts underwent additions, deletions or moves, and very often a combination of all of them. En-trans employed additions to 92% of the articles, deletions (90%) and moves (85%), while Ja-trans employed additions in 90%, deletions (100%) and moves (89%). These findings based on an empirical study supported the general claim in literature that news translation is target-oriented and employs shifts as a strategy (Bielsa & Bassnett 2009). In terms of the translation direction, Ja-trans indicated slightly higher rates.

The second main question had three sub-questions, as below.

- At what semantic rank do they occur?

Additions, deletions or moves occurred at all the semantic ranks; text unit, sequence, figure and element. However, most frequently observed were at the rank of element, which is smaller than the clause. At the rank of element, in En-trans, additions occurred with 77% of the articles, deletions (81%) and moves (65%), while in Ja-trans the various frequencies were additions (90%), deletions (95%) and moves (74%). On the other hand, at the rank of text unit, in En-trans, additions occurred to 43% of the articles, deletions (22%) and moves (8%), while in Ja-trans the various frequencies were additions (10%), deletions (76%) and moves (21%). The occurrence percentages for the ranks of sequence and figure generally fell in between those at the element and the text unit.

- What kinds of information are added, deleted or moved?

Information is added to enable the translated texts to function as a complete whole in the new context. Three types were identified for added information; to fill in the knowledge gap, to create a better discourse and to follow genre conventions. Hence, explanations, clarifications, and contextualisation were frequently added. To create discourse, Conjunctions or textual Adjuncts were added. To conform to the genre conventions, clarification of the place where the reported news came from, equivalent monetary values, and suffixes for addressing people, were examples of added information.

The majority of the information deleted was that which functioned in the original to contextualise the event in news or present comments from people involved in the event.

Information was moved from the text nucleus towards satellites, or towards the text nucleus from satellites, depending on what type of information was suitable in these sections in each language. For example, detailed information about the suspect of a crime was moved from the text nucleus in the Japanese ST and located in a satellite in the English translation.

- Is there a link between the ranks and types of information that are added, deleted or moved?

Clearly perceived links were between element and information that were shifted for conventional reasons. 'US', '- yen' and 'suspect' were examples of these. The rank of element was also linked to expressions of discourse, i.e. textual Adjuncts. Other than the rank of element, the figure could be linked to projecting clauses, though not exclusively. The rank of sequence also related to projecting and projected clauses. Although the ranks could suggest to some extent whether the assistance a given shift aims to provide is local or more global, proposing a link between the particular ranks and types of information is difficult.

The final question on this issue is **'do English translations and Japanese translations have similar patterns for additions, deletions or moves?'**

Other than the general similarities, such as the translation shifts which tended to occur at the rank of element, there were two major differences. Firstly, more additions were found in En-trans (up to 91%) than Ja-trans (up to 83%), and more deletions in Ja-trans (62%-95% at different ranks) than in En-trans (22%-87%). I argue that this tendency reflects the needs of the target audience of the translated news. The target readers of Ja-trans, translations of news in Reuters and CNN, are Japanese, who are likely to be residing in Japan. On the other hand, the readers of En-trans, translation of news in

Yomiuri and *Mainichi*, are speakers of English who live in Japan (*Yomiuri Shimbun* 2016: 8). Japanese wishing to improve their English skills are also reading these articles but they are unlikely to be the main target readers (*ibid.* 8). The degree of geographical and cultural distance is likely to have influenced the choice of strategies.

Secondly, with respect to moves, in particular those across the boundary of the text nucleus and satellites, and at the rank of element, En-trans and Ja-trans displayed a reverse movement. There were more element moves towards the text nucleus in Ja-trans, while in En-trans away from the text nucleus. This clearly suggests that certain types of information are preferred in the text nucleus in each language.

7.1.4 Theme in Japanese news

The questions here were set to gain a better contrastive understanding of the discourse structure of news articles written in Japanese and how Theme contributes to it. The main question was ‘**Are there any patterns of Theme realisation at specific locations in Japanese news articles as non-translation at the beginning of different structure components?**’ Firstly, I will report the findings to answer the question below, which is the sub-question of the above main question. Then I move on to report the findings on two more questions in this set.

- Are Themes realised at the beginning of a particular structure component distinguishable from those at the beginning of a different type of structure component?

I have uncovered several Circumstance Themes that were realised specifically either at the beginning of the text nucleus or at the beginning of satellites. At the beginning of the text nucleus, univariate multiple Themes (31%) and Matter (33%) appear frequently but rarely at the beginning of satellites (2% and 9% respectively). They can be used to signal the start of the text nucleus. On the other hand, Angle is one that can signal the start of a satellite. These were clear examples of Theme that were frequently in use in the register of news in Japanese.

In addition to those Circumstances, the employing of 同定文 *dōteibun*, identifying sentence (18%) was also associated with the location of the text nucleus. This structure deploys the order of New^Given to mean ‘what has become known is the fact ...’ and appeared only at the beginning of the text nucleus, a conventional way of breaking news in Japanese.

Those described above show the relationship between Theme and genres (cf. Fries 1995) by appearing at specific locations in the text. I will now turn to the second main question; **‘how does Theme realisation in different structure components compare with that in the whole corpus of the original Japanese texts with no reference to realisation locations in the text?’**

There were several noticeable differences in the way Theme was realised in general, i.e. in the whole sub-corpus of Ja-orig, and how it was realised at the beginning of the text nucleus or satellites. One such difference emerged in the way non-Subject Participant Themes were realised. They hardly ever appeared at the beginning of the text nucleus (4%) while in the whole Ja-orig their realisation was relatively frequent (22%). Since non-Subject Participant Themes are likely to be a result of Subject ellipsis, they are unlikely to occur at the beginning of the text nucleus, which is the very beginning of the text.

Another difference was in relation to the choice of Circumstance Themes. Specific Circumstances were realised selectively depending on the location within the text. At the onset of the text nucleus, univariate Theme (31%) and Matter (33%) appeared overwhelmingly more frequently than they did in general in the whole sub-corpus of Ja-orig (4% and 14% respectively). On the other hand, at the onset of satellites, Angle (46%) was realised more prominently than the whole Ja-orig (17%). These differences were observed in terms of the Theme choices at ‘higher delicacy’ (Teich 2003).

The final question here is **‘How does Theme realisation at the beginning of different structure components compare with the original ST corpus in English?’**

There was a clear difference between En-orig and Ja-orig in terms of how frequently Subject Themes are realised. Both at the beginning of the text nucleus and satellites, the Subject was a preferred choice for Theme in En-orig (94% and 85% respectively). In comparison, Subject Themes in Ja-orig were not as prominent as in English, both at the onset of the text nucleus (56%) and satellites (67%), although in the latter location, they were employed more frequently. Taking the Subject Theme as an example, their realisational difference according to locations in text indicated that specific locations influenced the Theme choice, although it is difficult to eliminate the possibility that this was also a more general linguistic feature, rather than a generic one, especially with less delicate choices of Theme.

7.1.5 Contributions

The present study confirmed that news translation both into English and Japanese was target-oriented in terms of textuality. However, there was a difference in the degree of TL normalisation depending on the translation direction. Ja-trans showed a stronger tendency to move closer to TL norms than En-trans did. Nonetheless, translation shifts both in terms of information content and information flow were motivated by the target reader's expectations of news articles written in the target language. In addition, it is possible that the translators' first language, which is probably Japanese, plays a role in this phenomenon, although this is a conjecture.

In addition to these particular findings, there are general contributions from the way the current study approached the topic. Firstly, this study is the first of its kind investigating shifts in news translation between English and Japanese using a reasonably sized corpus. The corpus I constructed in this study in itself constitutes a significant contribution to the field. It is representative of the specific register of news articles and is of an appropriate size for the purpose, consisting of four sub-corpora, each of which includes 85 news articles. There are 340 texts totalling 131,975 words in English and 180,738 characters in Japanese. Moreover, all the texts in the corpus are complete articles, paired with their translations and aligned.

One of the foci of the present study was Theme. Thomson (2001) explores Theme in Japanese broadly, making use of texts from several registers. Her case study centres around the general system of THEME, while I focus on the register of news articles. With respect to translation, there has been little research on Theme with the language pair of English and Japanese, except for Naganuma (2000). She points out thematic challenges translators may encounter based on the systemic differences between the two languages. Focusing on news translation, I investigated shifts quantitatively with respect to Theme.

Another contribution of this study is that it investigated the link between Theme and text structure in translation, which has not been done quantitatively before with respect to news translation between English and Japanese. Research on the text structure of news articles in Japanese has indeed been previously conducted (Thomson 2001, 2005b; Abe 2004; Washitake 2009); Washitake focuses on the structure itself while Abe and Thomson connect the structure and Theme. However, Abe (2004) is a small scale qualitative study and Thomson focuses on the different lexicogrammatical configurations of textual metafunction including Theme in texts in different genres. Hence, the present

study of mine contributes to addressing the topic of how Theme relates to the generic text structure of news translation quantitatively.

In terms of the approach to analysis, adopting only explicit elements as Theme instead of allowing implicit Theme was uncommon in previous studies of Japanese (cf. Thomson 2001; Teruya 2006) and proved beneficial. This approach made it possible to observe the frequency of the Subject Theme (78% in Ja-orig and 89% in Ja-trans) and non-Subject Themes (22% and 11% respectively) as well as how often non-Subject Themes were topicalised (21% in Ja-orig and 24% in Ja-trans) (Chapter 5).

As for practical contributions, the findings of this study are usefully applied to translator training, in particular, journalistic translation from English into Japanese, which is the translation direction and type I teach. The present study originated from the problem of the trainee translators' uncertainty regarding changes during translation. The findings of the present study will provide them with evidence of shifts that occur in professional news translation. These findings will enable trainees to analyse the ST and decide whether or not to employ a shift. For example, they will have examples of the type of information which is added or deleted, and how much is added or deleted. Also useful for them to know will be what kind of text nucleus is preferred in Japanese which will allow them to move information across the boundary of the text nucleus and satellites during translation from English, as well as knowing conventional ways of starting a text unit; univariate Themes at the beginning of the text nucleus and Circumstance of Angle at the beginning of satellites.

In short, incorporating the findings of this study into teaching will enable trainee translators to learn the professional practice in news translation, which they can then apply in their own practice. News translation is target-oriented (Bielsa & Bassnett 2009), which means recreating the meanings of the ST to fit into a new context (Matthiessen 2001) while balancing between ST and TT (Steiner 2001).

In addition to translator training, the findings can be applied to Japanese language teaching, in particular at the advanced level. News articles are often employed as materials in reading comprehension classes and understanding their linguistic and structural features facilitates students' comprehension. Being aware of the macro structure of typical hard news, i.e. a text nucleus and a set of satellites, as well as the Theme-Rheme structure within the T-unit and how information is organised, will support them in reading news in Japanese more effectively (cf. Martin & Rose 2008).

Applying the findings to translator training and language teaching is similar in terms of developing necessary skills. However, in producing texts in the TL, learners of Japanese simply need to be aware of the differences between the two languages while in translator training trainees need to balance between the ST and TT (Steiner 2001).

7.2 Limitations

The results of this study are encouraging, although inevitably there are some limitations and these are outlined below.

7.2.1 Corpus

Although the corpus I constructed in this project is a contribution in itself, it included eleven articles out of 170 texts in En-orig that do not have the orbital structure of nucleus and satellites. They may have affected the generalisability of results.

Firstly, the lack of the orbital structure meant that it was not possible to observe the Theme realisation at the beginning of these text units. It was also not possible to examine additions, deletions or moves in these articles at the rank of the text unit. I decided to separate them from the rest in the discussion of textuality in Chapter 6 to avoid the risk of skewing the results by including them.

Secondly, this lack of orbital structure may have affected the frequency of the occurrence of interpersonal Themes (Chapter 5). Although the analysis results indicated that the interpersonal Theme appeared both within and outside direct quotations, these atypical hard news articles may have included more direct quotations that included modal expressions appearing in Theme.

In addition to those in En-orig, there were also 16 articles out of 170 texts in Ja-orig that include extra structure units after the writer's name, i.e. after the news articles were complete. They conformed to the conventional structure of a text nucleus and a set of satellites within the article itself. This did not present a problem in terms of the analysis of thematic development in the text, as long as the extra section was excluded from analysis. However, information presented in those excluded sections in the ST was included in the TT. Since the notion of text units was essential in Chapter 6 to identify translation shifts, these were also excluded from analysis in Chapter 6.

Despite these limitations, those texts, in particular those found in En-orig, presented valuable findings. Firstly, they were evidence of a variety of news

articles that are translated into Japanese. Secondly, they indicated a clear shift of structure when translated. The ST without the nucleus and satellite structure was translated employing the conventional hard news structure, i.e. with a nucleus and satellites. This suggested that the translated news articles conformed to the culturally expected way a news article is structured in Japanese.

7.2.2 Methodology

7.2.2.1 Implicit Theme

This study did not allow the implicit Theme and adopted explicit elements as Theme, which was of particular relevance to Japanese (Chapter 4). The decision proved beneficial to confidently choosing Theme at the stage of annotation, in particular, between an implicit Subject and an explicit Circumstance. It also enabled an observation of the frequency of the Subject Theme and non-Subject Themes, as well as how often non-Subject Themes were topicalised, as described in 7.1.5.

However, there was one type of explicit Theme which posed a problem in terms of its functionality as Theme, namely the verbal Predicate Theme. A verbal Predicate Theme in Japanese is the beginning part of the Process and is practically one word. For consistency of analysis, it was the right decision. From the theoretical perspective, however, it felt artificial to divide a word into two parts, assigning one part the point of departure, while making the remaining part Rheme. This was one example of a case in which implicit Theme may work better.

7.2.2.2 Consistency of analysis

To explore Theme realisation at specific locations in text, the text nucleus and satellites were identified as an intermediate semantic unit between the figure and the whole text (Chapter 3 and Chapter 4 Section 3.4). Although adopting a functionally defined text unit was meaningful, deciding the boundary between satellites proved difficult at times, for two possibly functionally discrete satellites were often connected by reference chains.

This may have affected the consistency of data annotation since I was the sole coder. An intercoder agreement would have been useful to ensure the consistency. Moreover, in terms of the applicability of the findings to translator training, this may hinder accessibility. To enhance the applicability to translator training, employing the orthographical paragraphs as the unit of text structure would be an option. Although these are not semantic units, doing this will enable trainees to conduct text analysis easily.

7.2.2.3 Scope

This study explored Theme realisation in translation in terms of frequency (Chapter 5) and in relation to text structure (Chapter 6). With respect to the latter, the investigation was limited to the beginning of the text unit and the examination of Theme realisation within the text unit was not conducted. This provides scope for further study.

To gain an overall view of textuality of the news article, it will be useful to investigate how Theme is realised and developed within the text unit. This will potentially shed light on how non-Subject Theme is employed in Japanese, i.e. how Subject ellipsis occurs, which did not frequently appear at the beginning of the text unit.

7.3 Further research

This study has demonstrated evidences of translation shifts that take place in news translation between English and Japanese in terms of the textual metafunction. It also addressed the experiential metafunction in relation to the readability of translation in terms of information content with regards to additions, deletions or moves of information. However, the discussion of the interpersonal metafunction is left for future study, mainly because its relevance to Theme is limited in Japanese in that the major interpersonal functions are realised at the end of the clause (Teruya 2006: 135).

The investigations of interpersonal meaning will shed useful light on shifts in news translation. I am particularly interested in the use of an expression of evidentiality, an aspect of the interpersonal metafunction, in relation to projections. Direct quotations were translated from English into Japanese in three noticeable ways; as direct quotations, as indirect quotations and with the expression of hearsay '*to iu* [it is said]'. Considering that the projecting verbal clause is a 'favourite clause type in a news report' (Halliday & Matthiessen 2004: 586), this is an area where the present study could be further developed to facilitate understanding of news translation from English into Japanese.

Translation of direct quotations in news has been studied (Schäffner 2008; Matsushita 2013) but not yet extensively. Within SFL of Japanese, Ochi (2008) reports how news reports employ projections to present the news as objective. Research on projections in news translation between Japanese and English using the framework of SFL is yet to be conducted.

In terms of gaining a better understanding of translation shifts with respect to textuality, i.e. readability and acceptability, the approach adopted in the

current study could be applied to analysis of different text types. There is limited research on non-literary translation in Japanese (Chapter 3) and, therefore, analysing other types of texts is expected to shed light on possible translation shifts that occur in different registers.

On the other hand, further study could be in the area of translator training. Exploring trainee translators' translations in terms of textuality in comparison to the professionals' to identify gaps could potentially be applied both to teaching and to syllabus design. Kim (2007) claims that explicit explanation of textual meaning supports translator training. Identifying particular issues that trainee translators working with Japanese and English face will be valuable.

Another potential direction for further study is to investigate translation shifts from the point of the reader, i.e. how news translations are perceived by the audience. The present study has explored readability and acceptability of news articles by focusing on the end product, which has been produced for the reader based on the translator's perception of the needs of the audience. An investigation of the reader's point of view will be valuable.

In addition to Translation Studies, teaching Japanese as a foreign language is another area where further research can be conducted. Correct choice of lexis and correct syntax do not always guarantee a text with a good discourse flow, as has been noted in translation (Baker 1992). The same applies to compositions that the learners of Japanese produce. It will be valuable to explore whether explicit instruction enables students to communicate their ideas in writing in a way Japanese readers find comfortable, by following how information is organised in the T-unit in Japanese.

List of References

- Abe, S. 2004. Nihongo no janru-kōzō to goi-bunpō-teki shigen [Generic structure in Japanese and lexico-grammar resources]. *Gendai-shakai Bunka Kenkyū*. **30**, pp.179-195.
- Alekseyenko, N. 2013. *A Corpus-based study of Theme and Thematic Progression in English and Russian non-translated texts and in Russian translated texts*. [Online]. PhD Thesis, Kent State University. [Accessed 1 May 2018]. Available from:
https://etd.ohiolink.edu/rws_etd/document/get/kent1373497827/inli
- Aktan, O. & Nohl, A-M. 2010. International trans-editing: Typical intercultural communication strategies at the BBC World Service Turkish Radio. *Journal of Intercultural Communication*. **24**, [No pagination].
- Anzai, T., Inoue, K. & Kobayashi, A. 2005. *Hon'yaku o Manabu Hito no-tameni [For Those who Study Translation]*. Kyoto: Sekaishisō-sha.
- Aruz, J. 2010. On Theme in English and Spanish: A comparative study. In: Swain, E. ed. *Thresholds and Potentialities of Systemic Functional Linguistics: Multilingual, multimodal and other specialised discourses*. [Online]. Trieste: EUT Edizioni Università di Trieste, pp.23-48. [Accessed 11 Nov 2015]. Available from:
<http://hdl.handle.net/10077/3649>
- Baker, M. 1992. *In Other Word*. London: Routledge.
- Baker, M. 1996. Corpus-based translation studies: The challenges that lie ahead. In Somers, H. ed. *LSP and Translation: Studies in language engineering in honour of Juan C. Sager*. Amsterdam: John Benjamins Publishing Company, pp.175-186.
- Baker, M. & Saldanha, G. eds. 2009. *Routledge Encyclopaedia of Translation Studies*. 2nd ed. London & New York: Routledge.
- Bakker, M., Koster, C. & Van Leuven-Zwart, K. 2009. Shifts. In: Baker, M. & Saldanha, G. eds. *Routledge Encyclopaedia of Translation Studies*. 2nd ed. London & New York: Routledge, pp.269-274.
- Bartlett, T. & O'Grady, G. eds. 2017. *The Routledge Handbook of Systemic Functional Linguistics*. London & New York: Routledge.

- Bateman, J., Delin, J. & Henschel, R. 2007. Mapping the multimodal genres of traditional and electronic newspapers. In: Royce, T. & Bowcher, W. eds. *New Directions in the Analysis of Multimodal Discourse*. [Online]. New Jersey: Lawrence Erlbaum Associates, pp.147-172. [Accessed 5 August 2012]. Available from: <http://www.fb10.uni-bremen.de/anglistik/ling/ws06/gem-materials/mapping-multimodal-genres-neat.PDF>
- Bednarek, M. 2006. *Evaluation in Media Discourse: Analysis of a newspaper corpus*. London & New York: Continuum.
- Bednarek, M. 2009. Corpora and discourse: A three-pronged approach to analyzing linguistic data. *HCSNet Summerfest '08*, Somerville MA USA: Cascadilla Proceedings Project.
- Bell, A. 1991. *The Language of News Media*. Oxford: Blackwell.
- Bell, A. 1998. The discourse structure of news stories. In: Bell, A. & Garrett, P. eds. *Approaches to Media Discourse*. Oxford & Massachusetts: Blackwell Publishers Inc, pp.64-104.
- Berry, M. 1995. Thematic options and success in writing. In: Ghadessy, M. ed. *Thematic Development in English Texts*. London & New York: Pinter Publishers, pp.55-84.
- Bhatia, V.K. 1993. *Analysing Genre: Language use in professional settings*. London & New York: Longman.
- Biber, D. 1990. Methodological issues regarding corpus-based analyses of linguistic variation. *Literary and Linguistic Computing*. 5(4), pp.257-269.
- Biber, D., Connor, U. & Upton, T. 2007. *Discourse on the Move: Using corpus analysis to describe discourse structure*. Amsterdam & Philadelphia: John Benjamins Publishing Company.
- Biber, D. & Conrad, S. 2009. *Register, Genre and Style*. Cambridge: Cambridge University Press.
- Bielsa, E. 2010. The sociology of translation: Outline of an emerging field. *MonTI*. 2, pp.153-172.
- Bielsa, E. & Bassnett, A. 2009. *Translation in Global News*. London & New York: Routledge.
- Blum-Kulka, S. 1986/2000. Shifts of cohesion and coherence in translation. In: Venuti, L. ed. *The Translation Studies Reader*. London & New York: Routledge, pp.298-313.

- Caple, H. 2009. *Playing with words and pictures: Intersemiosis in a new genre of news reportage*. [Online]. PhD Thesis, University of Sydney. [Accessed 3 May 2013]. Available from: <http://hdl.handle.net/2123/7024>
- Catford, J. 1965. *A Linguistic Theory of Translation: An essay in applied linguistics*. Oxford: Oxford University Press.
- Cyrus, L. 2006. Source building a resource for studying translation shifts. In *Proceedings of the Fifth International Conference on Linguistic Resources and Evaluation LREC, 22-28 May 2006, Genoa*. Genoa: European Language Resources Association, pp.1240-1245.
- Cyrus, L. 2009. Old concepts, new ideas: Approaches to translation shifts. *MonTI*. 1, pp.87-106.
- Daneš, F. 1974. Functional sentence perspective and the organization of the text. In: Daneš, F. ed. *Papers on Functional Sentence Perspective, Janua Linguarum series minor 147*. Prague: Academia, The Hague & Paris: Mouton, pp.106-128.
- Downing, A. 1991. An alternative approach to theme: A systemic-functional perspective. *Word*. 42(2), pp.119-143.
- Evans, H. 2000. *Essential English: For journalists, editors and writers*. Revised ed. London: Pimlico.
- Fries, P. 1994. On Theme, Rheme and discourse goals. In: Coulthard, M. ed. *Advances in Written Text Analysis*. London: Routledge, pp.229-249.
- Fries, P. 1995a. Themes, methods of development and texts. In: Fries, P. & Hasan, R. eds. *On Subject and Theme*. Amsterdam & Philadelphia: John Benjamins, pp.317-360.
- Fries, P. 1995b. A personal view of Theme. In: Ghadessy, M, ed. *Thematic Development in English Texts*. London & New York: Pinter Publishers, pp.1-19.
- Halliday, M.A.K. 1961/2002. Categories of the theory of grammar. In: Webster, J. ed. *On Grammar. Collected Works of M.A.K. Halliday*. Vol. 1. London & New York: Continuum, pp.37-94.
- Halliday, M.A.K. 1967. Notes on transitivity and theme in English: Part 2. *Journal of Linguistics*. 3(02), pp.199-244.
- Halliday, M.A.K. 1977/2003. Text as semantic choice in social contexts. In: Webster, J. ed. *Linguistic Studies of Text and Discourse. Collected Works of M.A.K. Halliday*. Vol. 2. London: Continuum. pp.23-81.

- Halliday, M.A.K. 1978. *Language as Social Semiotic: The social interpretation of language and meaning*. London: Edward Arnold.
- Halliday, M.A.K. 1981/2002. Text semantics and clause grammar: How is a text like a clause? In: Webster, J. ed. *On Grammar. Collected Works of M.A.K. Halliday*. Vol. 1. London & New York: Continuum, pp.219-260.
- Halliday, M.A.K. 1985. *An Introduction to Functional Grammar*. London: Edward Arnold.
- Halliday, M.A.K. 1992. Language as system and language as instance: The corpus as a theoretical construct. In: Svartvik, J. ed. *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82, 4-8 August 1991, Stockholm*. Berlin & New York: Mouton de Gruyter, pp.61-77.
- Halliday, M.A.K. 2001. Towards a theory of good translation. In: Steiner, E. & Yallop, C. eds. *Exploring Translation and Multilingual Text Production: Beyond content*. Berlin & New York: Mouton de Gruyter, pp.13-18.
- Halliday, M.A.K. 2009. Methods – techniques – problems. In: Halliday, M.A.K. & Webster, J. eds. *Continuum Companion to Systemic Functional Linguistics*. London & New York: Continuum International Publishing Group, pp.59-86.
- Halliday, M.A.K. & Hasan, R. 1976. *Cohesion in English*. London: Longman.
- Halliday, M.A.K. & Hasan, R. 1989. *Language, Context, and Text: Aspects of language in the social semiotic perspective*. 2nd ed. Oxford: Oxford University Press.
- Halliday, M.A.K. & Matthiessen, C.M.I.M. 2004. *An Introduction to Functional Grammar*. 3rd ed. London: Hodder Education.
- Hasan, R. 1977. Text in the systemic-functional model. In: Dressler, W. ed. *Current Trends in Textlinguistics*. Berlin: Walter de Gruyter, pp.228-246.
- Hasan, R. & Fries, P. eds. 1995. *On Subject and Theme: A discourse functional perspective*. Amsterdam: John Benjamins Publishing Company.
- Hatim, B. & Mason, I. 1990. *Discourse and the Translator*. London: Longman.
- Hatim, B. & Mason, I. 1996. *The Translator as Communicator*. London & New York: Routledge.

- House, J. 1997. *Translation Quality Assessment*. Tübingen: Gunter Narr Verlag.
- Hayashi, M. 2007. Dōshi no te-kei to ren'yō-kei no shiyō-sa ni-kansuru keiryō-teki chōsa-kenkyū. [Quantitative study on relative usage of te-forms and infinitive forms of verbs]. *Ryūkoku Daigaku Kokusai-sentā Kenkyū Nempō*. **16**, pp.49-58.
- Hunston, S. 2002. *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Iedema, R., Feez, S. & White, P. 1994a. *Stage one: Media literacy. A report for the Write it Right Literary in Industry Research Project by the Disadvantaged Schools Program, N.S.W.* [Online]. New South Wales, Australia: Department of School Education. [Accessed 3 May 2013]. Available from: <http://www.grammatics.com/appraisal/MediaLit-Comment.pdf>
- Iedema, R., Feez, S. & White, P. 1994b. *Stage two: Media literacy. A report for the Write it Right Literary in Industry Research Project by the Disadvantaged Schools Program, N.S.W.* [Online]. New South Wales, Australia: Department of School Education. [Accessed 3 May 2013]. Available from: <http://www.grammatics.com/appraisal/MediaLit-Appraise.pdf>
- Ikegami, Y. 1981. *Suru to Naru no Gengogaku [The Linguistics of 'Do' and 'Become']*. Tokyo: Taishūkan-shoten.
- Ikegami, Y. 1996. *Eigo no Imi [The Meaning in English]*. Tokyo: Taishūkan-shoten.
- Jakobson, R. 1959. On linguistic aspects of translation. In: Achilles, F. & Brower, R. eds. *On Translation* 23. Boston: Harvard University Press, pp.232-239.
- Károly, K. 2012. News Discourse in Translation: Topical structure and news content in the analytical news article. *Meta*. **57**(4), pp.884–908.
- Katsuta, J. 2001. Nichiei shimbunkiji ni-okeru dai-1 danraku no kinō [the functions of the first paragraph in the newspaper articles in Japanese and English]. *Koide Kinen Nihongokyōiku Kenkyūkai Ronbunshū*. **9**, pp.59-71.
- Kawahara, K. 2009a. Analysis of translation shifts in the act of translating and interpreting between English and Japanese from the perspective of

cognitive linguistic typology. *Invitation to Interpreting & Translation Studies*. **3**, pp.29-49.

Kawahara, K. 2009b. Improvements in translation products through the teaching of translation. *Bunka Gakuin University Journal*. **17**, pp.91-105.

Kawahara, K. 2011a. Hon'yaku to wa nanika: Kenkyu to shite no hon'yaku no.5 [What is translation?: Translation as study no.5]. *Hon'yaku tsūshin*. Vol.104 [*Translation newsletter*. Vol.104]. [Online]. [Accessed 3 May 2013]. Available from: <http://www.honyaku-tsushin.net/bn/201101.pdf>

Kawahara, K. 2011b. Gaisetsusho ni miru honyakugaku no kihon ronten to zentaiteki taikai. [Basic issues and overall grasp of the whole system of translation studies explained in introductory translation studies books]. *Invitation to Interpreting & Translation Studies*. **5**, pp.53-80.

Kawahara, K. 2011c. Hon'yaku to wa nanika: Kenkyu to shite no hon'yaku no.7 [What is translation?: Translation as study no.7]. *Hon'yaku tsūshin*. Vol.106 [*Translation newsletter*. Vol.106]. [Accessed 11 May 2011]. Available from <http://www.honyaku-tsushin.net/bn/201103.pdf>

Kawahara, K. 2012. An attempt at a new approach to translation shift. *Reitaku University Academic Repository*. **20**(1), pp. 69-78.

Kenny, D. 2009. Equivalence. In: Baker, M. & Saldanha, G. eds. *Routledge Encyclopaedia of Translation Studies*. 2nd ed. London & New York: Routledge, pp.96-99.

Kim, M. 2006. The system of Theme of Korean: An interim report. In *Proceedings of 33rd International Systemic Functional Congress, 10-12 July 2006, São Paulo*. [Online]. pp.706-721. [Accessed 6 Jun 2015]. Available from: <http://www.pucsp.br/isfc>

Kim, M. 2007. Using systemic functional text analysis for translator education: An illustration with a focus on textual meaning. *The Interpreter and Translator Trainer*. **1**(2), pp.223-246.

Kim, M. 2011. A systemic functional approach to the tangled thread issues of Korean Theme study. *International Review of Korean Studies*. **8**(1), pp.101-138.

Kim, M. 2016. SFL, an empowering tool for translation studies. *The Journal of Translation Studies*. **17**(4), pp.5-10.

Kim, M. & Huang, Z. 2012. Theme choices in translation and target readers' reactions to different Theme choices. *T&I Review*. **2**, pp.79-111.

- Kim, M., Matthiessen, C.M.I.M. & Burns, A. 2009. Doctoral work in translation studies as an interdisciplinary mutual learning process. *The Interpreter and Translator Trainer*. **3**(1), pp.107-128.
- Kim, M. & Matthiessen, C.M.I.M. 2015. Ways to move forward in translation studies: A textual perspective. *Target*. **27**(3), pp.335-350.
- Knox, J. S. 2009. *Multimodal discourse on online newspaper home pages: A social-semiotic perspective*. [Online]. PhD Thesis, University of Sydney. [Accessed 6 Jun 2013]. Available from: <http://hdl.handle.net/2123/7696>
- Kokuritsu-kokugo-kenkyūjo. 1981. *Nihongo no Bunpō 2 [Grammar of Japanese 2]*. Tokyo: Kokuritsu-kokugo-kenkyūjo.
- Lavid, J., Arus, J. & Moraton, L. 2012. Genre realized in Theme: The case of news reports and commentaries. [Online]. *Discourse*. **10**, [no pagination].
- Lee, C. 2006. Differences in news translation between broadcasting and newspapers: A case study of Korean-English translation. *Meta*. **51**(2), pp.317–327.
- Li, C. & Thompson, S. 1976. Subject and topic: A new typology of language. In: Li, C. ed. *Subject and Topic*. New York: Academic Press, pp.457-490.
- McCabe, A. 1999. *Theme and Thematic Patterns in Spanish and English History Texts*. [Online]. PhD Thesis, Aston University. [Accessed 1 December 2015]. Available from: <http://www.isfla.org/Systemics/Print/Theses/McCabephd.pdf>
- Manfredi, M. 2008. Translating text and context. [Online]. Bologna: Centro di Studi Linguistico-Culturali (CeSLiC). [Accessed 01 April 2016]. Available from: https://internt.ht.lu.se/doc/1384248522.calendarEvents.5481.pdf.0.Manfredi_2008_Monografia.pdf/Manfredi_2008_Monografia.pdf
- Mann, W. & Thompson, S.A. 1988. Rhetorical structure theory: A theory of text organization. *Text*. **8**(3), pp.243-281.
- Martin, J.R. 1992. *English Text: System and structure*. Philadelphia & Amsterdam: John Benjamins Publishing Company.
- Martin, J. 1995. More than what the message is about: English Theme. In: Ghadessy, M. ed. *Thematic Development in English Texts*. London & New York: Pinter Publishers, pp.223-258.

- Martin, J.R. 2009. Genre and language learning: A social semiotic perspective. *Linguistics and Education*. **20**(1), pp.10-21.
- Martin, J.R. 2014. Evolving systemic functional linguistics: Beyond the clause. *Functional Linguistics*. **1**(3), pp.1-24.
- Martin, J. & Rose, D. 2008. *Genre Relations: Mapping culture*. London & Oakville: Equinox.
- Martin, S. 1975. *A Reference Grammar of Japanese*. New Haven: Yale University Press.
- Masuoka, T. 1997. *Fukubun [Compound Sentences]*. Tokyo: Kuroshio-shuppan.
- Masuoka, T. 2012. Nihongo no chūritsukei-setsuzoku to te-kei setsuzoku no kyōgō to kyōson [The competition and coexistence of the neutral form connection and the te-form connection]. [Online]. [leaflet]. Kyoto: Doshisha University. [Accessed 29 Feb 2016]. Available from: https://www1.doshisha.ac.jp/~cjitl210/data1/31_masuoka.pdf
- Matsushita, K. 2013. Obama-daitōryō no enzetsu o nihon no shimbun wa ikani hōjitaka: Chokusetsu-in'yō to shōryaku o-meguru kōsatsu [How Japanese newspapers reported President Obama's speech: Direct quotations and omissions]. *Media, English and Communication*. **3**, pp.31-46.
- Matthiessen, C.M.I.M. 1990. Two approaches to semantic interfaces in text generation. In: Karlgren, H. ed. *COLING '90 Proceedings of the 13th Conference on Computational Linguistics, 20-25 August 1990, Helsinki*, [Online]. Stroudsburg, PA, USA: Association for Computational Linguistics, pp.322-329. [Accessed 29 June 2018]. Available from: <http://www.aclweb.org/anthology/C90-2056>
- Matthiessen, C.M.I.M. 1995a. *Lexicogrammatical Cartography: English systems*. Tokyo: International Language Sciences Publishers.
- Matthiessen, C.M.I.M. 1995b. Theme as an enabling resource in ideational 'knowledge' construction. In: Ghadessy, M. ed. *Thematic Development in English Texts*. London & New York: Pinter Publishers, pp.20-54.
- Matthiessen, C.M.I.M. 2001. The environments of translation. In: Steiner, E. & Yallop, C. eds. *Exploring Translation and Multilingual Text Production: Beyond content*. Berlin & New York: Mouton de Gruyter, pp.41-124.

- Matthiessen, C.M.I.M. 2014. Choice in translation: Metafunctional considerations. In: Kuntz, K., Teich, E., Hansen-Schirra, S., Neumann, S. & Daut, P. eds. *Caught in the Middle – language use and translation*. Saarland: Saarland University Press, pp.271-334.
- Matthiessen, C.M.I.M., Lam, M. & Teruya, K. 2010. *Key Terms in Systemic Functional Linguistics*. London & New York: Continuum.
- McEnery, T., Xiao, R. & Tono, Y. 2006. *Corpus-based Language Studies: An advanced resource book*. London: Routledge.
- Moore, N. 2010. *Structuring Information in Written English: A re-appraisal of the systemic functional approach to information structure*. PhD Thesis, University of Liverpool.
- Morishige, S. 1973. Nihonbunpō no rekishi [The history of Japanese grammar], In *Bunkachō, Nihongo to Nihongokyōiku [Japanese Language and Japanese Language Teaching]*. Tokyo: Ōkurashō insatsukyoku, pp.223-240.
- Morita, Y. 1989. *Kiso Nihongo Jiten [Basic Japanese Dictionary]*. Tokyo: Kadokawa-shoten.
- Morita, Y. & Matsuki, M. 1998. *Nihongo Hyogen Bunkei [Japanese Expression Patterns]*. Tokyo: Alc.
- Munday, J. 1998a. A computer-assisted approach to the analysis of translation shifts. *Meta*. **43**(4), pp.542-556.
- Munday, J. 1998b. Problems of thematic analysis between Spanish and English. *Cadernos de Tradução*. **1**(3), pp.183-213.
- Munday, J. 2000. Using systemic functional linguistics as an aid to translation between Spanish and English: Maintaining the thematic development of the ST. *Revista Canaria de Estudios Ingleses*. **40**, pp.37-58.
- Munday, J. 2008. *Introducing Translation Studies*. London & New York: Routledge.
- Naganuma, M. 2000. Nichiei hon'yaku ni-okeru Theme ni-kansuru kadai [Thematic challenges in translation between Japanese and English]. [Online]. [Accessed 02 Aug 2012]. Available from: http://researchmap.jp/?action=cv_download_main&upload_id=12038
- Naganuma, M. 2008. A systemic functional approach to translation pedagogy in Japan. In: Nørgaard, N. ed. *Systemic Functional Linguistics*

- in Use: Odense working papers in language and communication*. 29. [Online]. Odense: University of Southern Denmark, pp.589-600. [Accessed 15 May 2018]. Available from https://www.sdu.dk/~media/Files/Om_SDU/Institutter/ISK/Forskningspublikationer/OWPLC/Nr29/Contents.ashx
- Negishi, Y. 1999. *Handbook of Japanese-English Translation: From the scene of news translation*. Tokyo: Taishūkan-shoten.
- Neumann, S. & Hansen-Schirra, S. 2005. The CroCo Project: Cross-linguistic corpora for the investigation of explicitation in translations. In *Proceedings from the Corpus Linguistics Conference Series*. 1(1), 14-17 July 2005, Birmingham, Birmingham: University of Birmingham, [no pagination]. [Accessed 1 July 2018]. Available from: <https://www.semanticscholar.org/paper/The-CroCo-Project-Cross-linguistic-corpora-for-the-Hansen-Schirra/bbc66f055f339539553a8f7595faa910ec47a2f9>
- Nitta, Y. 1993. *Nihongo no Kaku o Megutte [Concerning Cases in Japanese]*. Tokyo: Kuroshio-shuppan.
- Noda, H. 1996. *Wa to Ga [wa and ga]*. Tokyo: Kuroshio-shuppan.
- Ochi, A. 2008 Ideational projection in Japanese news reporting: Constructing reporter's experience above clause rank. In: Nørgaard, N. ed. *Odense Working Papers in Language and Communication*. 29. [Online]. Odense: University of Southern Denmark, pp.601-620. [Accessed 10 October 2018]. Available from: https://www.sdu.dk/~media/Files/Om_SDU/Institutter/ISK/Forskningspublikationer/OWPLC/Nr29/Contents.ashx
- O'Donnell, M. 2010. *UAM CorpusTool (2.8)*. [Software]. [Accessed 14 September 2012].
- Olohan, M. & Baker, M. 2000. Reporting that in translated English: Evidence for subconscious processes of explicitation?. *Across Languages and Cultures*. 1(2), pp.141-158.
- Pöttker, H. 2003. News and its communicative quality: The inverted pyramid – when and why did it appear?. *Journalism Studies*. 4(4), pp.501-511.
- Pym, A. 2007. Natural and directional equivalence in theories of translation. *Target*. 19, pp.271-294.
- Pym, A. 2016. *Translation Solutions for Many Languages*. London & New York: Bloomsbury.

- Rose, D. 2001. Some variations in Theme across languages. *Functions of Language*. **8**(1), pp.109-145.
- Sawanishi, T. 2003. Dōshi ren'yo-kei no seishitsu [The nature of the verbal ren'yo form]. *Nihongo, Nihon-bunka [Japanese Language and Japanese Culture]*. **29**, pp.47-66.
- Schäffner, C. 2008. The prime minister said ...: Voices in translated political texts. *SYNAPS*. **22**, pp.3-25.
- Scollon, R. 2000. Generic variability in news stories in Chinese and English: A contrastive discourse study of five days' newspapers. *Journal of Pragmatics*. **32**, pp.761-791.
- Scollon, R. & Scollon, S.W. 1995. *Intercultural Communication: A discourse approach*. Oxford UK, Cambridge USA: Blackwell.
- Sharoff, S. 2017. Corpus and systemic functional linguistics. In: Bartlett, T. & O'Grady, G. eds. *The Routledge Handbook of Systemic Functional Linguistics*. Milton Park & New York: Taylor and Francis, pp.533-546.
- Sharoff, S. 2018. Functional text dimensions for the annotation of Web corpora. *Corpora*. **31**(1), pp.65-95.
- Shibatani, M. 1990. Joshi no imi to kinō ni-tsuite [About the meanings and functions of particles]. In: Kunihiro Tetsuya kyōju kanreki taikan kinen ronbunshū henshūinkai. ed. *Bunpō to Imi no aida [Between Grammar and Meaning]*. Tokyo: Kuroshio-shuppan, pp.281-301.
- Sion, A. 2007. *Nihongo no terebi nyūsu no kenkyū [Research on TV news in Japanese]*. PhD Thesis, University of Tsukuba.
- Snell-Hornby, M. 1988. *Translation Studies: An integrated approach*. Amsterdam: John Benjamins Publishing.
- Sprung, R.C. & Vourvoulias-Bush, A. 2000. Adapting Time magazine for Latin America. In: Sprung, R.C. & Jaroniec, S. eds. *Translating into Success*. Amsterdam & Philadelphia: John Benjamins, pp.13-28.
- Steiner, E. 2001. Intralingual and interlingual versions of a text – how specific is the notion of translation? In: Steiner, E. & Yallop, C. eds. *Exploring Translation and Multilingual Text Production: Beyond content*. Berlin & New York: Mouton de Gruyter, pp.161-190.
- Steiner, E. 2004. *Translated Texts: Properties, variants, evaluations*. Frankfurt am Main: Peter Lang.

- Sunakawa, Y. 2005. *Bunpō to Danwa no Setten [Crossing Point of Grammar and Discourse]*. Tokyo: Kuroshio-shuppan.
- Suzuki, S. 1972. *Nihongo Bunpō, Keitairon [Japanese Grammar, Morphology]*. Tokyo: Mugi-shobō.
- Swales, J. 1990. *Genre Analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Tanaka, S. & Nakano, H. 2018. **ST**atistical **A**nalysis **R**escuers, js-STAR (9.0.6j). [Software]. [Accessed 1 February 2018].
- Tatsuki, M. 2000. Theme-Rheme no kaishaku to Supra Theme [Interpretation of Theme-Rheme and Supra Theme] In: Koizumi, T. ed. *Gengokenkyū ni-okeru Kinōshugi [Functionalism in Language Study]*. Tokyo: Kuroshio-shuppan, pp.49-73.
- Tatsuki, M. 2004. Analyzing the Japanese Thematic structure in a communicative unit: The Kyoto grammar approach. *Doshisha Studies in English*. **76**, pp.1-20.
- Teich, E. 2001. Cross-linguistic divergence and commonality. In: Steiner, E. & Yallop, C. eds. *Exploring Translation and Multilingual Text Production: Beyond content*. Berlin & New York: Mouton de Gruyter, pp.191-227.
- Teich, E. 2003. *Cross-linguistic Variation in System and Text: A methodology for the investigation of translations and comparable texts*. Berlin: Mouton de Gruyter.
- Teramura, H. 1991. *Nihongo no Shintakusu to Imi*. Vol. 3 [Syntax and Meaning of Japanese. Vol. 3]. Tokyo: Kuroshio-shuppan.
- Teruya, K. 2004. Metafunctional profile of the grammar of Japanese. In: Caffarel, A., Martin, J.R. & Matthiessen, C.M.I.M. eds. *Language Typology: A functional perspective*. Amsterdam & Philadelphia: John Benjamins, pp.185-254.
- Teruya, K. 2006. *A Systemic Functional Grammar of Japanese*. Vol. 1 & 2. London & New York: Continuum.
- Teruya, K. 2015. Fukubunkaiseki to bunpōteki shikō [Clause complex analysis and thinking grammatically: Systemic functional approach to intermediate and advanced Japanese language teaching]. *Nihon Gakkan*. **18**, pp.32-45.

- Thompson, G. 2004. *Introducing Functional Grammar*. London & New York: Arnold.
- Thompson, G. 2007. Unfolding Theme: The development of clausal and textual perspectives on Theme. In: Hasan, R., Matthiessen, C. & Webster, J. eds. *Continuing Discourse on Language: A functional perspective*. Vol. 2. London & Oakville: Equinox, pp.671-696.
- Thomson, E.A. 1998 Testing for Theme in Japanese. *Language Education around the Globe*. **8**, pp.219-242.
- Thomson, E.A. 2001 *Exploring the Textual Metafunction in Japanese*. [Online]. PhD Thesis, University of Wollongong. [Accessed 2 February 2013]. Available from: <http://ro.uow.edu.au/theses/689>
- Thomson, E.A. 2005a. Theme unit analysis: A systemic-functional treatment of textual meanings in Japanese. *Functions of Language*. **12**(2), pp.151-179.
- Thomson, E.A. 2005b. Informing curriculum design using genre analysis: A study of three genres in Japanese. In: Burke, B. ed. *Innovative Practice in Japanese Language Education*. [Online]. Queensland: Queensland University of Technology, pp.17-41. [Accessed 24 March 2018]. Available from: https://pdfs.semanticscholar.org/d881/bdabf971292c002763f25634d32be69be330.pdf?_ga=2.15079926.1918742501.1559495272-1872168458.1532341816
- Thomson, E.A. 2013. The system of theme in Japanese. In: Thomson, E.A. & Armour, W.S. eds. *Systemic Functional Perspectives of Japanese: Descriptions and applications*. London & Oakville: Equinox Publishing Ltd, pp.101-136.
- Thomson, E.A., White, P. & Kitley, P. 2008. "Objectivity" and "hard news" reporting across cultures: Comparing the news report in English, French, Japanese and Indonesian journalism. *Journalism Studies*. **9**(2), pp.212-228.
- Toury, G. 1995. *Descriptive Translation Studies and Beyond*. Amsterdam: Benjamins.
- Tsukada, H. 2001. *Nichieigo no Shudai, Shugo soshite Shōryaku [Themes, Subjects and Ellipsis in Japanese and English]*. Tokyo: Liebel-shuppan.
- Tuchman, G. 1973. Making news by doing work: Routinizing the unexpected. *American Journal of Society*. **79**(1), pp.110-131.

- Van Leuven-Zwart, K. 1989. Translation and original: Similarities and dissimilarities. part 1. *Target*. **1**(2), pp.151-181.
- Van Leuven-Zwart, K. 1990. Translation and original: Similarities and dissimilarities. part 2. *Target*. **2**(1), pp.69-95.
- Vasconcellos, M.L. 1992. Text and translation: The role of theme and information. *Ilha do Desterro A Journal of English Language, Literatures in English and Cultural Studies*. **27**, pp.45-66.
- Vasconcellos, M.L. 2009. Systemic functional translation studies (sfts): The theory travelling in Brazilian environments. *DELTA*. **25**, [no pagination].
- Ventola, E. 1995. Thematic development and translation. In: Ghadessy, M. ed. *Thematic Development in English Texts*. London & New York: Pinter Publishers, pp.85-104.
- Vinay, J. & Darbelnet, J. 1958/2000. A methodology for translation. In: Venuti, L. ed. *The Translation Studies Reader*. 2nd ed. London & New York: Routledge, pp.84-93.
- Washitake, M. 2009. Nihongo no shimbun hōdōkiji no janrukōzō [An analysis of Japanese news stories: A pilot description of its generic structure]. *Japanese Journal of Systemic Functional Linguistics*. **5**, pp.33-45.
- White, P.R.R. 1997. Death, disruption and the moral order: The narrative impulse in mass-media 'hard news' reporting. In: Martin, J.R. & Christie, F. eds. *Genres and Institutions: Social processes in the workplace and school*. London: Cassell, pp.101-133.
- White, P.R.R. 1998. *Telling Media Tales: The news story as rhetoric*. [Online]. PhD Thesis, University of Sydney. [Accessed 8 Mar 2017]. Available from:
[http://www.prrwhite.info/prrwhite,%201998,%20Telling%20Media%20Tales%20\(unpublished%20PhD\).pdf](http://www.prrwhite.info/prrwhite,%201998,%20Telling%20Media%20Tales%20(unpublished%20PhD).pdf)
- White, P.R.R. 2000. Media objectivity and the rhetoric of news story structure. In: Ventola, E. ed. *Discourse and Community: Doing functional linguistics*. [Online]. Tübingen: Guntar Narr Verlag, pp.379-397. [Accessed 31 Aug 2016]. Available from:
<https://www.researchgate.net/publication/288219658>
- Wilss, W. 1996. *Knowledge and Skills in Translator Behavior*. Amsterdam & Philadelphia: John Benjamins Publishing Company.

Yomiuri Shimbun. 2016. *The Japan News Quick Overview*. [Online].
[Accessed 11 May 2017]. Available from:
<https://adv.yomiuri.co.jp/thejapannews/download/TheJapanNews.pdf>

List of Notations and symbols

1. Morphological notations (see Teruya 2006)

ASP	aspect: (suspensive : -te) + iru, aru, oku, etc.
ABLE	ability
COND	conditional
h.SAY	hearsay: to iu
IMP	imperative
neg	negative: -nai, -masen
NOM	nominaliser: no, koto
pst	past
psv	passive
SUSP	suspensive
VOL	volitional

2. Systemic notations (see Halliday and Matthiessen 2004)

^	ordering
/// ///	clause complex
// //	clause
[[]]	rankshifted (embedded) clause

3. Structural notations

α , β ,	and other lower case Greek letters: hypotaxis
1,2,	and other Arabic numerals: parataxis
Theme	and other terms with initial capitals: names of function
THEME	and other terms in all upper cases: names of system