ΜΕΤΑ

Explicitation, Unique Items and the Translation of English Passives in Thai Legal Texts

Dorothy Kenny et Mali Satthachai

Aller au sommaire du numéro

Découvrir la revue

Citer cet article

Résumé de l'article

Malgré l'intérêt croissant pour la traductologie en Thaïlande, il n'existe que peu de publications sur les pratiques contemporaines de traduction thaïe. Cet article présente une première tentative d'exploration de la traduction juridique en thaï, en utilisant à la fois un corpus parallèle de traités internationaux traduits de l'anglais vers le thaï, et un corpus monolingue de législation thaïe non traduite. En nous appuyant sur des études de corpus antérieures portant sur les caractéristiques générales de la traduction, nous cherchons à savoir s'il existe des preuves d'explicitation (Becher 2010a) et si certains « éléments uniques » (Tirkkonen-Condit 2002) sont sous-représentés dans nos données traduites en thaï. Pour répondre à ces questions, nous étudions la traduction de la voix passive en thaï, tout en considérant les marqueurs passifs /thùuk/, /doon/ et /dai-rup/ comme des éléments uniques en thaï. L'étude révèle que, malgré les réticences de certains auteurs à utiliser la voix passive en thaï, la plupart des passifs anglais sont traduits par un passif en thaï. En outre, dans les cas où la voix active est utilisée, il est rarement question d'explicitation. Nous ne trouvons aucune preuve à l'appui de l'hypothèse selon laquelle les éléments uniques sont sous-représentés dans la traduction. Au contraire, les éléments uniques que nous étudions semblent être surreprésentés dans la traduction en thaï.

érudit

Explicitation, Unique Items and the Translation of English Passives in Thai Legal Texts

DOROTHY KENNY
Dublin City University, Dublin, Ireland
dorothy.kenny@dcu.ie

MALI SATTHACHAI
Dublin City University, Dublin, Ireland
mali.satthachai2@mail.dcu.ie

RÉSUMÉ

Malgré l'intérêt croissant pour la traductologie en Thaïlande, il n'existe que peu de publications sur les pratiques contemporaines de traduction thaïe. Cet article présente une première tentative d'exploration de la traduction juridique en thaï, en utilisant à la fois un corpus parallèle de traités internationaux traduits de l'anglais vers le thaï, et un corpus monolingue de législation thaïe non traduite. En nous appuyant sur des études de corpus antérieures portant sur les caractéristiques générales de la traduction, nous cherchons à savoir s'il existe des preuves d'explicitation (Becher 2010a) et si certains «éléments uniques» (Tirkkonen-Condit 2002) sont sous-représentés dans nos données traduites en thaï. Pour répondre à ces questions, nous étudions la traduction de la voix passive en thaï, tout en considérant les marqueurs passifs ถูก /thùuk/, โดน /doon/ et ได้รับ /dai-rup/ comme des éléments uniques en thaï. L'étude révèle que, malgré les réticences de certains auteurs à utiliser la voix passive en thaï, la plupart des passifs anglais sont traduits par un passif en thaï. En outre, dans les cas où la voix active est utilisée, il est rarement question d'explicitation. Nous ne trouvons aucune preuve à l'appui de l'hypothèse selon laquelle les éléments uniques sont sous-représentés dans la traduction. Au contraire, les éléments uniques que nous étudions semblent être surreprésentés dans la traduction en thaï.

ABSTRACT

Despite growing interest in the study of translation in Thailand, relatively little has been published on contemporary Thai translation practices. This study presents a first attempt to explore legal translation into Thai, using both a purpose-built parallel corpus of international treaties translated from English into Thai, and a monolingual corpus of non-translated Thai legislation. Drawing on previous corpus-based studies concerned with general features of translation, we ask whether there is evidence of explicitation (Becher 2010a) in our translated data and whether "unique items" (Tirkkonen-Condit 2002) are underrepresented. We pursue these questions through a study of the translation of the passive voice into Thai, and focus in particular on the Thai passive markers ถูก /thùuk/, โดน /doon/, and ได้รับ /dai-rup/, which we consider unique items in Thai. The study finds that, despite some writers' antipathy to the passive voice in Thai, most English passives are translated into Thai using the passive voice, and that in those instances where the active voice is used in the translation, there is rarely any explicitation involved, as explicit agents are rarely added in Thai. We do not find any evidence to support the hypothesis that unique items are underrepresented in translation. On the contrary, the unique items studied appear to be overrepresented in translation into Thai.

RESUMEN

A pesar del creciente interés por el estudio de la traducción en Tailandia, se ha publicado relativamente poco sobre las prácticas contemporáneas tailandesas de traducción. Este

estudio presenta un primer intento de explorar la traducción legal al tailandés, utilizando tanto un corpus paralelo de tratados internacionales traducidos del inglés al tailandés y construido específicamente para la presente investigación, como un corpus monolingüe de legislación tailandesa no traducida. Basándonos en estudios de corpus previos, relacionados con aspectos generales de traducción, exploramos si se recurre a la explicitación (Becher 2010a) y si los «elementos únicos» (Tirkkonen-Condit 2002) están subrepresentados en la traducción tailandesa. Planteamos estas preguntas a través de un estudio de la traducción de la voz pasiva al tailandés, teniendo en cuenta los marcadores pasivos tailandeses ถูก /thùuk/, โดน /doon/ y ได้รับ /dai-rup/ como aspectos que no tienen equivalente en inglés. Del estudio se desprende que, a pesar de la reticencia de algunos escritores a utilizar la voz pasiva en tailandés, la mayoría de las oraciones pasivas en inglés se traducen a dicha lengua usando esta voz, y que en los casos en que se usa voz activa, rara vez se recurre a una explicitación. No encontramos ninguna evidencia que respalde la hipótesis de que los elementos únicos estén subrepresentados en la traducción. Más bien se produce lo contrario: los elementos únicos que estudiamos parecen estar sobrerrepresentados en la traducción al tailandés.

1. Introduction

Despite growing interest in translation studies in Thailand,[1] there is still relatively little published on contemporary Thai translation practices in general[2] and contemporary Thai legal translation in particular. There is good reason to believe, however, that Thai legal translation might be a "rich field" (see Note 2) for translation studies scholars, given its atypical development: Siam, as contemporary Thailand was previously called, is often seen as exceptional in that although it was a "victim of European imperial aggression" (Loos 1999: 2), it was not actually colonized. During this same period it harboured imperial ambitions of its own and established a legal regime that mirrored aspects of the legal systems of European colonizing powers (Loos 1999). Of particular interest here is that fact that when legal reform was undertaken between the 1890s and the 1930s, partly in a bid to keep European colonial powers at bay, it was carried out by a cosmopolitan group of Siamese and foreign experts, who used continental (code) law as a prototype, but who drafted new laws in "functional" English (Kraivixien 2012), and then translated them into Thai. Even though modern Thai legislation is drafted directly in Thai, Kraivixien (2012: 43-50) argues that the functional English used during the earlier reform period continues to exert a positive influence over legal drafting in Thai, thanks to its valorization of concision and clarity, qualities that were subsequently brought over into legal Thai. The reform period could continue to exert influence over both Thai legal drafting and legal translation in other ways however: in translating from English, Thai legal reformers were expected to minimize the "inappropriate" influence of the English language on Thai law. The passive voice, for example, which was used routinely in English legal texts, was frowned upon in Thai translation on the basis that its use in Thai indicated adversative meaning, in other words, that the action or event being described was unpleasant or undesirable (Kraivixien 2012: 171).

Likewise, there is to date, to our knowledge, almost no translation studies scholarship involving Thai that uses corpus linguistic methods. This might be understandable in light of the difficulties that the computational processing of written Thai can present (see section 5), but we would contend that there are some questions in translation studies that are best broached using corpus techniques, despite the considerable hurdles that may first have to be overcome. In this article we present initial results from a study designed to investigate such questions. We focus specifically on how instances of the English passive voice are translated into Thai in legislative texts, and on how the use of the passive voice differs between translated and non-translated legislation in Thai. For reasons elaborated upon below, we attempt to view our linguistic data through the prism of two more abstract concepts, namely explicitation (Becher 2010a) and so-called "unique items" (Tirkkonen-Condit 2002). Previous studies (Olohan 2004; Loock 2016) have suggested that translators tend to make explicit in target texts information that is left implicit in source texts, or in comparable non-translated texts in the target language; and that target language items that do not have obvious counterparts in the source language (and hence are "unique" to the target language) tend to be under-represented in translation. Such tendencies are sometimes controversially viewed as "universal" features of translation, or as "translation inherent" but it is possible to use the analytical categories of explicitation and unique items—as we do here—without subscribing to the idea of translation universals or translation inherence. Several recent studies (Becher 2010a, 2010b; Cappelle 2012) have, in fact, undermined the universalist agenda by showing how contrastive linguistic factors and typological differences affect explicitation and the distribution of unique items, respectively. And even within a single target language, different domains and different target text functions may be associated with different translation strategies: translations of legislative texts that are not intended to have any legal effect, but merely to inform readers of the contents of the source text—corresponding to Cao's (2007: 11) "informative" translation—may, for example, tolerate explicitation more than "normative" (Cao 2007: 10-12) target texts, which are intended to be equally authentic to the source text (see related arguments in Dullion 2000 and Garzone 2000). We return to these issues again below.

The article starts with a description of the passive voice in Thai, which can be assumed to be unfamiliar to most readers. This is followed by a necessarily brief explanation of explicitation and Tirkkonen-Condit's (2002) unique items hypothesis, and of why we have chosen to investigate their possible operation in legal translation into Thai. We then present our methodology, describing the corpus used in our research, and the procedure followed in extracting data from the corpus. This section is preceded by some preliminary observations on the computational processing of Thai and the availability of corpora in the language pair English-Thai. Following the presentation of data from our corpus in the results and analysis section, we discuss their interpretation from the point of view of the explicitation and unique items hypotheses. The paper ends with some observations on the limitations of the study and outlines plans for future research.

2. The passive voice in Thai

This section presents a short introduction to the passive voice in Thai, based on general linguistic and typological studies (Keenan and Dryer 1985/2007; Prasithrathsint 2003), and integrating the very limited literature on the Thai passive conducted within specific linguistic frameworks such as Systemic Functional Grammar (Patpong 2006).

The first thing to note about the Thai passive is that some commentators express skepticism about its very existence. This is because the passive is often considered a Western grammatical calque that came into Thai only in the early Bangkok period, less than 200 years ago (Prasithrathsint 2006: 116). Sudmuk (2003), for example, claims that passive structures in Thai are just "so-called" ones influenced by English. Others, such as Prasithrathsint (2003, 2006), Kullavanijava (1974), and Iwasaki and Ingkapirom (2005), contend that a true passive voice does exist in Thai, with Prasithrathsint (2006) arguing that Thai passives are true passives since they are covered by Keenan's (1985) classification of world passives. In addition, the alignment of certain passive forms with adversative meanings in particular has led some linguists to consider the passive as a marginal phenomenon in Thai, and to deny that the passive can occur with neutral or favourable meanings (Prasithrathsint 2006: 116).

But whether the passive in Thai is considered "real" or "so-called" and whether it is used mostly to convey adversative meaning or not, Thai does have constructions used to express passive meaning on which Thai linguists agree to a large extent, and which are presented briefly in the following paragraphs.

Broadly speaking, there are two types of Thai passive structures: those with passive markers, also known as overt passives, and those without passive markers, also known as covert passives.

2.1. The passive with passive markers / overt passive

Overt passive constructions in Thai use one of three lexical items: ถูก /thùuk/, โดน /doon/ or ได้รับ /dai-rup/. These items are labelled differently by different Thai linguists: Patpong (2006: 273) describes them as "lexical verbs" expressing the "receptive" voice. Iwasaki and Ingkapirom (2005: 313) sometimes refer to them as "passive auxiliaries" and sometimes as "passive markers" (2005: 316; see also Prasithrathsint 2006). In the current study, the more neutral and broader term "passive marker" is used.

The basic patterns in passive sentences that use /thùuk/ and /doon/ are identical. However, there is a slight difference in the passive pattern that uses /dai-rup/. Examples of the use of the three passive markers are given below.

2.1.1. /thùuk/ and /doon/ passives

In passives using the passive markers /thùuk/ and /doon/, the patient-subject appears before the passive marker, followed by the agent (if present) and the lexical verb (Iwasaki and Ingkapirom 2005: 313-322). Auxiliary verbs (if present) come before the passive markers. The basic structure of this type of passive is thus:

> Patient-Subject + (Auxiliary Verb) + Passive Marker [/thùuk/ or /doon/] + (Agent) + Verb

where parentheses indicate the optional nature of an element, and square brackets introduce passive markers from which one option must be chosen.

Examples 1a to 1c compare the active and passive versions of equivalent sentences, where the passive constructions use the /thùuk/ passive marker:[3]

1) Active vs. passive voice

Active sentence
ตำรวจจับขโมย
[police arrest thief]
The police arrested the thief.

Passive sentence without agent
ขโมยถูกจับ
[thief /thùuk/ arrest]
The thief was arrested.

Passive sentence with agent
ขโมยถูกตำรวจจับ
[thief /thùuk/ police arrest]
The thief was arrested by the police

The /thùuk/ and the /doon/ passives have traditionally been used to express adversative meaning, in other words, to indicate that the action or event being described is unpleasant or undesirable for the patient-subject. /Thùuk/, however, is increasingly also used in neutral situations (Prasithrathsint 1985; Iwasaki and Ingkapirom 2005).

Examples 2a to 2c illustrate the usage of /thùuk/ and /doon/ in adversative and neutral situations.

2) Adversative vs. neutral

Adversative /thùuk/
แดงถูกแม่ตี
[dang /thùuk/ mom hit]
Dang was hit by Mom.

Adversative /doon/
แดงโดนแม่ตี
[dang /doon/ mom hit]
Dang was hit by Mom.

Neutral /thùuk/
เขาถูกแต่งตั้งให้เป็นหัวหน้าชั้น
[he /thùuk/ appoint to be leader class]
He was appointed as a class representative.

The literal meaning of /thùuk/ as a transitive lexical verb is *to touch* or *to hit on the point* (Prasithrathsint 2006: 115). The meaning of /doon/ as a lexical verb is similar to that of /thùuk/, but the /doon/ passive is more colloquial than the /thùuk/ passive, and it indicates a strong sense of adversity (Iwasaki and Ingkapirom 2005: 317).

2.1.2. /dai-rup/ passive

The /dai-rup/ passive is frequently used in both positive and neutral situations. The word /dai-rup/ is a transitive compound verb made up of /dai/ and /rup/, both of which mean *receive*, and when it is used as transitive verb, it can indicate either a favourable or an unfavourable situation (Iwasaki and Ingkapirom 2005: 317).

There are two passive patterns used with the passive marker /dai-rup/. The first is identical to that used with /thùuk/ and /doon/, where the patient-subject appears before the passive marker, followed by the agent (if present) and the lexical verb.

The second pattern is normally used in formal language where /dai-rup/ can appear before what is normally a verb, but which is nominalized using the prefixes การ /kaan/ or ความ /khwaam/ (Iwasaki and Ingkapirom 2005: 319).[4] The basic structure is:

Patient-Subject + (Auxiliary Verb) + Passive Marker /dai rup/ +[/kaan/ or /khwaam/]+ Verb[5]

As already indicated, the word /dai-rup/ is also used as a lexical verb meaning *receive* or *obtain*. When this type of structure in Thai is translated into English, it can be interpreted in two ways, either as a passive construction or as an active construction with the lexical verb *receive*, as illustrated in Example 3.

3) เขาได้รับการปกป้อง
[he /dai-rup/ /kaan/ protect]
He was protected. (passive)
He received protection. (active)

The /dai-rup/ passive is different from the /thùuk/ and /doon/ passives in that it is not possible to convert an active sentence to its passive equivalent with /dai-rup/ while retaining the agent (Iwasaki and Ingkapirom 2005: 318-319).

2.2. Passive without passive markers / covert passive

Thai passive constructions without passive markers, or "covert" passives,[6] fall into two sub-types. In the first one, the patient-subject appears before the verb and is then optionally followed by a complement. The structure is as follows:

Subject-Patient+ Verb + (Complement)

As illustrated by the passive sentence in Example 4:

4) บ้านนี้สร้างในปี 1994
[house this build in 1994]
This house was built in 1994.

If the agent is important and needs to be referenced, the word โดย /dooy/, meaning *by*, is added to indicate the agent.

Subject-Patient+ Verb+ /dooy/ Agent

As shown by the passive sentence in Example 5:

5) หนังสือเล่มนี้เขียนโดยเจมส์ จอยส์
[book this write /dooy/ James Joyce]
This book was written by James Joyce.

According to Kullavanijava (1974: 196-197), this passive pattern is often heard in broadcasting and is found in written works that contain translations of English passive sentences. Iwasaki and Ingkapirom (2005: 316) agree that it has become common to use /dooy/ to translate English *by* in contemporary written Thai and when /dooy/ appears, the passive marker can be left out.

It is thus possible that some Thai covert passives are the reduced form of overt passives, but to our knowledge, no studies to date have confirmed this possibility. In addition, at least one attempt has been made to categorise some Thai covert passives as middle constructions, with Inhongsa, Louie, *et al.* (2016) suggesting that any Thai construction should be considered a middle construction rather than a covert passive when it has a subject interpreted as theme/patient along with an agentive verb and obligatory adverbial (Inhongsa, Louie, *et al.* 2016: 220). The examples given in their paper are very limited however, and it would be premature to conclude that there is, indeed, a middle construction in Thai.

For the purposes of this study then, we recognise two types of Thai passive construction, based on the discussion above, namely:

(i)   the overt passive, which uses one of the three passive markers, /thùuk/ /doon/ or /dai-rup/

(ii) the covert passive, that is, the passive structure that does not use passive markers.

3. Explicitation

As already indicated, explicitation refers to a putative feature of translation whereby target texts are said either to represent information more explicitly than their source texts, or to represent information more explicitly than similar, non-translated texts, in the target language.[7] In the former case, studying explicitation requires access to a bilingual parallel corpus; in the latter case, it requires use of a monolingual, target-language comparable corpus. Studies exemplifying the former approach include Saldanha (2008) and Becher (2010a); those exemplifying the latter approach include Olohan and Baker (2000) and Cappelle (2012). It is, of course, also possible to combine parallel and comparable corpus methods in a single study, as, for example, in Pápai (2004).

Explicitation has been a controversial idea in translation studies, with Saldanha (2008), for example, taking issue with some of the assumptions that scholars adopting the notion appear to make about the "implicitness" of information in source texts, and the supposed increased "informativeness" of explicitated target texts. Becher (2010a), likewise, criticizes research in the area for its lack of rigour, and is particularly skeptical of the notion that explicitation is a "translation-inherent" phenomenon. He ultimately concludes that the explicitation hypothesis itself should be abandoned, "[…] since it is unmotivated, unparsimonious and vaguely formulated" (2010b: 1). That other scholars have not abandoned the notion is clear from the number of studies that continue to be published on explicitation, among them a small number that deal with legal translation, and to which we return below. For now, it suffices to say that our approach to the subject follows Becher's initial definition of explicitness as "the verbalization of information that the addressee would (most probably) be able to infer from the context, her world knowledge or from other inferential sources if it were not verbalized" (Becher 2010a: 2).

We thus view "explicitation" as the process of introducing verbalizations of such information to target texts. We take a source-oriented approach in the first instance, comparing levels of explicitness in our source and target texts in Section 7.1 below. We follow this with a comparison of the use of overt passive markers in translated and non-translated Thai in Section 7.2 below, where a higher frequency of passive markers is associated with higher levels of explicitness.

3.1. Explicitation in legal texts

Very few studies have addressed explicitation in translated legal texts. The exceptions include the work of Dorrit Faber and Mette Hjort-Pederson, who investigate explicitation in student and professional translations into English of a Danish law report (Faber and Hjort-Pederson 2009a, 2009b, 2013; Hjort-Pederson and Faber 2010), and Krogsgaard Vesterager (2017), who compares translations into Danish of a Spanish judgement completed by expert legal translators, on the one hand, and professional translators who do not specialize in legal translation, on the other.

Faber and Hjort-Pederson (Faber and Hjort-Pederson 2009a: 108-109; Hjort-Pederson and Faber 2010: 238-239) suggest one reason why explicitation might be expected to be rare in legal translation, namely that the addition of any information in target texts would be a risky strategy that could change the legal meaning and effect of the text, and it is therefore likely to be avoided. This argument is obviously particularly applicable to normative translations that, by definition, have legal effect. Conversely, they suggest a reason why explicitation might be expected, nonetheless, to appear in legal translations, especially those produced by novices: the linguistic complexity of legal texts makes them difficult to understand, and may trigger a process of mental inferencing which is ultimately evidenced in translated texts. This, they claim, is consistent with Pym's (2005: 8-9) speculation that, "The harder the source text, the harder the translator works, and the more likely they are to make their renditions explicit."

What is most interesting from our point of view are the points in source texts at which Faber and Hjort-Pederson (2010: 241) expect mental and concomitant linguistic explicitation to occur, including "nominalisations, passives, culture-bound terms and elliptical phrases." Krogsgaard Vesterager (2017) follows suit in her study, ultimately concluding that explicitations of passives and nominalisations are rare in her data.

Faber and Hjort-Pederson's work is also of interest to us as they make a useful distinction between "addition" and "specification." They use "addition" to denote a type of explicitation involving "the inclusion in the TT of extra lexical elements that either add or repeat meaningful elements" (Faber and Hjort-Pederson 2013: 44). Addition thus involves a quantitative increase in textual material. "Specification," on the other hand, is qualitative in nature. That is, "it adds meaning(s) by using lexical elements that are semantically more informative than the ST lexical elements" (Faber and Hjort-Pederson 2013: 44).

It appears, therefore, that while scholars have expected source-text passive constructions to generate potential instances of explicitation in legal translation, such potential has not been realized in the findings of the relatively small-scale studies of informative legal translation conducted to date. However, the passive voice remains

an intriguing issue for us, given the differences between English and Thai discussed above. If, for example, the passive voice is considered a "marginal phenomenon" or even, somewhat derogatively, a "foreign import" in Thai (Prasithrathsint 2006), then we might expect the English passive voice to be replaced by the active voice in translation into Thai. And if "basic" passives (Keenan and Dryer 1985/2007: 328) involve the erasure of agents, then conversion to the active voice could involve explicitation of agents (as long as they can be inferred from the source text/context) or even the "insertion" of agents (in the event that the translator inserts genuinely new information that could not have been inferred from the original text or context, in line with Becher's (2010a: 2) definition given above).[8] On the other hand, even if the passive voice is a relatively new phenomenon in Thai, it is also argued that Thai has by now well-established markers, namely ถูก /thùuk/, โดน /doon/, and ได้รับ /dai-rup/, which indicate the passive voice, but may also signal adversative, neutral or positive meanings (Prasithrathsint 2003, 2006). These markers have no obvious formal counterparts in English, the source language under consideration here. While English marks the passive voice using the auxiliary verb *be* and the past participle of the lexical verb, it does not have specific grammatical or lexical words that indicate, for example, both the passive voice and adversative meaning in the same token. Relative to English, /thùuk/, /doon/, and /dai-rup/ thus appear to be "unique items" in the sense first put forward by Tirkkonen-Condit (2002), and expanded upon in the next section. What is also interesting about these items from our point of view is that if they are used in translation into Thai, in instances where the adversative, positive or neutral nature of the situation or action described in English is not explicitly verbalized but may be inferred—again as per Becher's (2010a: 2) definition—then the presence of these items in Thai could constitute evidence of explicitation, whereby, for example, the adversative nature of an action in English is made explicit in the Thai passive marker /thùuk/ or /doon/.

The discussion so far suggests that we might expect explicitation in translation from English into Thai on two counts when it comes to the treatment of voice. Firstly, if translators into Thai have an aversion to the passive voice, then they may replace it by the active voice, and in doing so they may make agents explicit that were not explicit in the source text (although they could be inferred) or they may insert wholly new agents. Both of these actions would involve a quantitative increase in the information in the target text compared to the source text, and the former would certainly fall within Faber and Hjort-Pederson's (2013) category of explicitating "addition," while the latter would constitute an insertion, as described above. Secondly, and conversely, if translators preserve passives in translation into Thai, then they might use the above-mentioned passive markers and thus, in the target language, make pragmatic information (related to their attitudes to given events or actions, where such events or actions are seen as positive, neutral or adversative) explicit that was not made explicit in the source, although it could be inferred. Relatively frequent use of these passive markers in translation into Thai (compared to native Thai) would, however, run counter to another commonly hypothesized feature of translated texts, namely that they make less use of so-called "unique items" than do other texts in the same language (Tirkkonen-Condit 2002, 2004; Cappelle 2012).

4. Unique items

According to Tirkkonen-Condit (2004: 177-178):

> Every language has linguistic elements that are unique in the sense that they lack straightforward linguistic counterparts in other languages. These elements may be lexical, phrasal, syntactic or textual, and they need not be in any sense untranslatable; they are simply not similarly manifested (e.g. lexicalized) in other languages. Since they are not similarly manifested in the source language, it is to be expected that they do not readily suggest themselves as translation equivalents, as there is no obvious linguistic stimulus for them in the source text.

On this basis, Tirkkonen-Condit (2004) goes on to suggest that such "unique items" are likely to be under-represented in translated texts compared to non-translated texts in the same language. This "unique items hypothesis" can be criticized on a number of grounds (Chesterman 2007). One relates to the very definition of "unique" items. While the term suggests absolute exclusivity to the target language in question, in practice it is applied (by Tirkkonen-Condit and others) in cases where a target-language item does not have a "straightforward linguistic counterpart" merely in one particular source language. By way of illustration, take one of the above-mentioned passive markers in Thai: /thùuk/. There is no straightforward linguistic counterpart for /thùuk/ in English, so if we are analyzing translations from English into Thai, we might consider /thùuk/ as a unique item, which we might in turn expect to be underrepresented in such translations (compared to native Thai texts). If, however, we are analyzing translations from Vietnamese into Thai, we are obliged to recognize the existence of a more-or-less straightforward linguistic counterpart of /thùuk/ in the source language (Keenan and Dryer 1985/2007: 338; Prasithrathsint 2006: 116). In this case, /thùuk/ loses its status as a "unique item." Despite this and other issues with the definition and identification of unique items (Chesterman 2007), we will continue to use the term here in its relative sense of "occurring in a given target language but not in a given source language," for the sake of convenience—as corpus-based translation scholars are familiar with the term—and to maintain consistency with other studies that provide useful models for the current research (Cappelle 2012). Specifically, if we were to find that the above-mentioned passive markers are used in translation into Thai less frequently than in comparable original texts in Thai, we would have evidence to support the unique items hypothesis. If, however, we see more instances of the passive voice in translated Thai than in non-translated Thai, perhaps because the source texts "shine through" (Teich 2003), then we may expect to see more frequent use of passive markers in translated text than in non-translated Thai text, thus contradicting the unique items hypothesis.

5. Corpus resources in Thai

Thai is generally considered a "resource-poor" language (Abdelali *et al.* 2014), although efforts have been made to make monolingual Thai corpora available to the research community through a variety of channels. One of the earliest such initiatives appears to have been ORCHID, a part-of-speech tagged corpus of technical texts compiled to support Natural Language Processing (NLP) research in Thai

(Sornlertlamvanich, Charoenporn, *et al*. 1999). The Thai National Corpus (TNC) (Aroonmanakun 2007; Aroonmanakun *et al*. 2009) was subsequently designed as a general reference corpus of standard Thai and intended to be comparable to the written part of the British National Corpus. It is freely accessible online through Chulalongkorn University.[9] As of late 2017 it contained some 32 million words, over 1 million of which appear to be from legal texts, described as "regulation."[10] However, it is difficult to work out the exact composition of this subcorpus, which prevents us from integrating it into our study for the moment at least. The TNC is word tokenized—which is not a trivial matter, given that Thai does not use a white space to mark word boundaries—but it is not part-of-speech tagged. The corpus is accessed through a web interface that allows concordance searches and provides information on the collocations and distributions of the given keyword.

Other monolingual Thai resources of which we are aware are composed of either untranscribed telephone conversations[11] or transcribed children's speech,[12] or written texts harvested from websites, as in the case of the freely accessible 50-million word HSE Thai Corpus[13] and the 82.8 million word ThaiWaC corpus, accessible through Sketch Engine.[14] Finally, Thai was one of the South East Asian languages covered by the now-discontinued SEAlang Library Project, which provided access to dictionaries as well as monolingual and bilingual (that is to say parallel) corpus resources.[15]

In general, however, parallel corpora containing Thai are extremely rare. At least some data in Thai is included in the AMARA corpus of aligned educational video subtitles (Abdelali *et al*. 2014), and in Christodouloupoulos and Steedman's (2015) "massively parallel" corpus of the Bible, but we are not aware of any easily accessible general language parallel corpus that includes Thai, or any parallel corpus of legal language that involves Thai.[16] Nor are we aware of any Thai corpora, monolingual or bilingual, being used for translation studies research; the focus so far in Thai corpus development has instead been on the creation of corpora for use in general and computational linguistics (Kawtrakul *et al*. 2002) and especially in NLP applications such as speech recognition and machine translation (for example, AMARA).

Given the lack of suitable existing resources, the present study had to rely on a specially constructed parallel corpus of legal texts in English and Thai, and a specially constructed monolingual corpus of Thai legal texts, both of which are described below.

6. Methodology

In this section we briefly describe the corpora used in the current study, before outlining how they were queried. Sketch Engine was used to manage and process both corpora as it offers all standard corpus processing functions, it can handle the Thai character set, and has implemented the SWATH word tokenizer for Thai.[17] Neither corpus is part-of-speech tagged for the moment.

6.1. Corpora

Our self-built **parallel corpus** consists of the full texts of 25 international treaties in English and their translations into Thai. The corpus contains approximately four hundred thousand words (tokens). The source texts are all international treaties

entered into by Thailand and foreign countries, and came into force between the 1950s and the present day. Only the English source texts are considered authoritative texts and the Thai translations have no legal force. The Thai translations are, however, vital to the civil servants and others who use them for information purposes. All Thai translations were completed and published by Thai government agencies. The original texts and translations were obtained in electronic format from web sources. They were aligned manually in a format suitable for use with the Sketch Engine platform.[18]

Our self-built **monolingual corpus of comparable Thai legislative texts** consists of approximately one million words (tokens). The texts are all full, written, legislative texts enacted by a legislative body and came into force between the 1970s and the present day. The texts were obtained from the website of the Office of the Council of the State.[19]

Both our legal corpora thus consist entirely of legislation, one of the four categories of legal text type recognized by Cao (2007), and the Thai translations in the parallel corpus and the texts in the Thai monolingual corpus are deemed comparable along this and other dimensions (for instance, they were created roughly over the same time frame). We acknowledge here, however, that while our parallel corpus contains informative translations in Thai with no legal force, our monolingual corpus of legislation in Thai does have legal force, and so comparability in terms of legal function is somewhat undermined (Dullion 2000). Furthermore, as with other legal corpora, and especially parallel legal corpora "which require time-consuming alignment" (Biel 2014: 94), our corpora are relatively small and are dominated by legislation. They are, nevertheless, substantial resources by Thai standards, and indeed by the standards of many other less researched languages (Biel 2014: 93).

6.2. Corpus processing

Our empirical research begins with the identification in the English source texts in our parallel corpus of all instances of the passive voice constructed using the lemma *be* and a past participle. To do this, query (i) was created using Sketch Engine's Corpus Query Language (CQL):

[lemma="be"][tag="VVN"]

We wished to allow for cases where other constituents occurred between *be* and the past participle, and so also used the following CQL searches:

[lemma="be"][][tag="VVN"]

where [] indicates the presence of one token (for instance, *immediately* in the string *X is immediately withdrawn*) between *be* and the past participle, and:

[lemma="be"][][]{2}[tag="VVN"]
[lemma="be"][][]{3}[tag="VVN"]
[lemma="be"][][]{4}[tag="VVN"]

which allow for two, three or four tokens between *be* and the past participle.

Queries (i) and (ii) yielded the vast majority of instances of the passive voice in the English corpus, with 2,291 and 219 hits respectively. Hits in this case took the form of bilingual concordance lines. Queries (iii), (iv), and (v) generated considerable

noise, with only 20%, 9%, and 1% of hits actually constituting passive patterns. We thus decided to focus entirely on hits for queries (i) and (ii). Using Sketch Engine's random sample function, a sample of 230 instances (~10%) was taken from the 2,291 hits for query (i). Likewise, a sample of 22 instances (again, ~10%) was taken from the 219 hits for query (ii).[20] Finally, all 252 concordance lines in our random sample were double checked to ensure that they were, indeed, passive constructions. The next step in the procedure involved manually inspecting the Thai translation for each of the 252 instances of the English passive voice in our random sample, to find out how they were translated. The findings of this analysis are presented in section 7.1 below.

As our research was concerned not just with how the English passive is translated into Thai, but also with whether the distribution of the "unique items" ถูก /thùuk/, โดน /doon/, and ได้รับ /dai-rup/ differs between translated Thai and non-translated Thai, we also needed to search for instances of these lexical items on the target side of our parallel corpus and in our monolingual corpus of legal Thai. In this case, "simple" searches were conducted for the items ถูก, โดน, and ได้รับ. Unsurprisingly, given its very colloquial nature, no hits were returned for the form โดน /doon/ in either corpus, whereas 1,509 and 3,867 hits were returned for the forms ถูก /thùuk/ and ได้รับ/dai-rup/ respectively. As these forms are homographic with full lexical verbs or else may occur as part of certain compounds that were not accurately tokenized by SWATH, many hits for /thùuk/ and /dai-rup/ were false positives and needed to be weeded out. This was done manually, but the process was greatly assisted by the flexibility of the Sketch Engine user interface. The frequency of /thùuk/ and /dai-rup/ used as passive markers in the two Thai corpora was thus established. Given the difference in size between the two Thai corpora, the results for each were normalized by calculating frequency per million words. Further details are given in section 7.2 below.[21]

7. Results and analysis

In this section we present first the results from our analysis of the translation into Thai of instances of the English passive voice (Section 7.1). We then present the results from our comparison of the distribution of the passive markers /thùuk/ and /dai-rup/ in the Thai target texts in our parallel corpus, on the one hand, and our non-translated Thai legislative texts, on the other (Section 7.2).

7.1. Bilingual parallel corpus

Table 1 shows how English passive clauses are translated into Thai in our parallel corpus. In roughly two thirds of these cases (159 out of 252 instances or 63% of the time), Thai translators translate English passives using Thai passives. In such cases, the covert passive is used almost twice as often (105 out of 252 instances, or 42% of the time) as the overt passive (54 instances, or 21%). The second most common technique is to convert the English passive clause into a Thai active clause, which happens just over one quarter of the time (66 cases, or 26%). A relatively small number of other techniques are also employed, including omission of the passive clause or the shifting of the passive clause to a lower rank, such as a nominal group. There are 27 such instances in our data, accounting for 11% of the 252 cases studied.

Table 1

Three main techniques employed by Thai translators when translating English passives

Translation Techniques		Number of occurrences	Percentage (%)
English passive to Thai passive	Covert passive	105	42
	Overt passive	54	21
	/thùuk/	*23*	
	/dai-rup/	*31*	
	Sub-total	**159**	**63**
English passive to Thai active	**Sub-total**	**66**	**26**
Others (omission, change, clause to phrase)	**Sub-total**	**27**	**11**
	TOTAL	**252**	**100**

A more detailed, qualitative analysis of the data depicted in Table 1 follows.

7.1.1. *English passive to Thai covert passive*

As previously mentioned, 105 instances (or 42%) of the English passives were translated into Thai covert passives. The high frequency of this passive pattern in the Thai translations is consistent with Kullavanijava's (1974: 196-197) and Iwasaki and Ikaphirom's (2005) claim that Thai covert passives are characteristic of translations from English. In the absence of a comparative study of the distribution of covert passives in comparable translated and non-translated Thai texts, however, it is not possible to establish a causal relationship between the translation condition and the prevalence of covert passives in translated Thai.

The Thai covert passive in our parallel corpus was found in both patterns described above, namely:

Subject-Patient + Verb + (Complement)

and:

Subject-Patient + Verb + /dooy/ Agent

as highlighted in Examples (6) and (7) respectively.[22]

6) *…a certified copy shall be sent, through diplomatic channels, to each of the States Members of the Hague Conference on Private International Law…*
…สำเนาเอกสารดังกล่าวจะส่งตามวิถีทางการทูตไปยังแต่ละรัฐสมาชิกของที่ประชุมกรุงเฮกว่าด้วยกฎหมายระหว่างประเทศแผนกคดีบุคคล…

[…**copy document such will send** through channel diplomat to each state member of the conference Hague on law international department case private…]

7) ***Any organization*** *referred to in paragraph 1 above which becomes a Party to this Convention without any of its member States being a Party* ***shall be bound by*** *all the obligations under the Convention.*
องค์การใดที่อ้างถึงในวรรค 1 ข้างต้นซึ่งเข้าเป็นภาคีอนุสัญญานี้โดยที่ไม่มีรัฐสมาชิกใดของตนเป็นภาคีต้องผูกพันโดยพันธกรณีทั้งปวงภายใต้อนุสัญญานี้

[**organization any** that refer in paragraph 1 that become party convention without state member as party **shall bind /dooy/** obligation all under convention this]

Based on our findings, the verbs used in this pattern are very varied, but none of them is particularly associated with adversative situations.

7.1.2. English passive to Thai overt passive

As already indicated, 54 instances of the overt passive were found in our Thai translation data (accounting for 21% of our original 252 cases). Of these, 23 are passive constructions with the passive marker /thùuk/, of which 9 express adversative meaning while the others express either positive or neutral meaning. Examples (8) and (9) illustrate the /thùuk/ passive used to express adversative and neutral meaning respectively.

8) ***An alien*** *lawfully in the territory of a State Party to the present Covenant* ***may be expelled*** *therefrom only in pursuance of a decision reached in accordance with law.*
คนต่างด้าวผู้อยู่ในดินแดนของรัฐภาคีแห่งกติกานี้โดยชอบด้วยกฎหมายอาจถูกไล่ออกจากรัฐนั้น ได้โดยคำวินิจฉัยอันได้มาตามกฎหมายเท่านั้น

[**alien** who live in territory of state party this lawfully **may /thùuk/ expel** from state that by decision from law only]

9) *Any such plan should be transmitted to the Secretariat as soon as **it has been developed**.*
แผนใด ๆ ดังกล่าวควรจะถูกกล่งให้กับสำนักเลขาธิการ โดยเร็วที่สุดหลังจากที่ถูกพัฒนาขึ้น

[plan any such should will /thùuk/ send to secretariat as soon as **/thùuk/ develop**]

Our data thus appear to confirm the claim, reported above, that /thùuk/, although traditionally used in adversative situations, is now increasingly also used in neutral contexts. Such observations may indicate the increasing grammaticalization of forms such as /thùuk/.[23]

The /dai-rup/ passive appears in the other 31 cases where an English passive is translated by a Thai overt passive. In 30 out of 31 instances it is used to convey a positive or neutral meaning.

There are 8 instances where /dai-rup/ is followed by a verb, and 23 instances where it is followed by [/kaan/ or /khwaam/] + Verb. Example (10) illustrates the use of the /dai-rup/ passive in the former pattern while Example (11) illustrates the latter.

10) *One competent authority **shall be designated** to receive the notification in case of a State of transit.*
ศูนย์ประสานงานหนึ่งแห่ง ต้องจะได้รับแต่งตั้งให้รับการแจ้งในกรณีเป็นรัฐที่ถูกนำผ่านแดน

[centre co-ordinate one **shall will /dai-rup/ designate** to receive notification in case be state which transits]

11) *…and **their defence counsel**, as well as persons participating in the investigation, **are protected** against all ill-treatment or intimidation…*
และทนายของบุคคลดังกล่าว ตลอดจนบุคคลอื่นๆ ที่มีส่วนร่วมในการสอบสวน ได้รับการคุ้มครองจากการปฏิบัติที่เลวร้าย หรือการข่มขู่

[…and their defence counsel as well as person other who participate in investigation **/dai-rup/ /kaan/ protect** from treatment bad or intimidation]

7.1.3. English passive to Thai active

As indicated above, the second most common technique used by the translators when translating the English passive to Thai is to convert English passive clauses into Thai

active clauses, which occurred in 66 (26%) out of 252 instances. The resulting Thai active clauses can be divided into three sub-categories as described below.

7.1.3.1. Translation to Thai full active clause

In 21 cases where an explicit agent in the English clause was available, translators converted the English passive clause into an active Thai clause with SVO word order as illustrated in Example (12).

> 12) *…no further communication by any state Party shall be received under the present article after **the notification of withdrawal of the declaration has been received by the Secretary-General**, unless the state Party concerned has made a new declaration.*
> ภายหลังเลขาธิการได้รับเรื่องแจ้งขอถอนประกาศแล้ว จึงไม่ต้องมีการรับคำร้องเรียนอื่นใดโดยรัฐภาคีภายใต้มาตรานี้อีก ยกเว้นแต่รัฐภาคีที่เกี่ยวข้อง ได้ทำประกาศขึ้นใหม่
>
> […after **secretary general receive the notification of withdrawal** so must not have receipt of communication any by state party under article this unless state party that relate make declaration new]

In a further 5 instances translators not only converted English passive to Thai active, they also put ให้ /hây/ at the beginning of the sentence as in Example (13).

> 13) *The frequency of subsequent submission of information required under this Article **shall be determined** by the Conference of the Parties serving as the meeting of the Parties to this Protocol…*
> ให้[ที่ประชุมสมัชชาประเทศภาคีอนุสัญญาฯ ที่เป็นการประชุมของประเทศภาคีพิธีสารนี้]กำหนด(ระยะเวลาในการนำเสนอข้อมูลครั้งต่อๆ ไปที่กำหนดภายใต้พิธีสารนี้)
>
> [**Hây** [meeting conference party serve as meeting of party protocol] **determine** (period of time to submit information subsequent require under protocol this)]

A detailed discussion of /hây/ is beyond the scope of this article. Suffice it to say here that the placing of /hây/ at the beginning of sentences like that in Example (13) creates a causative construction (Iwasaki 1998) in which one entity, the "trigger," causes, enables or obliges another entity, the "instigator," to carry out some action. In Example (13) the trigger is only implied (but is understood to be the law itself), while the instigator is *Conference of the Parties*. Here /hây/ also expresses a degree of obligation similar to that of the modal *shall* in the English source text, and the analysis of the translation of the passive voice overlaps with the analysis of modality, which we treat in depth elsewhere (Kenny and Satthachai 2018).[24]

7.1.3.2. Translation to Thai active clause without subject

In 26 cases where there was no explicit agent in the English passive clause, the translators translated the English passive to a Thai active clause without a subject using the pattern: Verbal group + Complement, as illustrated in Example (14).

> 14) *The persons concerned **shall be informed** of these rights before or at the latest at the time the decision is rendered.*
> ต้องแจ้งสิทธิเหล่านั้นแก่บุคคลที่เกี่ยวข้องก่อน หรืออย่างช้าที่สุดในเวลาที่ออกคำตัดสินนั้น
>
> [**Shall inform** right those to person related or latest in time issue decision that]

In a further 10 instances the translators used a /hây/ causative construction, but unlike in the instance referred to in Example (13) above, in this case the translations

did not express an instigator, but followed the structure: /hây/ + Verb + Complement, as illustrated in Example (15).

15) *Any dispute arising from the interpretation of the Charter* **shall be settled** *in accordance with the relevant provisions in Chapter VIII.*
ให้ระงับข้อพิพาทใดๆ ที่เกิดจากการตีความกฏบัตรตามบทบัญญัติที่เกี่ยวข้องในหมวด 8

[/**Hây**/ **settle dispute** any arise from interpretation according to provision related in chapter 8]

The omission of an instigator, and hence of a grammatical subject, in the Thai translation in Example (15) (and which makes the gloss back translation read like an imperative) may be linked to the absence of an agent in the passive construction used in the source text.

7.1.3.3. Translation to Thai active clause with added subject

Third, there are four instances in our data where there is no agent specified in the English passive, but the translator added an agent as subject in the Thai translation, as illustrated in Example (16).

16) **External parties may be invited to ASEAN meetings** *or cooperative activities without being conferred any formal status, in accordance with the rules of procedure.*
อาเซียนอาจเชิญภาคีภายนอกให้เข้าร่วมการประชุมหรือกิจกรรมความร่วมมือโดยมิต้องกำหนดให้สถานภาพ
อย่างเป็นทางการใดๆ ตามกฏว่าด้วยขั้นตอนการดำเนินงาน

[**asean may invite party external** to join meeting or activity without confer status formal any in accordance with rule of procedure]

In these cases, we assume that the translator has verbalized "information that the addressee would (most probably) be able to infer from the context, her world knowledge or from other inferential sources if it were not verbalized" in line with Becher's definition of explicitation given above (Becher 2010a: 2), and so these four cases constitute genuine instances of explicitation. We return to this issue again below, after results from our monolingual comparable corpus investigations have been presented.

7.2. Translated vs non-translated Thai

Table 2 shows the distribution of /thùuk/ and /dai-rup/ used as passive markers (see section 6.2 above) in the Thai translations in our parallel corpus and in non-translated texts in our monolingual Thai corpus.

Table 2

Occurrences of passive markers /thùuk/ and /dai-rup/ in Thai translation and comparable non-translated Thai

	Corpus	Number of occurrences	Corpus size (tokens)	Frequency (per million)
/thùuk/	Thai translation	465	222,556	2,089.36
	Non-translated Thai	1,471	1,139,743	1,290.64
/dai-rup/	Thai translation	1,271	222,556	5,710.90
	Non-translated Thai	3,713	1,139,743	3,257.75

Figure 1 provides an alternative visualization of the relative frequencies of /thùuk/ and /dai-rup/ in the same two corpora.

Figure 1

Frequency per million words of the passive markers /thùuk/ and /dai-rup/ in Thai translated and Thai non-translated legal corpora
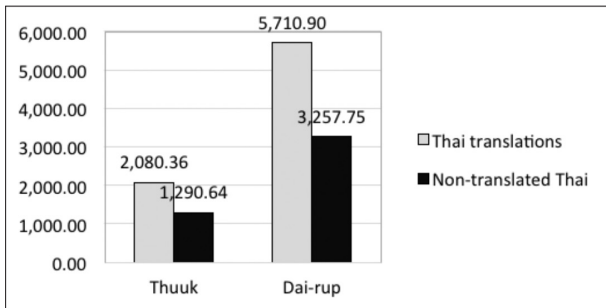

Figure 1 suggests that /thùuk/ and /dai-rup/ are overrepresented in legislative translations into Thai compared to comparable legislative texts originally written in Thai. The implications of this finding for the unique items hypothesis are discussed below.

8. Discussion

This study began with the observation that explicitation might be expected to be a feature of legal translation into Thai if translators (a) sought to avoid the passive voice and subsequently replaced it by the active voice on occasion, thereby making explicit agents in the target language that were not explicit in the source text or (b) if they preserved passives in translation into Thai, but in doing so added Thai passive markers that in turn made explicit the favourable, neutral or adversative nature of the situation described. Our first major finding is that our data do not support the expectation that the passive is avoided in translation into Thai. In 63% of the cases, English passives are translated by Thai passives. In slightly more than 26% of the cases, however, English passives are translated by Thai active constructions, but in only a tiny number of cases (4 out of 252, or slightly less than 1.6%) are agents that are only implicit in the source made explicit in the target. This means that in general, the conversion of passives into actives in our parallel corpus does not generally coincide with the kind of quantitative explicitation called "addition" by Faber and Hjort-Pederson (2013: 44).

In the 63% of the cases where English passives are translated by Thai passives, we found that this is mostly achieved using the covert passive, which, again, does not involve explicitation. In fact, uses of the covert passive outnumber uses of the overt passive almost two to one. In the 54 cases in our data (or 21% of cases) where an English passive is translated using a Thai overt passive, in other words the use of /thùuk/ or /dai-rup/, we tentatively claim that explicitation can be observed. We are tentative in this claim as it is difficult to say whether there is a self-evident increase in the amount of textual material in the target language vis-à-vis the source language,

as required in the additive kind of explicitation recognized by Faber and Hjort-Pederson (2013: 44) and discussed above: /thùuk/ or /dai-rup/ coincide with morphological markers of the passive in English after all – namely forms of the auxiliary verb *be* and past participles. Having said that, the particular markers that Thai uses often appear to reinforce or repeat meanings already implicit in the lexical verbs with which they collocate, as well as signalling the passive voice, and on this basis their use might be seen as additive and explicitating. Given the fairly marginal use of these passive markers in translations of English passives into Thai, however, along with the extremely infrequent addition of subject-agents in cases where English passives are translated by Thai active clauses, we can say that the translation of the passive voice from English into Thai provides only very weak evidence to support the hypothesis that explicitation is a feature of translation in the text type, language pair, and direction under consideration.

Our monolingual comparable data, on the other hand, seem to provide strong evidence *against* Tirkkonen-Condit's (2002) unique items hypothesis as Thai passive markers are over-represented rather than under-represented in our translated data compared to comparable target-language data. The relative frequency of the /thùuk/ passive in our Thai legal translations is almost twice that of our non-translated texts, and the relative frequency of the /dai-rup/ passive in the translations is around 40% higher than that of the monolingual Thai legal corpus. This finding could be due to the source language "shining through," to use Teich's (2003) expression.[25] That is, the over-representation of passive markers in legislation translated into Thai could simply be due to the over-representation of the passive voice in general in such translations, brought about by the influence of their passive-laden source texts. Pursuing this idea would require a larger-scale study of passives of all kinds (that is, not just of overt passives) in our Thai monolingual comparable corpus. It would also require us to investigate whether there are instances of the passive voice in our Thai translations that do not coincide with the passive voice in English.

9. Conclusions

In this article we have described one of the first corpus-based studies, to our knowledge, of English to Thai translation, based on two self-built corpora: an English-Thai parallel corpus of legislative texts and a monolingual corpus of non-translated legislative texts in Thai. Our first exploration of these corpora involved investigation of how the passive voice is translated into Thai, and we have provided more empirical evidence on this topic than previous discussions, which were not based on corpora. However, our study leaves many questions unanswered, such as how the use of the *covert* passive differs between translated and non-translated Thai. It also investigates unique items and explicitation through very narrow windows, and assumes comparability of normative non-translated legislative texts with merely informative translations. Having already created the corpus resources used in this study however, we are in a position to widen our investigation to consider other manifestations of the passive voice as well as related linguistic features in translation from English to Thai, and to continue refining our analyses, taking leads from the wide-ranging corpus studies conducted by Teich (2003) and Biel (2014), among others.

NOTES

1. This interest is manifest in the number of Thai Universities now offering full programmes in translation studies—these include Chulalongkorn University, Mahidol University and Thammasat University (all in Bangkok)—or translation modules as part of a modern languages degree, for example, at Chiang Mai University. Recent years have also seen major translation studies conferences and workshops held in Bangkok, including what was billed as the First International Conference on Translation Studies in June 2014 (http://icts.utcc.ac.th/), and an ARTIS Workshop in July 2017 (https://artisinitiative.org/events/upcoming-events/artisbangkok-translatorship/).

2. Wakabayashi, Judy (2014): Positionality, Representation and Questions on Translation in Thailand, unpublished. *Translating Asia: Then and Now. First International Conference on Translation Studies: Translating Asia*. Bangkok, 5-6 June 2014. Visited 11 November 2017, <http://icts. utcc.ac.th/wp-content/uploads/2014/12/Positionality-Representation-and-Questions-on-Translation-in-Thailand.pdf>.

3. Note that, unlike English, Thai is not inflected, and there is no morphological marking of tense, case or number. Linguistic glosses are thus kept to a minimum in these examples. The basic word order in Thai is Subject-Verb-Object.

4. The words การ /kaan/ and ความ /khwaam/ are classified as noun-forming prefixes: prefixes that attach to a non-nominal word to create nouns (Iwasaki and Ingkaphirom 2005: 28-29). The prefix /kaan/ forms an activity noun with a verbal root (either with action verbs or mental-state verbs), and can also precede a verbal phrase while the prefix /khwaan/ forms an abstract noun from an adjectival or verbal root and may appear before a verb phrase (Iwasaki and Ingkaphirom 2005: 28-29).

5. Entire clauses can be nominalized in this way (that is, verb + complement) (Iwasaki and Ingkaphirom 2005: 28-319).

6. According to Iwasaki and Ingkapirom (2005: 316), passive constructions that do not use passive markers are sometimes also called "verb passives."

7. The first understanding of explicitation sees the phenomenon as a manifestation of a "potential S-universal," while the second understanding suggests the operation of a "potential T-universal" to use Chesterman's (2004) terms.

8. The word "insertion" is used here to avoid confusion with Faber and Hjort-Pederson's (2013:44) "addition."

9. See *Thai National Corpus*. Bangkok: Department of Linguistics (Chulalongkorn University). Visited 3 April 2018, <http://www.arts.chula.ac.th/~ling/TNCII/>.

10. See *Thai National Corpus*. Bangkok: Department of Linguistics (Chulalongkorn University). Visited 3 April 2018, <http://www.arts.chula.ac.th/~ling/TNC/contents/File/tncstat-3-23-2013.txt.pdf>.

11. See NIST Multimodal Information Group (16 November 2011): *2006 NIST Speaker Recognition Evaluation Training Set*. Philadelphia: Linguistic Data Consortium. Visited 18 December 2017, <https://catalog.ldc.upenn.edu/LDC2011S09>. See also NIST Multimodal Information Group (21 October 2011): *2008 NIST Speaker Recognition Evaluation Training Set*. Philadelphia: Linguistic Data Consortium. Visited 18 December 2017, <https://catalog.ldc.upenn.edu/LDC2011S08>.

12. The Thai Childes corpus, which contains 243,939 words of transcribed speech, is available through Sketch Engine. See <https://www.sketchengine.co.uk/user-guide/user-manual/corpora/by-language/thai-text-corpora/>.

13. Accessible here: <http://web-corpora.net/ThaiCorpus/search/?interface_language=en>. The HSE Thai Corpus is part-of-speech tagged and each token in the corpus is assigned an English translation (based on bilingual dictionary entries). The corpus is thus searchable by word form, lemma and English word-level translation, although it remains a monolingual corpus. It appears to have been developed sometime after 2014.

14. See <https://www.sketchengine.co.uk/thaiwac/>. ThaiWaC is tokenized using the SWATH word tokenizer.

15. Bickner, Robert J. and Cooper, Doug (Last update: 4 January 2013): *SEAlang Library*. Madison: Center for Southeast Asian Studies (University of Wisconsin-Madison). Visited 5 January 2018, <http://sealang.net/library/>.

16. We also note that GitHub does not list any parallel corpus resources for Thai. See <https://github.com/kobkrit/nlp_thai_resources>.

17. The SWATH (Smart Word Analysis for Thai) word tokenizer was developed by Paisarn Charoenpornsawat and is available for download. See Charoenpornsawat, Paisarn (7 October 2003): SWATH. Version 2.0.1. Visited, <http://www.cs.cmu.edu/~paisarn/software.html>.

18. Manual intervention is necessary as Thai does not use punctuation to mark sentence boundaries. Nor is there any reliable way to use information from whatever white space might be found in a text to recognize sentence boundaries (see Palmer 2000/2010: 22). This makes it extremely difficult to automate sentence tokenization and subsequent alignment.

19. *Law Translation*. Bangkok: Thai Council of State. Visited, 8 August 2017, <http://www.krisdika.go.th/wps/portal/general_en>.

20. Random sampling ensures that, in cases where there is too much data for the linguist to analyze manually, a more manageable, but unbiased sample can be analyzed instead. Given the lack of bias in the sampling strategy, there is no reason to believe that the 252 instances extracted are not representative of the approximately 2,500 instances of passive voice identified in the English source texts in the corpus.

21. In this case, the intention was to conduct a purely quantitative analysis of the passive markers in question, once identified, and no further sampling to reduce the data set was necessary.

22. Note that only the instance of the passive voice under discussion is highlighted in these examples, and the concordance lines may contain further instances of the passive voice that are not discussed here.

23. Put simply, grammaticalization concerns "the development of lexical items into markers of grammatical categories" (Bisang 2017: n. p.).

24. Kenny, Dorothy and Satthachai, Mali (2018): Deontic Modality in English-Thai legal translation. *TC18: Transius Conference*. Geneva, 18-20 June 2018.

25. We use the term "shining through" rather than the related term "translationese" here, as the latter is sometimes used pejoratively (see Olohan 2004: 90) and we wish to avoid any negative connotations.

REFERENCES

Abdelali, Ahmed, Guzman Francisco, Sajjad, Hassan, *et al.* (2014): The AMARA Corpus: Building Parallel Language Resources for the Educational Domain. *In*: Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, *et al.*, eds. *LREC 2014 Proceedings*. (LREC 2014: Ninth International Conference on Language Resources and Evaluation, Reykjavik, 26-31 May 2014). Paris: European Language Resources Association. Visited 11 November 2017, <http://www.lrecconf.org/proceedings/lrec2014/index.html>.

Aroonmanakun, Wirote (2007): Creating the Thai National Corpus. *In*: Montira Rato and Raquel A. G. Reyes, eds. *Gender and Sexuality in Modernising Southeast Asia, 16th-20th Centuries. Manusaya*. 13:4-17.

Aroonmanakun, Wirote, Tansiri, Kachen, and Nittayanuparp, Pairit (2009): Thai National Corpus: A Progress Report. *In*: Hammam Riza and Virach Sornlertlamvanich, eds. *Proceedings of the 7th Workshop on Asian Language Resources*. (ACL-IJCNLP 2009: The 7th Workshop on Asian Language Resources, Singapore, 6-7 August 2009). Stroudsburg: Association for Computational Linguistics, 153-160.

Becher, Viktor (2010a): Towards a More Rigorous Treatment of the Explicitation Hypothesis in Translation Studies. *trans-kom*. 3(1):1-25.

Becher, Viktor (2010b): Abandoning the Notion of 'Translation-Inherent' Explicitation: Against a Dogma of Translation Studies. *Across Languages and Cultures*. 11(1):1-28.

Biel, Łucja (2010): Corpus-Based Studies of Legal Language for Translation Purposes: Methodological and Practical Potential. *In:* Carmen Heine and Jan Engberg, eds. *Reconceptualizing LSP. Online proceedings of the XVII European LSP Symposium 2009*. (XVII European Symposium on Languages for Specific Purposes, Aarhus, 17-21 August 2009). Aarhus: Aarhus School of Business (Aarhus University). Visited 3 November 2017, <https://pdfs.semanticscholar.org/74a8/ee84fb30d6e5593fba1fa0d1cb02cc3f9030.pdf >.

Biel, Łucja (2014): *Lost in the Eurofog: The Textual Fit of Translated Law*. Frankfurt am Main: Peter Lang.

Bisang, Walter (2017): Grammaticalization. *In*: Mark Aronoff, ed. *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press. Visited 11 November 2017, <http://dx.doi.org/10.1093/acrefore/9780199384655.013.103>.

Cao, Deborah (2007): *Translating Law*. Clevedon: Multilingual Matters.

Cappelle, Bert (2012): English is less rich in manner-of-motion verbs when translated from French. *Across Languages and Cultures*. 13(2):173-195.

Chesterman, Andrew (2004): Beyond the Particular. *In:* Anna Mauranen and Pekka Kujamäki, eds. *Translation Universals: Do They Exist?* Amsterdam/Philadelphia: John Benjamins, 33-49.

Chesterman, Andrew (2007): What is a unique item? *In:* Yves Gambier, Miriam Schlesinger, and Radegundis Stolze, eds. *Doubts and Directions in Translation Studies*. Amsterdam/Philadelphia: John Benjamins, 3-13.

Christodouloupoulos, Christos and Steedman, Mark (2015): A massively parallel corpus: the Bible in 100 languages. *Language Resources and Evaluation*. 49(2):375-395.

Dullion, Valérie (2000): Du document à l'instrument: les fonctions de la traduction des lois. *In*: *La traduction juridique: Histoire, théorie(s) et pratique*. (Colloque international "La traduction juridique: Histoire, théorie(s) et pratique," Geneva, 17-19 February 2000). Bern: Association suisse des traducteurs, terminologues et interprètes, 233-253.

Faber, Dorrit and Hjort-Pedersen, Mette (2009a): Manifestations of Interference Processes in Legal Translation. *In:* Susanne Göpferich, Arnt Lykke Jakobsen, and Inger Mees, eds. *Behind the Mind: Methods, Models and Results in Translation Process Research*. Copenhagen: Samfundslitteratur, 107-124.

Faber, Dorrit and Hjort-Pedersen, Mette (2009b): Translation Preferences in Legal Translation: Lawyers and Professional Translators Compared. *In:* Inger Mees, Fabio Alves, and Susanne Göpferich, eds. *Methodology, Technology and Innovation in Translation Process Research: A Tribute to Arnt Lykke Jakobsen*. Copenhagen: Samfundslitteratur, 339-357.

Faber, Dorrit and Hjort-Pedersen, Mette (2013): Expectancy and Professional Norms in Legal Translation: A Study of Explicitation and Implicitation Preferences. *Fachsprache*. 35(1-2):42-62.

Garzone, Giuliana (2000): Legal Translation and Functionalist Approaches: A Contradiction in Terms? *In*: *La traduction juridique*: *Histoire, théorie(s) et pratique*. (Colloque international "La traduction juridique: Histoire, théorie(s) et pratique," Geneva, 17-19 February 2000). Bern: Association suisse des traducteurs, terminologues et interprètes, 395-414.

Hjort-Pederson, Mette and Faber Dorrit (2010): Explicitation and implicitation in legal translation: A process study of trainee translators. *Meta*. 2(55):237-250.

Inhongsa, Kanokwanwalai, Louie, Meagan, and Ruangjaroon, Sugunya (2016): Is there a middle construction in Thai? *In*: *Proceedings of the International Conference on Culture, Art, Language, and Literature in the Mekong Basin (I-CALL: 2016)*. (I-CALL2016: The International Conference on Culture, Art, Language, and Literature in the Mekong Basin, Nongkhai, 20-21 July 2016). Kantharawichai: Faculty of Humanities and Social Sciences (Mahasarakham University), 206-221.

Iwasaki, Shoichi (1998): Causative and benefactive constructions in Thai. *In*: Shobhana L. Chellian and William J. D. Reuse, eds. *Papers from the Fifth Annual Meeting of the Southeast Asian Linguistics Society: 1995*. (SEALS V: Fifth Annual Meeting of the Southeast Asian Linguistics Society, Tucson, 19-21 May 1995). Tempe: Program for Southeast Asian Studies (Arizona State University), 201-210.

Iwasaki, Shoichi and Ingkaphirom, Preeya (2005): *A Reference* Grammar *of Thai*. Cambridge: Cambridge University Press.

Kawtrakul, Asanee, Suktarachan, Mukda, Varasai, Patcharee, *et al.* (2002): A State of the Art of Thai Language Resources and Thai Language behavior analysis and modeling. *In*: *Proceedings of the 3rd Workshop on Asian Language Resources and International Standardization*. (COLING-02: The 3rd Workshop on Asian Language Resources and International Standardization, Taipei, 24 August-1 September 2002). Vol. 12. Stroudsburg: Association for Computational Linguistics, 1-8.

Keenan, Edward (1985): Passive in the world's languages. *In:* Timothy Shopen, ed. *Language Typology and Syntactic Description. Clause Structure*. Vol. I. Cambridge: Cambridge University Press, 243-281.

Keenan, Edward and Dryer, Matthew (1985/2007): Passive in the world's languages. *In:* Timothy Shopen, ed. *Language Typology and Syntactic Description.* Vol. I. 2nd ed. Cambridge: Cambridge University Press, 325-361.

Kraivixien, Thanin (2012): *Thai legal language.* Bangkok: Thammasat Press.

Krogsgaard Vesterager, Anja (2017): Explicitation in legal translation – a study of Spanish-into-Danish translation of judgments. *JoSTrans.* 27(1):104-123.

Kullavanijava, Pranee (1974): *Transitive Verb in Thai.* Doctoral dissertation, unpublished. Hawaii: University of Hawaii.

Laviosa, Sara (2002): *Corpus-Based Translation Studies: Theory, Findings, Applications.* Amsterdam: Rodopi.

Loock, Rudy (2016): *La traductologie de corpus.* Villeneuve d'Ascq: Presses Universitaires du Septentrion.

Loos, Tamara (1999): *Subject Siam: Family, Law, and Colonial Modernity in Thailand.* Ithaca/London: Cornell University Press.

Olohan, Maeve (2004): *Introducing Corpora in Translation Studies.* London/New York: Routledge.

Olohan, Maeve and Baker Mona (2000): Reporting that in Translated English: Evidence for Subliminal Processes of Explicitation? *Across Languages and Cultures.* 1(2):141-158.

Palmer, David (2000/2010): Text Preprocessing. *In:* Nitin Indurkhya and Fred J. Damerau, eds. *Handbook of Natural Language Processing.* 2nd ed. Boca Raton: CRC Press, 9-30.

Pápai, Vilma (2004): Explicitation. A universal of translated text? *In:* Anna Mauranen and Pekka Kujamäki, eds. *Translation Universals: Do They Exist?* Amsterdam/Philadelphia: John Benjamins, 143-164.

Patpong, Pattama (2006): *A Systemic Functional Interpretation of Thai Grammar: An Exploration of Thai Narrative Discourse.* Doctoral dissertation, unpublished. Sydney: Macquarie University.

Prasithrathsint, Amara (1988): Change in the passive constructions in Standard Thai from 1802 to 1982. *Language Sciences.* 10(2):363-393.

Prasithrathsint, Amara (2003): A typological approach to the passive in Thai. *Manusya: Journal of Humanities.* (6):1-17.

Prasithrathsint, Amara (2006): Development of the thùuk passive marker in Thai. *In:* Werner Abraham and Larisa Leisiö, eds. *Passivization and Typology: Form and Function.* Amsterdam/Philadelphia: John Benjamins, 115-131.

Saldanha, Gabriela (2008): Explicitation Revisited: Bringing the Reader into the Picture. *transkom.* 1(1):20-35.

Sornlertlamvanich, Virach, Charoenporn, Thatsanee, and Isahara, Hitoshi (1999): ORCHID: Thai Part-of-Speech Tagged Corpus. *Journal of the Acoustical Society of Japan.* 20(3):189-198.

Sudmuk, Cholthicha (2003): The thuuk construction in Thai. *In:* Miriam Butt and Tracy Holloway King, eds. *The Proceedings of the LFG'03 Conference: University at Albany, State University of New York.* (LFG03: 8th International Lexical Functional Grammar Conference, Saratoga Springs, 16-18 July 2003). Stanford: CSLI Publication, 402-423.

Teich, Elke (2003): *Cross-linguistic Variation in System and Text. A Method for the Investigation of Translations and Comparable Texts.* Berlin: Mouton de Gruyter.

Tirkkonen-Condit, Sonja (2002): Translationese – a myth or an empirical fact? *Target.* 14(2):207-220.

Tirkkonen-Condit, Sonja (2004): Unique Items – Over- or Under-represented in Translated Language? *In:* Anna Mauranen and Pekka Kujamäki, eds. *Translation Universals: Do they Exist?* Amsterdam/Philadelphia: John Benjamins, 177-184.