The University Of Sheffield.

Translation Universals: a Usage-Based Approach

By:

Nina Szymor

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

The University of Sheffield

Faculty of Arts and Humanities

School of Languages and Cultures

25th April 2017

Abstract

The language used in translated texts is said to differ from the language used in other communicative contexts. Translation-specific linguistic behaviour (*translation* universals) has been shown to explain those differences at the levels of syntax, lexicon, discourse, and semantics. Scholars seem to disagree as to the roots of this behaviour - some turn to socio-cultural and economic factors such as risk-avoidance while others argue that cognitive processing inherent in translation and unique to it affects the linguistic choices made by translators.

The aim of this thesis is to shed new light on translation universals from a usage-based perspective. The plausibility of universal translational behaviour is assessed with reference to what we know about implicit and explicit linguistic knowledge: how it is acquired and how it affects language use. I argue that there is little support for the idea that the process of translation constrains the linguistic choices of translators. Instead, I will show that the differences between translated and non-translated texts observed in many studies, which have been attributed to translation universals, are likely to result from differences between the content of translated and non-translated components of comparable corpora. My hypothesis is supported with corpus and experimental evidence which shows that differences in the use of modality and aspect in translated and non-translated Polish texts can be explained with frequency effects: the two corpora contain different verbs whose frequency of occurrence affects translators' and authors' aspectual choices, resulting in the observed differences.

The thesis has important methodological and theoretical implications for Translation Studies. First, it shows the importance of looking at the comparability of comparable

corpora before turning to translation universals to explain the linguistic choices made in translation. Second, it casts doubt on the plausibility of translation universals as a factor in linguistic decision-making in translation and thereby simplifies the theoretical account needed to explain choices in translation.

Table of contents

List of tables

List of figures

Acknowledgements

Writing this thesis was life-changing and would not have been possible without the help and support from many wonderful people. First and foremost, I want to thank my supervisors, Dr Dagmar Divjak and Prof. Neil Bermel, for guiding me through this process with such patience, kindness, and professionalism. While Dagmar shared all her knowledge with me and always pushed me to do better, Neil helped with my doctoral development and enabled me to gain so many skills beyond what is evident from this thesis. I would also like to thank Dr Gert de Sutter from Ghent University for his support on the matters of translation theory. It has been a wonderful journey and I could not have had a better academic team supporting me.

I owe special thanks to the experimental participants and to those who helped me recruit them, particularly Dr Edyta Więcławska from University of Rzeszów and my dear friend Agnieszka Rupa.

My family and friends supported me in many different ways and I owe every one of them my deepest gratitude. In particular, I want to thank my wonderful mother Bożena who inspired me to strive for more and was there for me in all the possible ways, my sister Izabela for making her home my home every time I travelled to Sheffield for my supervision meetings, my brother Daniel whose software engineering skills made the corpus and experimental analyses possible, and my dear friend Katie for all the 'quick coffees' that turned into hours-long, you-can-do-it talks. Last but not least, I want to thank my husband Marcin - there are simply no words that would express how thankful I am to you for being the best person I could possibly have in my life. I will be forever grateful for your calmness, kindness and patience that kept me sane all these years.

Introduction

Some patterns of linguistic behaviour exhibited by translators are said to be unique to the process of translation, inherent to it, and universal. That is, these patterns are not observable in other types of verbal communication, they arise as a natural consequence of the translation process, and they occur regardless of the language pairs and type of translation involved. The past two decades have witnessed an increase in the search for these unique, inherent, and universal patterns of linguistic behaviour in translation, commonly but also controversially termed *translation universals* (Baker, 1993). These patterns are said to exist at the level of syntax, lexicon, and discourse, for example, the language of translated texts is said to be more explicit than that of source texts and non-translated texts written in the target language (e.g. Olohan, 2001, p. 424). Translators are also claimed to transform idiosyncratic features of source texts to conform to the conventions of the target language to the point of exaggerating them (e.g. Laviosa, 2002, p. 54). That is, translated texts are said to contain more standard language than a typical non-translated text. The approaches to investigating translation universals taken in these studies differ, as do the proposed motivations, reasons, and explanations as to why translators choose linguistic features differently to other language users producing texts. Thus far, no convincing proof has been offered for the existence of translation universals, bringing some researchers to the conclusion that there is simply no such thing. Halverson (2003) attempted to explain the observed patterns with reference to characteristics of human cognition and proposed the Gravitational Pull Hypothesis. She concluded that many of those patterns are most likely to be natural effects of bilingual language production, rather than evidence of a unique character of the translation process (Halverson, 2013, p. 50). House (2008, p. 11) argues that

translation universals cannot exist because translation is an act that operates on language so any behaviour observed in the translation process is a behaviour that applies to all language use.

In most cases, when differences are observed in the linguistic make up of translated texts as compared to non-translated texts, they are immediately attributed to the unique nature of the translation process. A few scholars have turned to the organisation of languages in a bilingual/multilingual mind and how it can affect the linguistic output in translation (e.g. Halverson, 2003; Lanstyák & Heltai, 2012). Although the knowledge of two languages and the constraints of bilingual communication will undeniably have influence on linguistic choices in translation, the processes that underlie all language use, including a variety of frequency effects, will arguably also be at play. Not considering the influence of such language-general processes in explaining the differences between translated and non-translated texts may have led to the mistaken conclusion that the process of translation causes these differences. For example, the increased use of optional *that* in translated texts was attributed to the translation universal of explicitation (Olohan & Baker, 2000). However, certain types of constructions have been shown to attract specific lexical items, and specific lexical items have been shown to rely on certain constructions more than other types of grammatical structures for their occurrence. For example, Schmid (2010) shows that the noun *fact* constitutes 18.45% of the nouns that occur in a 'noun + that' construction, while *idea* only constitutes 3.4%. That is, we are more likely to find *that* after the noun *fact* than the noun *idea*. Perhaps certain verbs in Olohan and Baker (2000) are more attracted to the 'verb+that' construction than other verbs, and the translated and non-translated texts contained different distributions of such verbs thus resulting in the different frequencies of occurrence. That is, the

increased use of *that* in translated texts may have nothing to do with translation, but rather with the type of verbs used in the translated and non-translated texts, occurrence of which is, in turn, determined by the contextual/communicative requirements of different texts in comparable corpora. It may be the case that no support for translation universals would have been found, had the differences in the content of comparable corpora were considered.

The aim of this thesis is to re-visit the notion of translation universals and to assess its plausibility from a usage-based perspective. We first turn to insights about language from usage-based linguistics to show that there is very little support for translation universals. We then consider an alternative explanation for the differences between translated and non-translated texts that should be considered before exploring the role of translation universals. It will be argued that translated and non-translated texts in comparable corpora are likely to differ at a very fine-grained level that can only be discovered by looking at language-general processes, particularly frequency effects. Such limited comparability of comparable corpora can provide a more psychologically plausible and verifiable explanation of differences between translated and non-translated texts, supporting the conclusion that perhaps there is little reality to universals of translation. This analysis will hopefully lead to more rigorous investigations of translated texts, and a more psychologically and linguistically realistic model of translational behaviour.

The thesis is divided into four parts. Part I introduces the theoretical background for the investigation. Chapter 1 reviews the existing literature on translation universals and assesses the notion from a usage-based perspective. It starts with the evidence for existence of translation universals from Corpus-Based Translation Studies,

followed by a discussion of the proposed reasons for universal translational behaviour, which are assessed with reference to what we know about language knowledge from usage-based linguistics. Chapter 2 discusses the usage-based approach to translation universals that is taken in this thesis - it begins with an introduction to frequency effects, which play a crucial role in the current account, and ends with an example of how such effects are likely to underlie the differences between translated and non-translated texts observed in one of the most widely cited studies on universals. Chapter 3 introduces the notions of modality and aspect, which will be used to demonstrate the importance of considering comparability of comparable corpora before attributing any unusual features of translated texts to translation universals.

Part II consists of a corpus analysis that uncovers differences between translated and non-translated legal Polish and interpretation of the results. Chapter 4 discusses the methodology and the results of the comparison of translated and non-translated texts, while Chapter 5 contains the interpretation of the results and offers a number of explanations for the differences observed in the choice of aspect in modal context - the type of situations that the two sources of data contain, frequency effects of the verbs contained in the two sources of data (analogical mapping and chunking hypothesis), and translation universals (explicitation and normalisation). It is argued that the former two explanations should be considered first, before turning to the latter.

Part III explores the chunking hypothesis formulated in Part II. First, the chunking hypothesis is tested by extracting frequency information form the National Corpus of Polish and analysing it by fitting regression to the extracted data (Chapter 6). Then,

the results of the corpus analysis are validated by engaging 45 native speakers of Polish in three experimental tasks - judgements of well-formedness, self-paced reading, and forced-choice (Chapter 7). Chapter 7 also assesses how well chunking explains the observed differences in aspectual choices made by translators and authors of non-translated texts.

Part IV discusses the implications of the current investigation for Translation Studies and Cognitive Linguistics.

PART I. THEORETICAL BACKGROUND

Chapter 1. Descriptive Translation Studies

Like any other scientific field, Translation Studies seeks to make generalisations about its object of study in order to make predictions about future or unstudied cases (Chesterman 2004, p. 33). The generalisations offered by translation scholars include prescriptive statements about the desirable and undesirable features that all translations should or should not manifest, and descriptive generalizations about the unique nature of translated language (ibid p. 34). The current study focuses on the latter type of generalisations, commonly - but also controversially - referred to as *translation universals*[1]. One of the objectives is to assess the psychological plausibility of translation universals, i.e. whether the process of translation is likely to constrain the linguistic behaviour of translators, resulting in features of translated texts that differ from comparable non-translated texts. The second objective is to assess the usefulness of the methodological approach to investigating the notion, i.e. whether comparable corpora actually tell us what we think they tell us about language processing in translation.

1.1. Laws and universals of translational behaviour

The focus on prescriptive statements about what translation should (not) be originated in the idea that translated language is inferior and represents a distorted version of its non-translated equivalent. The descriptive turn in Translation Studies shifted the focus from what translation should be to what the process of translation involves: if translated language is indeed different from non-translated language, it is worth exploring why it is the case and what underlies this unusual linguistic

[1] The term *translation universals* will be used for convenience but this does not imply that the investigation conducted in this thesis is limited to Baker's definition.

behaviour of translators (Chesterman 2010, p. 40). It was argued that the process of translation imposes certain constraints on the translator's linguistic choices, which results in linguistic patterns that are specific to translation (Toury 1991, p. 50; Baker 1993, p. 246).

Early descriptive work resulted in a number of generalisations about linguistic behaviour in translation, based on the analysis of the features of translated texts. For example, Toury (1995) proposed two laws of translational behaviour: the law of growing standardization and the law of interference. According to the former, translators have a tendency to suppress the use of original and creative language they find in the source texts in favour of using more habitual items in the translated texts. According to the latter law, translated texts exhibit interference from their originals, either as deviations from target language conventions, or as overuse of typical conventional features of target language. Blum-Kulka (1986) formulated the Explicitation Hypothesis, according to which translated texts are characterised by "an observed cohesive explicitness from [source language] to [target language] texts regardless of the increase traceable to differences between the two linguistic and textual systems involved" (Blum-Kulka 1986, p. 19). The author argues that apart from language-dependent instances of additions in the use of cohesive markers, caused by the stylistic preferences of the languages involved in translation, there is also translation-inherent explicitation, which results from the processing complexities involved in translation.

The increased popularity of corpus methods had important implications for the study of translational behaviour: it enabled translation scholars to investigate features of translated texts - and thus the nature of translational behaviour - in a quantitative

manner, without relying on intuitive assessments such as those above. In her seminal paper, Baker (1993) called for the development of corpus tools that would enable translation scholars to identify "universal features of translation, that is features which typically occur in translated text rather than original utterances and which are not the result of interference from specific linguistic systems" (Baker 1993, p. 243). Baker (1993, p. 246) argued that these features are a product of the constraints that are inherent in the process of translation, making them potentially universal and present in all translation, regardless of the languages and cultures involved. She proposed a number of different translation universals, for example, explicitation (an overall tendency to spell things out rather than leave them implicit), simplification (tendency to simplify the language used in translation), normalization (tendency to exaggerate features of the target language and to conform to its typical patterns), and levelling out (tendency of translated texts to gravitate towards the centre of the continuum) (1996, pp. 180-184).

The approaches of Baker and Toury/Blum-Kulka differ: they focus on two different types of processing in translation. By comparing the features of translated texts to the features of their source texts, Blum-Kulka and Toury are interested in the way the process of translation constrains de-coding of the source text message and re-coding it in the target language. Any patterns discovered this way are referred to as s-universals, where 's' stands for 'source' (Chesterman 2004, p. 39). Baker is interested in the way the process of translation constrains the translator's use of the target language in comparison with the use of that language by non-translators; this can be investigated by comparing the features of translated texts with the features of comparable non-translated texts. Patterns discovered this way are referred to as t-universals, where 't' stands for 'target' (Chesterman 2004, p. 8). The same proposed

universal can be investigated from both perspectives. For example, explicitation is understood as an s-universal whereby information that is implicit in the source text is encoded explicitly in the target text (e.g. increased use of cohesive markers in translation, as argued in Blum-Kulka 1986), but also as a t-universal whereby translators opt for explicit encoding of information more frequently than authors of non-translated texts, even though the target language allows them to encode that information implicitly (e.g. increased use of optional *that* in translated English texts, as argued in Olohan and Baker 2000). This difference in approaches reflects an important split that underlies investigations of the nature of translation process, which will be discussed in more detail in Section 1.1.3.

Baker's formulation of translation universals was criticised by Pym (2008) who argued that Baker simply reformulated Toury's law of growing standardization, making it into four universals that are difficult to discerns from one another. He argues that the same features are listed under explicitation and simplification (e.g. preference for finite structures), with both universals supposedly making the text easier to read. Pym points out that perhaps all of Baker's universals are really just different aspects of one underlying universal: the law of growing standardization, as formulated by Toury. Baker's failure to add the source text interference into her definition is also criticised by Pym: how do we know that the features of translated texts are independent of the influence of the source (and target) language if we do not look into the source text?

Evidence for the proposed universals has only been found for certain language pairs, certain translation directions, and certain genres, so Baker's use of the term *universals* has also been questioned. It seems to be the case that this universal

translational behaviour is not universal at all and only pertains to some types of translation. This suggests that the term 'universals' was used prematurely and that the unique features of translated language should instead be considered as "less-than-universal" tendencies or generalisations that apply to some types of translation, not to all translation universally (Chesterman 2010, p. 46).

Halverson (2003) turns to Croft's (1990) approach to linguistic explanation in order to assess Baker's universality claims and bring together the body of research into the nature of translated language. According to the approach, there are three levels of explanation that can be made about languages: (1) level of observation, (2) level of internal and cross-linguistic generalization, and (3) level of external generalization. Halverson proposes that comparisons of translated texts and their source texts (i.e. investigations of s-universals) belong at the first level of generalizations, helping us understand what happens at the level of specific language pairs. The observations made about specific language pairs can then inform second-level, cross-linguistic generalizations, made on the basis of the first-level generalizations. These second-level generalizations neutralize the role of the specific languages involved and are made by comparing translated and non-translated language. Halverson proposes that Baker's translation universals belong at this level of generalizations, which "are *explanatory* with respect to individual studies of particular linguistic realizations and/or language pairs" (ibid). The third-level generalizations explain second-level generalization by turning to language-external factors that are rooted in human psychology, biology and sociology. Halverson's proposal suggests that what Baker considers to be *universal* features of translated texts are in fact second-level generalizations made on the basis of observations from a number of language pairs and phenomena. These generalizations do not tell us what happens at the cognitive

level to evoke such linguistic behaviour and for that we need to consider the nature of human cognition. Halverson (2003, 2013, 2017) herself turns to bilingual cognition to explain certain atypical features of translated language and the generalizations about such features. She concluded that many of the supposedly translation-specific patterns are most likely to be natural effects of bilingual representation, rather than constraints of the translation process itself (Halverson 2013, p. 50).

A discussion of the psychological plausibility and cognitive reality of the claims that the process of translation constraints linguistic behaviour can be found in Section 1.2, followed by a statement of research question and methodology in Section 1.3. First, however, we briefly introduce two of the proposed universals that will be relevant at later stages of the current study: explicitation (Section 1.1.1) and normalization (Section 1.1.2). This is not to say that the methods and findings presented in this thesis apply to these two universals only; the purpose here is to assess the validity of the claim in general and the results can hopefully be extended to all posited laws or universals of translational behaviour, not only the two described below.

1.1.1. Explicitation

As mentioned above, explicitation as a cognitive process that constraints the linguistic choices of translators was already identified by Blum-Kula (1986) who defines it as an increased explicitness in translated texts in comparison with their source texts. In Baker's understanding, explicitation results in "a marked rise in the level of explicitness compared to specific source texts and to original texts in general" (Baker, 1993, p. 243) and an "overall tendency to spell things out rather

than leave them implicit in translation" (Baker 1996, p. 180). That is, Baker suggests that translators tend to encode information more explicitly than authors of non-translated texts in the same language.

It is clear from both definitions that *explicitation* in translation is understood as a cognitive process which results in increased *explicitness* in the translated text, either in comparison with the source text (as in Blum-Kulka's approach) or in comparison with a comparable non-translated text (as in Baker's approach). That is, in order to establish whether the process of explicitation took place, we need to be able to identify instances of explicitness in translated texts.

Explicitness refers to overt linguistic encoding of information, as opposed to implicitness, which refers to information that is not encoded with linguistic means but can still be inferred from what is encoded (Baumgarten, Meyer, & Özçetin 2008, p. 177). Languages differ in their conventions and preferences when it comes to explicit and implicit encoding of information; some languages require more explicitness than other languages for an utterance to be deemed acceptable (ibid, p. 178). It is therefore crucial to distinguish between instances of explicitness in the target text, which occur because of such conventions, from instances of truly translation-inherent explicitness. Klaudy (2008) argues that there are three types of language-dependent explicitness: obligatory (additions in the TT that are required by the conventions of the target language for the target text to be grammatical), optional (additions in the target text that are due to differences in text-building strategies and stylistic preferences between the target language and the source language), and pragmatic (additions in the target text that are dictated by differences between cultures). If these can be ruled out, the observed explicitness can be interpreted as

translation-inherent. Becher (2010, pp. 8-9) argues that this is problematic because it leaves one 'hoping' that after finding all of the language-dependent instances, the remaining occurrences of additions will be instances of translation-inherent explicitness. To mitigate the risk of misinterpreting explicitness as translation-inherent just because we cannot explain it with reference to the three types of language-dependent explicitness suggested by Klaudy, we can apply Krüger's (2013; 2015) model for identifying translation-inherent explicitness. According to the model, translation-inherent explicitation only takes place if we can link explicitness in the translated text to implicitness in the source text. In order to do that, we need to determine whether the explicitly encoded information in the translated text falls within the knowledge structures that underlie the source text utterance in question. To achieve that, we need to take apart the source text utterance and establish what these knowledge structures are - which domains are foregrounded and what information is salient in the communication. This makes it possible to determine whether the additions in the translated text fall within that information (Krüger 2013, p306). If they do, then we can say explicitation took place in the process of translation. If the information does not belong to the domains foregrounded by source text communication, then there is no translation-inherent explicitation.

The two definitions of explicitation mentioned at the beginning of this section have divided the pool of researchers into two groups: those who operationalize explicitation in line with Blum-Kulka's approach (e.g. Krüger 2013; Øverås 1998), and those who apply Baker's definition (e.g. Olohan & Baker 2000; Olohan 2001). There are also those who combine both approaches (e.g. Baumgarten, Meyer, & Özçetin 2008). According to these studies, explicitation is manifested through lexical, grammatical, stylistic and discourse-related features of translated texts. For

example, Krüger (2013) observed that translators fill elliptical constructions in the source texts (e.g. 'half-closed or half-open eyes' translated as 'half-closed **eyes** or half-open eyes'), insert conjunctions not present in the source text (e.g. 'long, slender' translated as 'long **and** slender'), and insert pro-adverbs in order to explicitate semantic roles of time and place (e.g. 'she crossed her legs' translated as '**then** she crossed her legs'). Olohan & Baker (2000) observed that translators are more likely to insert optional *that* than authors of comparable non-translated texts. Other examples of explicitness include distributing the meaning of a source text unit over several units in the target text, replacing nominalisations with verb phrases, disambiguating metaphors with similes, including additional explanatory remarks, use of parentheticals, and so on (Baumgarten, Meyer, & Özçetin 2008, p. 182). Explicitation results in the translated text being more wordy (longer) and including more explanatory words and more optional subordinators (Baker 1996). Although evidence for translation-inherent explicitation has indeed been found, some authors conclude that explicitation is far from a universal feature in translation.

1.1.2. Normalization

Like explicitation, normalization as a process that constraints the linguistic choices made by translators was also identified before the era of corpus-based translation studies. Toury's law of growing standardization, mentioned earlier, proposes that translators have a tendency to modify the idiosyncratic language use in the source text "to the point of being totally ignored, in favour of [more] habitual options offered by a target repertoire" (Toury 1995, p. 268). That is, translators tend to normalize unusual linguistic items present in the source texts and replace them with more conventional and less creative choices available in the target language. Baker

defines normalization as a "tendency to exaggerate features of the target language and to conform to its typical patterns" (Baker, 1996, p. 183) without referring to the source texts.

The body of corpus-based investigations into normalization is split between those that apply Toury's understanding of normalization (e.g. May 1997; Kenny 2001, Malmkjaer 1998) and those that follow Baker's definition (e.g. Mauranen 2000; Williams 2005; Dayrell 2008; Kruger and van Rooy 2012; Delaere, de Sutter and Plevoets 2012). Others combine both approaches and compare translated texts to both their sources and to comparable non-translated texts (Hansen and Teich 2001; Bernardini and Ferraresi 2011). These studies offer evidence that normalization is manifested through various lexical, grammatical and stylistic features of translated texts, resulting in the use of conventional features. For example, unusual collocations in the source text are replaced by more conventional collocations in the translated text (e.g. Øverås 1998; Malmkjær 1998); creative use of punctuation in the source text is often disregarded by the translator (May 1997; Malmkjær 1997) and so is the creative use of upper and lower case letters, non-standard spelling, or idiosyncratic hyphenation in source texts (Kenny 2001); translated texts contain more instances of recurring lexical phrases than comparable non-translated texts (Baker 2004; Dayrell 2008) and at the same time they contain fewer coined words (Williams 2005). Although many instances of normalizing behaviour have been observed in these studies, researchers often stress that normalization is also far from universal (e.g. Kenny, 2001; Delaere, De Sutter, & Plevoets, 2012; Krüger & Van Rooy, 2012).

1.1.3. S-universals *vs.* t-universals

As mentioned earlier, the two major approaches to investigating features of translated language focus on two different types of processing in translation. By comparing the features of translated texts to the features of their source texts, some scholars are interested in the constraints inherent during processing of the source text message and re-formulating it in the target language (s-universals). Other scholars investigate the way translation constrains the use of the target language by comparing the features of translated texts with the features of comparable non-translated texts (t-universals).

Some scholars argue that the constraints of the translation process can only be investigated in one and not the other way. On the one hand, Baker (1996, p. 177) argues that if the atypical features of translated texts are to be considered inherent to translation, then they will be absent (or present to a lesser degree) in non-translated texts. As such, we can only establish whether certain linguistic behaviour is universal and inherent in the process of translation by comparing translated language to non-translated language. According to Krüger (2015, p. 235), comparable texts written originally in the language of the translation have nothing to do with the process of translation, and establishing whether a translator exhibits a certain behaviour based on the characteristics of a text that he or she had no influence over is highly problematic. Another issue with not comparing translated texts to their source texts is that we cannot rule out source text interference as an explanation for the observed atypical features (Pym 2008, p. 321).

It may be the case that this split into s-universals and t-universals is artificial and the nature of translation process can only be reliably understood by analysing both types

of processing. For example, it was mentioned in the section on explicitation that language-dependent instances of explicitness have to be identified and removed first, before remaining explicitness can be attributed to translational explicitation. For that reason, we first need to establish a "base-line for the linguistic make-up expected in the target language genre" (Baumgarten, Meyer, & Özçetin 2008, p. 199) by analysing the features of comparable non-translated texts. The explicitness identified in translated texts can then be analysed with reference to such base-line. That is, the inclusion of comparable non-translated texts is crucial in investigations of explicitation (Hansen and Teich 2001; Baumgarten, Meyer, & Özçetin 2008). Once language-dependent instances of explicitness are identified, we have to ensure that the remaining instances of explicitness are translation-inherent by linking the explicitness in the translated text with implicitness in the source text; this can only be done by comparing translated texts to their sources. That is, inclusion of source texts is also crucial in investigations of explicitation. The same can be said of normalization: without comparable non-translated texts we will not know what the standard and conventional features of non-translated language are so we cannot establish whether they are exaggerated or not in the translated text, but without the source texts we cannot rule out the influence of the features of the source texts, rather than normalization, on the linguistic choices exhibited through the exaggerated features of the translated texts. It is therefore only in "a triangular set-up of source text, translation text and comparable text(s) in the target language that the phenomenon of translational explicitation [and normalization - N.S.] can be reliably identified" (Baumgarten, Meyer, & Özçetin, 2008, p. 199).

1.1.4. Explaining translational behaviour

It is clear from the above that the body of research that aims to understand the nature of translation process is far from unified. Observations made about the features of translated language are called by different names (universals, laws, generalizations, tendencies, hypotheses) and investigations are conducted in ways that do not allow for direct comparison of results so generalizations cannot be made reliably.

Whether we talk of universals, laws, hypotheses, tendencies or generalisation, the fact of the matter is that translation scholars attempt to identify certain constraints that the process of translation places on translators and their linguistic behaviour. The overall goal is to explain why certain features of translated texts seem atypical and unusual in comparison to non-translated communication. It is not the aim of the current thesis to dispute or verify any single approach to investigating the features of translated language. Instead, the interest lies in whether it is cognitively realistic to posit that the process of translation itself constrains on linguistic behaviour of translators. That is, an attempt will be made here to establish whether it is justified to explain internal and cross-linguistic generalizations (cf. Halverson 2003) by positing cognitive constraints rooted in human psychology and biology (i.e. language-external generalizations) that are inherent and unique to the process of translation.

It is by no means the first attempt at understanding what it is about translation that seemingly results in unusual linguistic behaviour. As mentioned earlier, Halverson (2003) pointed out that although Baker's universals can explain observations about certain patterns in specific language pairs, they are themselves in need of an explanation. That is, universals are a higher-level generalizations about certain constraints of the translation process, but they do not explain why these constraints

happen in the first place. Some explanations as to why translators are constrained in their linguistic choices have been offered; these can be broadly divided into two groups: socially-oriented and cognitively-oriented (Halverson, 2017, p. 10). The former rely on socio-cultural and economic factors in order to understand what it is about the process of translation that constraints the way translators use the target language. Such factors include translator's understanding of his/her role, the requirements of the audience, the status of the languages involved in translation, work conditions (e.g. payment), and so on. For example, Øverås (1998, p. 569) explains explicitation observed in her study as a strategy applied by conscientious translators who want to reduce the loss of information in the transfer between source text and target text. Similarly, Pym (2008) argues that translators tend to avoid the risk of producing texts that are unacceptable for their audiences by either standardizing the language use or by channelling interference from the source text. This risk aversion, according to the author, produces 'a deceptively universal behavioral disposition' (ibid., p. 326). Mauranen (2008) suggests that "translations are supposed to avoid margins or periphery and remain safely within the mainstream" (Mauranen 2008, p. 40) which influences the tendency to conform to the conventions of the target language, to the point of "exaggerating" them. Kenny (2001, p. 67) points out that translations that deviate from the accepted linguistic norms and conventions may be criticized and rejected by the target audience, so normalisation may be influenced by trying to avoid such criticism. Moreover, it may even be the case that translators' manuscripts undergo editing processes imposed by publishing houses, and it may be those editing processes that impose the more conventional and standard variety of language (Delaere, de Sutter, & Plevoets 2012, p. 221). We will return to this point in section 1.2.2.

Other scholars turn to cognitive processing in order to understand what it could be about the process of translation that constrains linguistic behaviour. For example, Chesterman (2004, p. 44) argues that since translation takes place in the mind of the translator under certain causal constraints, we should look for explanations in cognitive processing. Olohan (2001, p. 427) proposes that the patterning observed in her study, i.e. the tendency *not* to omit optional syntactic elements, may result from the cognitive processing that the process of translation requires. Neither of the authors suggest what type of cognitive processes they have in mind and in what way they are inherent in translation. Faber & Hjort-Pedersen (2009, p. 108) make a suggestion in relation to linguistic explicitation observed in target texts: they propose that it results from the mental comprehension process that translators go through. When translating a text, translators will explicitate certain information in their minds in order to comprehend the information contained in the source text, and this will lead them to subconsciously explicitate this information linguistically in the translated text.

Proponents of the abovementioned explanations do not seem to provide much evidence to support their claims and it has been suggested that perhaps translators' linguistic behaviour can be explained by exploring cognitive mechanisms that underlie language in general (House, 2008; de Sutter & Velde, 2008), mechanisms more specific to bilingual communication (Halverson, 2013), or the nature of all constrained (Lanstyák & Heltai, 2012) or mediated (Ulrych & Murphy, 2008) communication, not only translation. Lanstyák & Heltai (2012) draw parallels between the proposed generalizations about translated language and characteristics of bilingual communication: they argue that the supposedly unique features of translated texts are not in fact specific to the process of translation but are a result of

communication taking place under various cognitive, social, or linguistic constraints. House (2008, p. 11) claims that translation universals cannot exist because translation is an act that operates on language so any behaviour observed in the translation process is a behaviour that applies to all language use. For example, De Sutter & van de Velde (2008, pp. 14-15) suggest that the differences they observed in their study could have resulted not from normalisation but from strategies translators - just like other language users - have for dealing with linguistic alternatives. Finally, Halverson (2003) attempted to explain certain differences with reference to the characteristics of human cognition and proposed the so-called Gravitational Pull Hypothesis. She concluded that many of the patterns that were proposed to be unique to translation are most likely to be natural effects of bilingual production, rather than the nature of the translation process (Halverson 2013, p. 50).

1.2. Can the process of translation constrain linguistic behaviour?

In order to explore whether the process of translation can constrain linguistic behaviour of translators, resulting in law-like tendencies, we will turn to insights about implicit and explicit linguistic knowledge and conscious and unconscious cognitive processing (Section 1.2.1). We will also explore the idea that comparable corpora used to investigate translational behaviour may not necessarily be well suited to do so (Section 1.2.2).

1.2.1. Linguistic knowledge and cognitive processing in translation

The process of translation involves 'changing an original written text [...] in the original verbal language [...] into a written text [...] in a different verbal language' (Munday, 2001, p. 5). That is, the process of translation involves comprehending content in one language and verbalising it in another language. On a deeper level,

that involves a variety of mechanisms and decisions that result from two types of linguistic knowledge a speaker has - implicit and explicit.

Implicit linguistic knowledge is the organisation of language in the speaker's mind, resulting from the acquisition of one's mother tongue. We know that this knowledge is implicit because by the age of five, a child acquires the main structures of her/his language without consciously knowing that there is such thing as grammar (Dienes, 2012, p. 337). This type of knowledge is acquired and operates independently of the speaker's awareness and therefore belongs to the mode of *unconscious cognitive processing* (Evans, 2008; Reber, Allen, & Reber, 1999). There are various theories about the way this knowledge is acquired and structured. In the current thesis, we take the usage-based cognitive-linguistic view of implicit linguistic knowledge, according to which, language is shaped by usage, and language acquisition and use are facilitated by a number of general cognitive abilities, such as abstraction, categorization, chunking, and entrenchment. These abilities are general to all cognitive systems, not only language, which means that the idea of language universals has no place in this usage-based approach. More on this approach to language in Chapter 2.

Explicit linguistic knowledge, on the other hand, is the conscious awareness of grammatical rules and metalinguistic descriptions of linguistic constructions, acquired as a result of language instruction (Dienes, 2012, p337; Roehr-Brackin, 2015, p125). This explicit knowledge covers such areas as syntax, semantics, morphology, phonology and so on. Conscious awareness of rules enables explicit reasoning and decision-making, which belong to the mode of *conscious cognitive processing* (Evans, 2008). Conscious cognitive processing is subject to individual

differences conditioned by general intelligence, working memory capacity, cognitive ability, dispositions for critical and reflective thinking, and even cross-cultural differences in thinking styles (Evans, 2008; Reber, Allen, & Reber, 1999). The conscious language choices made by speakers will therefore largely differ on an individual basis. Although implicit linguistic knowledge can also be subject to individual differences (cf. Dąbrowska, 2015), it is caused by differences in linguistic experience of individual speakers, rather than general intelligence and their conscious knowledge of language. It means that speakers that have been exposed to different language varieties and environments will exhibit differences in the way their implicit language knowledge is structured and subsequently, the way they speak.

The two types of linguistic knowledge are separate, but they are intertwined: implicit knowledge is by default relied upon in language comprehension and production, but explicit knowledge may come into play when difficulties arise and the speaker needs to make conscious effort to analyse the input and control the output (Roehr-Brackin, 2015, pp. 118-119). It is argued here that these two types of knowledge also interact during translation, resulting in two types of linguistic behaviour: universal (in the usage-based sense) behaviour guided by the translator's implicit linguistic knowledge, and translation-specific behaviour guided by explicit linguistic knowledge. I argue that the universal behaviour exhibited by translators cannot be translation-specific, and the translation-specific behaviour exhibited by translators cannot be considered universal. That is, I propose that the universal or general tendencies in linguistic behaviour of translators will not result from the constraints of the translation process but rather from the constraints that bilingualism and human cognition place on the said linguistic behaviour (more details in Section 1.2.1.1). I

also propose that translation-specific linguistic behaviour, such as selecting features based on risk-avoidance, are not universal and cannot be generalised due to the reliance on explicit linguistic knowledge and conscious cognitive processing of such decision-making, which are very individual in nature (more details in Section 1.2.1.2).

1.2.1.1. Universal but not translation-specific

Like other speakers, translators are first and foremost guided by their implicit linguistic knowledge of the two languages. From work on bilingualism we know that a bilingual's two languages are not completely separate structural entities: a bilingual's use of either language will differ from the use of each of those languages by monolingual speakers (Cook, 2003). For example, bilingual speaker's mother tongue words are activated during second language processing (Sunderman & Kroll, 2006) and performance in a native context is influenced by the speaker's knowledge of a second language (Hell & Dijkstra, 2002). When speaking one language, bilinguals exhibit mechanisms and processes that monolinguals simply do not have, such as translation and code-switching (Bassetti & Cook, 2011). All of this suggests that the comprehension of the source text during translation is likely to be influenced by the knowledge of the target language, and the production of target text is likely to be influenced by the source language. This may result in linguistic output produced by bilingual/multilingual speakers (including translators) that differs from the output of monolingual speakers. That is, it may seem that translators' linguistic behaviour differs from the behaviour of non-translators, but it is likely that such behaviour would also be exhibited by other bilingual speakers, rather than it being influenced by the act of translation. This bilingual/multilingual linguistic behaviour may differ

from monolingual behaviour but it is not translation-specific because it arises from the commonalities of the bilingual/multilingual mind. The position taken here is that bilingual/multilingual knowledge is likely to underlie the differences observed in comparisons of translated and non-translated texts: translators are bilingual/multilingual and as such they use all their languages differently to monolingual speakers of those languages, hence the observed differences. We can therefore make language-external generalizations about these internal and cross-linguistic generalizations that t-universals are. However, we cannot say that these generalization arise from the unique nature of the translation process and as such cannot be considered translation-specific. Instead, they are characteristic of all production and comprehension by multilingual speakers. This is indeed what Halverson (2013) and Lanstyák & Heltai (2012) suggest - the unique patterns of linguistic behaviour in translation result from the constraints of the bilingual linguistic organisation in the speaker's mind, not from translation-specific mechanisms or processes.

1.2.1.2. Translation-specific but not universal

Certain constraints on linguistic behaviour in translation may arise when linguistic input is causing processing difficulties and translators resort to conscious problem-solving by relying on explicit linguistic knowledge. For example, when translators come across problematic implicitness or idiosyncratic language use in the source text (which does not necessarily mean the same thing for different translators; what is problematic for one may not be for another), they turn to their explicit linguistic knowledge, educational and professional experience, and past experiences of dealing with similar problems. Solutions may be informed by a number of factors: the type

and purpose of the translation, the intended readership, the differences between the languages and cultures involved in translation, the translator's perception of her/his role, her/his assumptions about the audience, risk-avoidance, and so on. For example, depending on the translator's assumptions about the educational background of the audience, the choice of linguistic alternatives may vary (Newmark, 1988, p. 15). Similarly, depending on the type and function of the translated text, the translator may have to apply different linguistic conventions in the target language (Nord, 1997, p. 38).

These linguistic decisions are conscious and as such are subject to individual differences associated with conscious cognitive processing (as described in the introduction to Section 1.2.1), but also depend in large part on the knowledge acquired by translators in the course of professional training and professional experience, which is also subject to individual differences. Translation literature abounds with text typologies, translation typologies, strategies and techniques for dealing with problems related to the different text types and functions, advice on what the translator should and should not do, and so on. Different recommendations represent different approaches to what translation is and how it should be performed, and will therefore results in individually motivated linguistic choices made by translators. Some of these strategies and techniques may have been applied by translators so many times that they have become automated, and therefore belong to the mode of unconscious cognitive processing in the form of 'if X then Y' strategies, e.g. if translating a legal text from English into Polish for instrumental purposes, always use impersonal constructions. Such strategies would not need to be explicitly applied as a result of conscious decision-making, but would constitute quick,

intuitive and automated responses to particular circumstances. However, they would - again - depend on all the individual factors identified above.

In both cases - whether conscious and controlled or unconscious and automated - the solutions applied in these situations are interpreted by translation scholars as tendencies, leading to certain generalizations about the process of translation. However, the position taken in the current thesis is that such solutions applied in cases of processing difficulties, and generalizations made on their basis, would reflect the prescriptive approaches to translation that are taught to translators during vocational training, not the constraints of the process of translation itself. Such position is supported by the fact that these tendencies have only been observed in some language pairs, directions and genres. Even within the same study, conclusions have been made that normalization and explicitation are not universally applied in all cases of implicitness or idiosyncratic language use. This may reflect individual differences in what causes processing difficulties and in the solutions applied when dealing with them, resulting from different educational background and theoretical approaches, as well as other differences relating to explicit linguistic knowledge and conscious cognitive processing, as previously mentioned.

Though translation-specific, it would be misleading to regard such constraints as universal or law-like, and to make generalizations about them; any generalizations would only result in re-formulations of prescriptive approaches that translators learn and apply in professional work. I propose that this is what underlies a number of s-universals identified in the literature: translators come across processing difficulties and solve them with reference to their translation-specific but inevitably individual in nature explicit knowledge. In other words, it is argued here that what various

researchers call s-universals are in fact translation-specific responses to processing difficulties, which are only generalizable to the extent to which prescriptive approaches to translation can be generalized. We are therefore unable to make language-external generalizations about s-universals, suggesting that although translation-specific, they are not law-like or universal.

1.2.1.3. Summary

To sum up, there seems to be little support from cognitive point of view for the idea of universal or law-like linguistic behaviour in translation that results from the constraints of the process of translation itself. Linguistic behaviour in translation can be constrained by the nature of bilingualism, which is universal to all bilingual speakers rather than translation-specific and will result in language use in bilingual/multilingual that differs from the language use of monolingual speakers (as explained in Section 1.2.1.1). Linguistic behaviour in translation can also be constrained by processing difficulties encountered during translation which are solved by relying on individual translation-specific approaches, resulting in translated texts whose features do not fully correspond with the features of their sources (as explained in Section 1.2.1.2). Such translation-specific approaches cannot and should not be generalized into law-like or universal statements about linguistic behaviour because they reflect the strategies and techniques for dealing with problematic input acquired during training and professional experience.

It will be shown in the current thesis that the nature of bilingualism may not be the only reason why comparisons of translated and non-translated language result in differences being observed. Comparable corpora used in such investigations may be the main culprit and translation scholars should re-consider whether such corpora are

well-suited for investigating the nature of translated language in comparison with non-translated language. Comparable corpora and the texts contained in them are inevitably different, which may lead to differences being observed which have nothing to do with the process of translation and its constraints but are mistakenly interpreted as such. This is discussed in more detail in the next section.

1.2.2. Comparable corpora: do they tell us what we think they tell us?

The emergence of Corpus-Based Translation Studies can be attributed to Baker's (1993) seminal paper in which the author called for the application of corpus techniques in translation studies. It was argued that language corpora will enable researchers to discover the nature of translated language, and universal features of translation in particular. There are many advantages to corpus methods, the main being the naturalness of the data which prevents reliance on intuition and introspective analyses that can be abused and manipulated to suit the researcher's needs (Pullum, 2007, pp. 38-39). The quantitative nature of corpus analyses allows for a consistent and objective study of large number of examples, leading to the formation of objective generalisations (Biber & Jones, 2009, p. 1287). Corpus analyses within Translation Studies can involve various types of corpora, most notably parallel and comparable corpora. The former consist of collections of translated texts in one or more languages together with their source texts, while the former consist of texts translated into a given language and texts written originally in that language but which are comparable to the translated texts in terms of genre, topic, time span and communicative function (Laviosa, 2002, p. 36). Parallel corpora are typically used for investigations of s-universals while comparable corpora are

used for t-universals. More recently, both types of corpora are being used jointly in a triangular set-up, as discussed in Section 1.1.

The use of comparable and parallel corpora for investigations of translated language has been questioned. Doubts relate to the assumed but questionable comparability of comparable corpora (e.g. Laviosa, 1997; Bernardini and Zanettin, 2004; Bernardini and Ferraresi, 2011), to the criteria for inclusion of translated texts in a corpus (e.g. Chesterman 2010), and to the influence of various verification and editing processes on the features of translated texts, and subsequently, on the results of comparisons of these texts to the supposedly comparable non-translated texts (e.g. Kruger, 2012; Delaere, De Sutter, & Plevoets, 2012).

As mentioned above, comparable corpora are matched for a number of characteristics, which is generally assumed to ensure comparability of the translated and non-translated component. However, it was shown that the type of texts that are generally selected for transaltion into and from various languages can have significant consequences for the features of translated language (Bernardini and Zanettin, 2004; Bernardini and Ferraresi, 2011). For instance, source texts selected for translation into a given target language may contain more formal language than comparable texts written originally in the target language, making the comparable coprpora less comparable than is typically assumed (Bernardini and Ferraresi, 2011, p. 228). If the features of the source texts are not looked at, the analysis of such translations can lead to the mistaken conclusion that translated language is more formal than non-translated language. One of the solutions for the limited comparability of comparable corpora is therefore to include source texts in the analysis (ibid.). Others turn to register-controlled corpora which supposedly ensure

closer comparability (e.g. Teich 2003). It will be shown in this thesis that such steps do not fully mitigate the limited comparability: comparable corpora differ in the type of situations they describe and in the linguistic experience of the authors and translators, both of which significantly affect the way language is used by individual speakers. This may lead to differences being observed, which if not analysed carefully enough, may be interpreted as evidence for translation-specific constraints on the linguistic behaviour of translators. It may therefore be the case that no matter how closely matched for genre, time span, topic and communicative function, comparable corpora will never be comparable enough to warrant any reliable generalizations about differences between translated and non-translated language.

Additional limitations of language corpora relate to the criteria for including translated texts. There are various decisions to be taken: do we only include translations done by native speakers working into their mother tongue or translations out of one's mother tongue too, only translations done by professionals with training and experience or amateur translations too, only "good" translations (however that is measured) or less than ideal ones too, and so on (Chesterman, 2010, p. 44). There is also the issue of translations done with the help of translation memories, which exhibit an increased level of syntactic interference from source texts (Pym, 2008, p. 323). This is particularly important when comparable corpora are used and the researcher has no access to the source texts: there is no way of ensuring that the atypical features of translated texts do not result from the interference from the source text rather than translation universals. Moreover, in many languages and countries some text types, genres or lexical innovations are heavily influenced by translations from other languages (Mauranen, 2008, p. 35) or texts written in other languages (Delaere, De Sutter, & Plevoets, 2012, p. 220), which makes comparisons

of translated and non-translated language futile. Last but not least, it has been pointed out that translated texts undergo verification and editing procedures not necessarily performed by the original translator which means that the final product may not reflect to a full extent the cognitive processes and language use of the translator, but language norms and expectations of the editors or publishing houses (cf. Kruger, 2012; Delaere, De Sutter, & Plevoets, 2012, p. 221).

All of these issues raise important questions about the validity of language-internal and corss-linguistic generalizations (cf. Halverson 2003) made on the basis of comparisons between translated and non-translated language. It may be recalled that such generalizations are an important step in understanding the nature of translated language and the translation process: although they can help us understand observations made on the basis of language-specific pairs, language-internal and corss-linguistic generalizations are themselves in need of explaining (Halverson, 2003, p. 231). That is, in order to truly understand the nature of translation and how it constrains linguistic choices, we have to explain language-internal and cross-linguistic generalizations with reference to language-external factors, such as human psychology and biology (ibid). However, if language-external generalizations are to be reliable and valid, we need to ensure that the methodology used to arrive at the language-internal generalizations is reliable and valid. As such, we need to ensure that comparable corpora really do tell us what we think they tell us. However, all of the above suggests that it may well be the case that comparable corpora are not as helpful in discovering the patterns and nature of translated language as is assumed among translation scholars. It will be demonstated here that corpora lack comparability at very fine-grained levels of linguistic analysis, casting even more doubt on the usefulness of corpus methods in translation.

1.3. Aims and objectives

The aims of the current thesis are two-fold: to assess (1) whether there is likely to be translation-specific linguistic behaviour governed by the constraints of the translation process, and (2) whether comparable corpora actually tell us what we think they tell us about language processing in translation. The former question was to a theoretical extent answered in Section 1.2. In what follows, we explore the usefulness of comparable corpora as a tool for investigating the nature of translated language. Section 1.3.1 discusses the research question in more details while Section 1.3.2 introduces the methodology.

1.3.1. Research question

It was established in the previous sections that there is little likelihood that translation-specific linguistic behaviour can be explained with reference to language-external factors and as such we cannot talk of law-like or universal tendencies of linguistic behaviour in translation. It was proposed that if linguistic behaviour in translation differs from non-translated linguistic behaviour, as observed in many corpus studies, there are three potential explanations: (1) the effects of a bilingual/multilingual mind, which will hold universally for all bilingual/multilingual speakers and will therefore not be translation-specific (see Section 1.2.1.1 for more details); (2) the effects of educational and professional experience of translators, which will be translation-specific but cannot be considered universal or law-like (see Section 1.2.1.2 for more details); and (3) the effects of corpus methods applied in the studies of such linguistic behaviour, particularly the limited comparability of comparable corpora (see Section 1.2.2 for more details).

Comparability of comparable corpora is crucial to ensuring that any attempts at understanding and explaining the tendencies observed in such corpora are reliable and valid. That is, if we want to make sure that the language-external explanations (cf. Halverson 2003) for atypical linguistic behaviour of translators observed in comparable corpora are psychologically and biologically plausible, we first need to ensure that the atypical features observed do not result from the limited comparability of corpora. The current thesis will therefore explore the third explanation in more detail. If we show that differences observed in comparisons of translated and non-translated corpora can be explained by looking at very fine-grained differences in the content of the corpora, not considered previously, then this will cast doubt on the validity of all previous work conducted with the use of such corpora, and consequently on the evidence that supposedly supports the idea of translation universals. In other words, the current thesis will explore the differences between translated and non-translated corpora in a way that has not been done before in order to answer the question of whether these differences in content, not the constraints of the translation process, underlie the apparently different linguistic behaviour of translators and non-translators. The next section explains how this will be achieved.

1.3.2. Method

In order to investigate the research question, we will compare translated and non-translated Polish texts of legal nature. More specifically, we will compare the use of modal verbs in legal texts written originally in Polish and legal texts translated into Polish from English. We will apply the usage-based approach to lexical semantics called Behavioral Profiling (Divjak, 2004; Divjak & Gries, 2006), which will enable

us to compare at a very fine-grained level the linguistic features of utterances containing modal verbs, randomly selected from the translated and non-translated corpora. It will be shown that translated and non-translated texts differ in the distribution of verbal aspect. These differences could be attributed to either explicitation or normalization but we will first analyse them by considering the content of the translated and non-translated corpora. We will be interested in whether the atypical aspectual choices in translated utterances differ from the typical choices in the non-translated utterances because the translated corpus contains texts that differ in content and therefore the distribution of aspectual forms.

We will consider two factors that are likely to affect aspectual choices but do not rely on the speakers' preferences: the communicative context (i.e. situation types) and implicit linguistic knowledge that affects language use. As mentioned earlier, we will take the usage-based approach to implicit linguistic knowledge, and the main mechanism looked at will be frequency effects. For example, increased explicitness in translated texts as compared to non-translated texts may result from the fact that the linguistic items in the translated texts are more frequently encoded explicitly. Such items would be more deeply entrenched in the translator's memory in that explicit form and would therefore be activated more quickly and selected in production. At the same time, the non-translated corpus may contain linguistic items that are more frequently encoded implicitly and therefore the implicit encoding is selected. This would result in more explicitness in the translated texts than in the non-translated texts, but would have nothing to do with the constraints of the process of translation but rather with the differences in the content of the two corpora.

Only if the role of the three above-mentioned factors can be ruled out, can we start looking at language-external explanations for the differences. That is, once we rule out the role of the source language, the individual in nature strategies applied by translators to deal with processing difficulties, and differences in the content of translated and non-translated corpora, we can start making language-external generalizations about the atypical linguistic behaviour of translators, and to establish whether such generalizations are unique to translation. However, if any of the three factors can be said to underlie the observed differences, it would be misleading to posit the existence of law-like tendencies or universals of translation.

Part II discusses the comparison of translated and non-translated texts (Chapter 4) and the interpretation of the results (Chapter 5). Various explanations for the observed differences will be proposed, including translation universals, frequency effects, and situation types, with the latter shown to explain the atypical aspectual choices of translators to a large extent. The psychological plausibility of the frequency effects proposed to explain the remaining differences will then be assessed in Part III.

1.3.2.1. Modality in legal language

One of the reasons for investigating legal language in the current thesis is to ensure closer match between translated and non-translated texts (cf. Teich 2003). However, there are other reasons why looking at legal texts, particularly at modality in legal texts, is beneficial.

Legal texts are a promising test bed for translation universals - legal communication is likely to include instances of implicitness, vagueness, and idiosyncratic features, all of which have been argued to cause the unique linguistic behaviour witnessed in

translation. For example, legal texts are characterised by ambiguity and vagueness - "[s]ignificant portions of the institutional legal system, especially courts at the appellate level and supreme courts, are for the most part concerned not with disentangling the facts of cases but with the indeterminacies of the law" (Poscher, 2015, p. 128). Expressions and syntactic structures can be considered ambiguous if they have multiple meanings. For example, a modal verb will be ambiguous because it has more than one meaning - the epistemic and the root meaning. Although ambiguity can usually be disambiguated by looking at the context, vagueness occurs when it is impossible to determine the intended meaning. As mentioned in Section 1.1.4, Faber and Hjort-Pedersen (2009: 108) suggest that translators explicitate certain information in order to comprehend the meaning of the source text utterance - this will be particularly true of legal texts, which are linguistically complex and contain the abovementioned ambiguity and vagueness. Moreover, an investigation of legal translations (particularly EU texts) can also prove fruitful, if normalisation exists. Each culture has its own legal system, and each legal system has its own legal language, with an individual apparatus, conceptual structure, sources of law, and principles (Šarčević, 1997). The two most common legal families are the Common Law and the Civil Law, and all of the Western national legal systems belong to either of those (Merryman, 1981, p. 358). The two families differ substantially in terms of concepts, institutions and domains of law. The European Union is becoming a mixed jurisdiction, with two legal traditions converging – the Civil Law of the continental countries and Common Law of England, Wales and Ireland (Tetley, 2003, p. 24). Legal translators are therefore required to not only have excellent linguistic skills, but also solid knowledge of the two legal systems involved in translation (Groot G.-R. d., 1987, p. 797). The texts analysed in this thesis have all been translated from

English, one of the drafting languages of the EU, and the *lingua franca* of Common Law jurisdictions. Concerns have been raised about the English language being used for drafting purposes - some claim that it will inevitably impose Common Law concepts on Civil Law jurisdictions (Tetley, 2003, p. 31), resulting in 'loose and sloppy assumptions' about legal concepts, if not enough attention is paid to the structural differences between the legal systems (Rossini, 1998, p. XXI). Legal concepts expressed in English may be understood differently by lawyers belonging to different legal traditions. For example, the word 'contract' in the mind of a Common Law lawyer means something radically different than its apparently equivalent French concept 'contrat' (Kjaer, 2004, p. 386). According to the European Commission, however, 'English as a drafting language of the EU texts is […] distinct from the common law systems and from the cultural constraints of countries using English as a national language' (European Union, 2010, p. 89). Nevertheless, it is easy to see how translating legal texts might result in the normalising behaviour observed in many studies and discussed in Section 1.1.2. It was shown that idiosyncratic features of the source text are likely to be normalised during translation, resulting in less creative use of language. In case of translation of EU instruments, concepts that have specific meaning in the Common Law systems, or neutral concepts which are specific to the EU but unknown in the Civil Law systems of the member states, may become normalised.

Modality plays an important role in legal communication - modal verbs are the most coherent group of markers that express meanings of obligation, permission, and prohibition, and are therefore crucial to normative texts, investigated here. Moreover, modality is a good test case for normalisation. This is because the way modal markers are used in legal English and Polish differs: Matulewska and Gortych (2009)

show that *must* is the most frequent item after *shall* in English legal texts (2009, p71). In Polish legal texts, however, *musieć* is just one of many other, much more frequent expressions (ibid, p69). If the studies on normalisation (Section 1.1.2) are right, the frequent use of *must* in English should be normalised in a Polish translation, and a more conventional expression in Polish will be overrepresented to avoid mistakenly using *musieć*, a formal but obviously not a functional equivalent of *must*. Or instead, the formal equivalent of *must* in Polish (i.e. *musieć*) will be overrepresented, suggesting that shining through took place. In either case, the distribution of other necessity modals (*należy, powinien, trzeba*) will also be affected by this under- or overrepresentation of *musieć.* Modality is also likely to have its semantic structure influenced by the process of translation. Vandevoorde (2016) found evidence of levelling out, normalisation, and shining through in the semantic field of Dutch inchoativity. This suggests that apart from affecting lexical, syntactic and discourse features of texts, the translation process can also affect the semantics. The semantic field of modality has a complex structure. The two basic senses - possibility and necessity - can be expressed in Polish by seven modal verbs. Possibility is expressed with *móc, można, wolno* and necessity is expressed with *musieć, powinien, należy* and *trzeba.* Each of these modals have at least three meanings, e.g. *musieć* has deontic, dynamic, and epistemic meaning. Some of these related meanings can be more central than others. For example, *musieć* can be considered a more typical expression of necessity in comparison to *powinien* - the latter occurs in deontic and dynamic contexts (expressing necessity) but also in epistemic contexts (expressing probability), while the former occurs in deontic and dynamic contexts, expressing necessity only (Divjak, Szymor i Socha-Michalik, 2015). Each individual modal is also linked to other lexical items that synonymously

express the same senses. For example, *musieć* expresses the necessity senses of deontic and dynamic modality, and *powinien, należy,* and *trzeba* express those senses too. Even though superficially synonymous, however, the items are likely to construe the same concept from different perspectives (Divjak & Gries, 2006). Szymor (2015) has shown that *musieć* and *powinien*, even though considered synonymous in legal language, are in fact used in varying sentential contexts - *musieć,* for example, is more likely to be chosen for abstract subjects, while *powinien* is more likely to be used with concrete ones. The semantic field of modality is obviously a very complex one. If the process of translation does affect the semantic structures of lexical items then we are likely to observe some changes in this complex semantic network.

To sum up, by investigating modal verbs in legal texts, we are likely to come across various instantiations of what may be interpreted as normalisation and explicitation. This will in turn allow us to analyse whether we can indeed attribute those instantiations to the posited translation universals or whether the differences between the content of the corpora used in the investigation are the more likely reason behind any observed differences.

1.3.2.2. Usage-based approach

It was mentioned previously that the current thesis takes the usage-based cognitive-linguistic approach to language. This approach to language is based on what is known about the human mind and brain from other disciplines, such as Cognitive Sciences. It is therefore more psychologically plausible and cognitively realistic than other approaches. For this reason, this view of language underlies the analyses conducted in the current thesis.

Broadly speaking, there are two main approaches to how language is acquired and how it is organised in the speaker's mind. According to the generativist approach, even though languages differ on the surface, beneath the surface they are all guided by a common language blueprint, i.e. universal principles that are innate and specific to language. This notion seems to be accepted as fact and pervades a lot of work done in various branches of linguistics and cognitive sciences (Evans & Levinson, 2009), p429). This idea has been opposed by cognitive linguists and various cognitive scientists and psychologists (for an overview, I refer the reader to Evans and Levinson 2009). Evans and Levinson (2009, p429) show that languages differ on so many levels that it is very difficult to find any property that they would all truly share. Moreover, they argue that the universality claims made by Chomskyan Universal Grammar are "either false, unfalsifiable, or misleading in that they refer to tendencies rather than strict universals" (ibid).

Cognitive linguists argue that instead of a universal and innate language blueprint, all humans share cognitive, neuro-anatomical, and environmental constraints, which influence what can be expressed in language, resulting in patterns that are shared by majority of languages (Evans & Green, 2006, pp. 63-64). In other words, the common patterns identified across languages will not result from some universal language-inherent blueprint with which all humans are born, but rather from universal cognitive abilities and mechanisms shared by all humans, which are not exclusive to language itself, but which enable us to learn languages, and other skills, from exposure to them. Three main assumptions underlie usage-based cognitive linguistics: language is shaped by usage, it is rooted in general cognitive abilities, and all of its elements are meaningful, including grammar (Dąbrowska & Divjak, 2015, p. 1). Frequency of occurrence plays an important role in a linguistic system

shaped by usage (Barlow & Kemmer, 2000, p. x) and will be central to the usage-based approach to translation universals advocated in the current study. This is discussed in more detail in Chapter 2.

Chapter 2. Frequency in language

Effects of frequency of occurrence are prevalent at all levels of linguistic analysis and should therefore be taken into account when trying to explain linguistic behaviour, including linguistic behaviour in translation. Generally speaking, the more frequently a linguistic item - a phoneme, a word, a phrase - is encountered by a speaker, the more entrenched it becomes in the speaker's memory, and the easier it is to subsequently retrieve and process that item. Frequency effects are not exclusive to language - frequency seems to be relevant whenever humans "perceive, reason, think, judge, and make decisions" (Sedlmeier, Betsch, & Renkewitz, 2002, p. 3). For example, Aarts and Dijksterhuis (2000, p. 55) argue that goal-directed actions become habitual if they are frequently performed in similar situations. That is, the more often a given action is performed to achieve a given goal, the more associated the two become, leading to the formation of a habit and to automatic activation of actions associated with a given goal. The actions required to achieve a goal that is pursued regularly will be performed habitually and without attention paid to the process. As a result, the frequency with which a given situation was encountered in the past enables us to predict future choices and behaviours in similar situations. Similar frequency effects to those observed in psychological research have been encountered in various aspects of language acquisition, use, and change. Frequency of occurrence and repetition is therefore an important aspect of human behaviour that needs to be taken into account whenever we try to understand why humans speak the way they do in all circumstances, including translation.

2.1. Frequency effects

In language acquisition research, frequency with which a child is exposed to a given word/phrase was shown to correlate with the age of acquisition of that word/phrase. That is, such words as *daddy, mommy, bye, hi,* etc., will be produced before forms that are encountered less frequently, such as *coffee* or *computer* (Ambridge et al., 2015, p. 243). The order in which children acquire certain syntactic structures is also associated with frequency. For example, the order of acquisition of individual verbs in the verb-object, subject-verb-object, and intransitive structures, is mainly influenced by how often children hear the individual verbs, but also by how often those verbs are used by mothers in the different syntactic structures (Theakston et al., 2004). In other words, the frequency with which a child encounters a given verb in a given syntactic structure will influence the order in which the child starts using that verb in that syntactic structure. Frequency of occurrence can prevent or reduce errors produced by children. For example, children are less likely to over-regularize irregular forms, e.g. to say **blowed* instead of *blew* or **foots* instead of *feet*, that they hear often than irregular forms that they hear less often (Maslen et al., 2004, p. 1325). Effects of frequency of occurrence can also be observed in the acquisition of word order. Matthews et al. (2005) observed that 2-year-old English-speaking children are more likely to repeat the experimenter's ungrammatical subject-object-verb order if the verb has low frequency. If the verb has high frequency, however, 2-year-olds are more likely to correct the experimenter's ungrammatical word order to the grammatical subject-verb-object. Ambridge et al. (2015) argue that based on evidence from a number of child language acquisition studies, any successful account of child language acquisition must take frequency effects into consideration.

In adult speakers, frequency effects are also pervasive - entrenchment of frequently encountered items makes them more easily accessible (Bybee, 2007, p. 10; p. 324). This facilitates faster, easier and more accurate processing, as well as resistance to noise (Divjak & Caldwell-Harris, 2015, p. 56). For example, Rubenstein et al. (1970) show that recognition of high-frequency words is faster than recognition of low frequency words. They asked participants to discriminate between English words and nonsense words that followed English orthographic and phonological rules - participants had to decide whether a given item was an English word or not. The authors found that English words of high frequency were recognised significantly faster than English words of low frequency. This effect also takes place when two or more words co-occur frequently. For example, Snider & Arnon (2012) used a phrasal decision task to measure the processing times of four-word phrases that differed in their frequencies. Participants saw one phrase at a time on the computer screen and were asked to decide as quickly as possible whether the phrase was a possible sequence in English (e.g. *I saw the man*) or not (e.g. *I saw man the*). All the phrases were matched for the frequency of their individual parts so it was only the frequency of the whole phrases that differed. It was found that the higher-frequency phrases were decided on faster than lower-frequency phrases (Snider & Arnon, 2012, p. 134). This effect results from *chunking* - if two or more items co-occur frequently, they will form a unit and will become entrenched and retrieved from memory as a unit (Bybee, 2010, p. 34). Instead of retrieving the individual component parts of a chunk and putting them together into a phrase on the spot, speakers retrieve the chunks as a whole (Divjak & Caldwell-Harris, 2015, p. 61). Prefabricated expressions, constructions, conventionalised collocations, and multi-word compositional phrases have been shown to emerge as a result of chunking of

frequently co-occurring items (Arnon & Snider, 2010; Bybee, 2010; Snider & Arnon, 2012).

Frequency of occurrence can also help to interpret the meaning of an ambiguous word or structure when no other clues are available to the speaker. For example, Teddiman (2012) asked participants to decide whether words that they see are nouns or verbs. The stimuli included unambiguous nouns and verbs (e.g. *bird* or *earn*) and categorically ambiguous words (e.g. *work* or *walk*). Half of the ambiguous words occurred more frequently as verbs and the other half were more frequent as nouns. In 75% of cases, participants categorised the ambiguous words into the category, in which that word occurs more frequently (ibid., p.241). That is, if a word is more frequent as a verb than a noun, participants were more likely to categorise that word as a verb, and if a word is more frequent as a noun than a verb, then it was more likely to be categorised as a noun. Juliano and Tanenhaus (1993) conducted a self-paced reading task in order to investigate the influence of frequency in the interpretation of the meaning of *that* in various syntactic environments. The word *that* can have various interpretations - an analysis of the Brown Corpus that the authors conducted showed that at the beginning of a sentence, *that* serves most frequently as a pronoun (54%), a determiner (35%), and a complementizer (11%), but when it follows a verb, it is most often a complementizer (93%), then a determiner (6%), and rarely a pronoun (1%). The self-paced reading study found that speakers experience processing difficulties if the structure that is being processed runs counter to what the speakers would expect based on the frequencies of the regular patterns in language (Juliano & Tanenhaus, 1993, p. 598). For example, when *that* occurs at the beginning of a sentence, speakers interpret it automatically as a pronoun - the reading times for conditions in which that was the case (e.g. *That*

experienced diplomat would be very helpful to the lawyer) are shorter than reading times for conditions in which the sentence-initial *that* is a complementizer (e.g. *That experienced diplomats would be very helpful made the lawyer confident*). Similarly, when *that* follows a verb (e.g. *The lawyer insisted that*...), speakers interpret *that* as a complementizer that will introduce a sentential complement - the reading times for conditions in which that was the case (e.g. *The lawyer insisted that experienced diplomats would be very helpful*) are shorter than reading times for conditions in which post-verbal *that* was a pronoun (e.g. *The lawyer insisted that experienced diplomat would be very helpful*).

Although the above are only few of many examples of frequency effects in language, it is already clear that repetition plays an important role in the formation of linguistic system and in its use. Frequency of occurrence influences our use of lexical items and grammatical structures, helps us deal with ambiguity, and leads to chunking of co-occurring items, to name but a few of its effects. These effects seem to have been neglected by translation scholars in attempts to explain linguistic choices in translation, despite their clearly important role.

2.2. Measures of frequency

Frequencies of occurrence can be obtained in various ways, for example, they can be gathered from computer-readable corpora or by asking speakers to estimate the frequency of a given item (Divjak & Caldwell-Harris, 2015, p. 54). Deciding how to count frequencies is not straightforward because it is still not clear how the frequency effects described above come about - are we simply counting the number of times we encountered a given word/phrase/construction, or are other factors important, for example, the context in which the word/phrase/construction occurred

(cf. Divjak & Caldwell-Harris, 2015). Various measures of frequency have therefore been proposed - some rely on counting the occurrences of isolated words while others consider the various ways in which to include the role of the contextual environment of the item(s) under investigation.

Two commonly used measures of frequency are token and type frequencies. Token frequency refers to the number of times a given word, e.g. *broken* or *have*, or a sequence, e.g. *I don't' think*, occurs in the input. This type of frequency is said to lead to entrenchment and automation - the more often a speaker is exposed to a given word or sequence, the more entrenched it becomes and the easier it is for the speaker to access and use that word or sequence (Bybee & Hopper, 2001, p. 16). Type frequency refers to the number of different items that are used within a given pattern or construction, e.g. the number of English verbs that have the regular past tense *-ed* form (Bybee, 2007, p. 269). This type of frequency is said to lead to the formation of general categories for constructions that occur with many lexical items (ibid., p. 275). Until recently, psycholinguists relied on token frequencies in their investigations and gave little attention to contextual factors, unlike corpus linguists who always realised the importance of context (Divjak & Caldwell-Harris, 2015, p. 58). A number of contextualised frequency measures have been proposed, including the attraction and reliance (Schmid, 2000) which measure to what extent different types of constructions attract specific types of nouns, and to what extent nouns rely on certain constructions; collostructional analysis (Stefanowitsch & Gries, 2003) which investigates the interaction between lexemes and the grammatical structures associated with them; dispersion (e.g. Baayen, 2010) which measures the spread of a word across texts; or conditional probabilities (e.g. Jurafsky, 1996) which calculates the probability of an item given the lexical, syntactic, semantic or other factors

surrounding it. These contextualised frequencies have been shown to result in better predictions than frequencies of isolated items (Divjak & Caldwell-Harris, 2015, p. 60).

2.3. Frequency, entrenchment, and representation

An important question in interpreting the role of frequency is to what extent it relates to implicit linguistic knowledge and to the representation of linguistic structures in the speakers' minds. In 2000, Schmid proposed the 'From-Corpus-to-Cognition Principle' according to which "frequency in text instantiates entrenchment in the cognitive system" (Schmid, 2000, p. 39). He argued that not only linguistic preferences, but also cognitive functions and processes, can be investigated by analysing language corpora. The assumption that patterns observed in corpora can be linked to patterns in the mind of the speaker underlies many usage-based studies (Blumenthal-Drame, 2012, p. 30). Hilpert (in Arppe et al., 2010, p. 15) points out, however, that many psycholinguists do not consider corpus data to have explanatory power and deem frequency data useless in drawing inferences about language representation in the mind. Biber, Conrad & Cortes (2004, p. 376) argue that frequency data enable researchers to identify patterns that would otherwise go unnoticed, but those patterns are in need of explanation, rather than having any explanatory power themselves. Gilquin (in Arppe et al., 2010, p. 9) argues that raw frequency in corpora should not be blindly accepted as representative of salience in the mind. Schmid (2010, p. 125) himself admits in his later work that the relationship between frequency of occurrence in corpora and entrenchment is still unclear, with frequency being a far less objective measure than initially assumed. As pointed out by Zeschel (in Arppe et al., 2010, p. 10), however, frequency as observed in corpora

has been shown to influence many aspects of language processing and learning. The corpus-to-cognition assumption, the author argues, is therefore a well-documented fact. Divjak & Arppe (2013, pp. 229-230) also justify using corpora to draw inferences about mental grammars - they argue that corpus data are related to speaker's linguistic knowledge and as such are a reflection of its characteristics, even if indirectly.

2.4. The role of frequency in translation

The importance of looking at frequency effects that underlie the linguistic structures contained in comparable corpora becomes clear when we consider the fact that corpora of translated texts and corpora of non-translated texts, used in investigations of t-universals, will never be 100% comparable. If we are to assume that any differences observed between translated and non-translated texts are caused by unique translational behaviour, we would need to ensure that the translated and non-translated texts are perfectly comparable and that they only differ on the translation dimension, and not on any other dimension (Bernardini & Ferraresi, 2011, p. 228). That is, we would need to be able to analyse texts that are identical in content and purpose of production, with the only difference being in the way they were produced - one by means of being translated from another language, and the other by means of monolingual production. That is hardly possible, unfortunately. Instead, we use corpora of translated texts and comparable non-translated texts, which although matched for many characteristics (genre, mode, dates of publications of the texts, and so on), differ along other dimensions, which can influence the phenomena we investigate. For example, our ability to chunk items that co-occur frequently means that we are likely to be guided by chunks in our linguistic choices - in some cases,

such chunks can be up to four words long (Snider & Arnon, 2012). If we do not consider the sentential context of the phenomena we investigate, we run the risk of attributing any differences in their use between translated and non-translated texts to the process of translation, while in actual fact these difference may be related to the context of the sentences in which these phenomena occur, and the linguistic chunks that these phenomena form with various other items in a sentence, as well as other frequency effects. This is illustrated below by looking at the study of optional *that* in English, carried out by Olohan & Baker (2000).

Olohan & Baker (2000) investigate the use of optional *that* after the verbs *say* and *tell*. Their corpus analysis showed that optional *that* is more frequent in the translated corpus than in the non-translated corpus, suggesting that explicitation took place. Among other verb forms, they discuss the occurrence of *that* with the verb form *tells* (ibid., p. 154). The translated corpus contains 55 occurrences of *that* and 25 occurrences of *that*-drop with *tells*. In the non-translated corpus the reverse is observed - 28 sentences with *tells* contain *that*, and 52 do not. This is summarised in Table 1.

construction	non-translated	translated
that	28	55
that-drop	52	25

Table 1: Occurrence of optional *that* in Olohan & Baker (2000, p. 154)

The authors then analyse the type of constructions in which *tells that* occurs and report that straightforward reported speech structures with an animate object (e.g. My doctor tells me that...) account for 71% of *tells that* in the non-translated corpus, but only about 47% in the translated corpus. Reported speech structures with a modifier inserted before the object clause (e.g. He tells her during their conversation

that...) account for 10.7% of *tells that* in the non-translated corpus, and 14.5% in the translated corpus. Finally, figurative uses of *tells* with inanimate objects (e.g. What little knowledge of history I have tells me that...) account for 14.3% of *tells that* in the non-translated corpus, and 38% in the translated corpus. This is summarised in Table 2.

construction	non-translated	translated
reported speech	20 (71%)	26 (47%)
reported speech + modifier	3 (10.7%)	8 (14.5%)
figurative	4 (14.3%)	21 (38%)

Table 2: Type of structures in which *tells* occurs (Olohan & Baker, 2000, p. 154)

The authors state that 'certain lexical or syntactic patterns used with *that* appear to be more prevalent in translated than in original English' (ibid., p. 157). In other words, the sentential contexts of *tells* in the two corpora are different, yet the authors do not factor this into their interpretation of the results, still suggesting explicitation to be the reason for these differences. They do not seem to consider that those lexical or syntactic patterns that are more prevalent in the translated corpus are the ones that chunk with optional *that* and that is why translators chose *that* instead of dropping it. Since there are fewer of those structures in the non-translated corpus, the occurrence of *that* which chunks with those structures is also lower. Unfortunately, the authors do not provide a breakdown of the structures in which *that* was dropped, which prevents the extraction of co-occurrence of the different patterns with *that* to confirm that this is indeed the case. From what we know about chunking and entrenchment, however, it is likely that the 52 sentences in the non-translated corpus, in which *that* was dropped, contain structures that do not chunk with *that*, while in the translated corpus there were only 25 such instances. Investigating the sentential context of each

occurrence of *tells that* would shed a different light on the results, perhaps ruling out explicitation as an explanation.

Chapter 3. Modality and aspect

It was argued in the two previous chapters that the notion translation universals is not plausible and that comparability of comparable corpora should be re-considered as it is likely to underlie the differences observed in comparisons of translated and non-translated texts. This may have led to mistakenly attributing differences in linguistic features of translated texts, as compared with non-translated texts, to translation universals.

This chapter introduces two grammatical categories that will be central to the investigations conducted in Part II and III: modality and aspect.

3.1. Modality

Linguistic modality is a complex and widely discussed category, and it is difficult to find one comprehensive description. The category has roots in modal logic and its interest in the concepts of possibility and necessity. Linguists, however, look at the two concepts from a different perspective than logicians - the former are interested in how possibility and necessity are expressed in natural languages, while the latter investigate the logic of reasoning behind them (Portner, 2009, pp. 10-11).

3.1.1. Modal types

Scholars look at modality various perspectives, which results in many different classifications and definitions of modal types. Some authors propose two types (e.g. root and epistemic), others suggest there are three (e.g. deontic, dynamic, epistemic), and some insist on four (e.g. deontic, participant-internal, participant-external, epistemic).

Two-way classifications

Several scholars propose two modal categories: epistemic modality and root modality (e.g. Coates, 1995; Hofmann, 1976; Sweetser, 1982). Epistemic modality refers to the speaker's judgment about the truth of the proposition; the speaker can be confident (epistemic necessity) that, or uncertain (epistemic possibility) whether, the state of affairs uttered is true, based on his or her assessment of the circumstances, for example:

(1) *I must have a temperature* (Coates, 1995, p. 56)

Here, based on the symptoms, the speaker makes a judgement about their condition, and by using *must* they express certainty that what they are is saying must necessarily be true.

Root modality, on the other hand, refers to obligation, permission and ability, both internally and externally placed upon the subject:

(2) *You must finish this before dinner* (ibid.)

Here, the obligation (root necessity) to finish the dinner is placed upon the agent by an external source.

Divjak, Szymor & Socha-Michalik (2015) also propose a two-way classification, but look at modality from the perspective of its functions. Their usage-based corpus and experimental analysis of Polish modals suggest that in the native speaker's mind, modal meanings are likely organised as two semantic fields - *nessecity* and *possibility*, without any further subdivisions into epistemic and root meanings.

Three-way classifications

The most commonly cited classification of modality involves three types: epistemic, deontic, and dynamic. The meaning of epistemic modality is the same as in the previous classification - this is the only modal type that seems uncontroversial. The abovementioned root modality, however, is divided up into two types. Deontic modality refers to situations when a certain state of affairs is (un)desired, be it from a legal, ethical, or moral perspective (Nuyts, 2006, p. 4), for example:

(3) *We should be thankful for what he has done for us, so we must find a way to show our gratitude to him* (ibid., p. 5)

In terms of dynamic modality, Nuyts (2006) distinguishes situations, where capacity and needs/necessities are ascribed to the participant in the clause (*participant-inherent*), and situations where capacity and needs/necessities are determined by the circumstances (*participant-imposed*), for example:

(4) *I have unlocked the back door, so you can enter the house there.*

Here, the possibility for the participant to enter the house is external to them, rather than internal.

Four-way classifications

Van der Auwera & Plungian (1998) recognise four modal types: participant-internal, participant-external, deontic, and epistemic. Epistemic and deontic modality remain the same as above. The only difference is that the authors divide dynamic modality into two separate categories - participant-*internal* and *participant-external*. The former refers to possibilities and necessities internal to the participant, while the latter covers situations in which the conditioning factors are external to the participant. Their deontic modality is a more specialised extension of participant-external modality.

Bybee, Perkins, & Pagliuca (1994) argue for agent-oriented, speaker-oriented, subordinating, and epistemic categories. Agent-oriented and speaker-oriented modality cut across the traditional notion of deontic and dynamic modality, with the former referring to any conditions (external and internal) that exist on the agent of the sentence, including obligation, necessity, desire, ability, permission and root possibility:

(5) I need to hear a good loud alarm in the mornings to wake up. (Bybee & Fleischman, 1995, p. 177)

Speaker-oriented modality, on the other hand, covers all directives that enable the speaker to modify the listener's behaviour:

(6) You can start the revels now. (ibid., p. 179)

Finally, subordinating modality refers to forms that are used to mark the verbs in certain types of subordinate clauses, for example:

(7) I suggest that he should call you immediately. (ibid., p. 180)

3.1.2. Disputed members

Apart from the disagreement about the internal structure of the modal category, there is also uncertainty about the membership of some modal types, such as evidentiality, mood, volition, and alethic modality. The description here is only brief - a thorough discussion of the modal types can be found in Nuyts (2006).

Evidentiality

Evidentiality refers to utterances, in which the truth-value of the described state of affairs is explained with reference to certain sources of knowledge, including general

knowledge, visual experience, auditory evidence, hearsay, and reasoning (Palmer, 2001, p. 8; van der Auwera & Plungian, 1998, p. 85). Evidentiality is closely related to epistemic modality - they both rely on evidence. Palmer (2001) and Bybee & Fleischman (1985) include evidentiality in their accounts, together with epistemic modality, while van der Auwera & Plungian (1998, p. 85) only include 'inferential' evidentiality.

Mood

Mood has traditionally been associated with different types of utterances that exist in language - those expressing facts (indicative mood), questions (interrogative), orders (imperative), wishes (optative), etc. Palmer (2001, p. 4) argues that normally languages deal with modality by means of modal systems and mood, i.e. he places the two together at the same level. According to Hengeveld (2004, p. 1190) and van der Auwera & Plungian (1998, p. 83), on the other hand, mood is subdivided into two smaller categories: illocution and modality; that is, modality is a subcategory of mood, rather than its equal 'partner'. Bybee, Perkins, & Pagliuca (1994) also consider mood to be a subcategory of modality under the heading of speaker-oriented modality.

Alethic modality

Alethic modality refers to the truth of the propositions; propositions can either be necessarily true, i.e. true in all logically possible worlds, or they can be possibly true, i.e. not necessarily false, or true in at least one possible world (Lyons, 1977, p. 791). Alethic modality is related to epistemic modality, the difference between the two lying in objectivity. Epistemic modality is considered subjective, and alethic modality - objective and equivalent with statements of fact. Notion of alethic

modality is rarely used - it coincides with the notion of objectivity, which is not always considered modal (cf. Palmer, 1986).

Volition

Within discussions of modality, volition refers to one's will, desire and often intention. In English, it can be expressed by the auxiliary *will*, or with other means, such as *want, hope, wish, would rather*, etc. As with the other disputed categories, volition is sometimes included and sometimes excluded from modality. Bybee, Perkins, & Pagliuca (1994) include desire, intention, and willingness in their agent-oriented modality. Van der Auwera & Plungian (1998, pp. 84-86) excludes these types of statements, arguing that modality should only be limited to domains that show the opposition of possibility vs. necessity.

3.1.3. Expressing modality

There is a wide range of linguistic means to express the abovementioned modal types. Modal verbs are studied most heavily because they constitute "the most coherent class of modal expressions in English" (Perkins, 1983, p. 19). Other exponents of modality include adverbs (e.g. *perhaps, probably*), adjectives (e.g. *it is likely*, *it is necessary*, *you are obliged*), and nouns (e.g. *there is a possibility*, *there is a necessity*) (Von Fintel, 2006, p. 20). If one accepts evidentiality and mood as members of the modal category, then the markers of those two modal types should also be included in this list.

In Polish, modality can be expressed in the following ways (Grzegorczykowa, 2001):

(i) indicative and imperative moods;

(ii) particles, e.g. *na pewno* [certainly], *z pewnością* [with certainty], *chyba* [maybe], *niewątpliwie* [undoubtedly], *prawdopodobnie* [probably], etc.;

(iii) modal verbs, e.g. *móc* [can, be able to], *musieć* [must, have to], *należy* [it is necessary], *trzeba* [it is required], *powinien* [should], *wolno* [it is allowed], *można* [it is permitted].

Hansen (2004) proposes a model for categorising various types of modal markers. He argues that the modal category consists of a small core with specific semantic and syntactic properties, and a periphery, which overlaps with other categories. The internal core of the category consists of fully-fledged modal verbs, which show both the central and the peripheral features. The external core of the category, on the other hand, consists of modal constructions and content words that display only the central characteristics and not necessarily the peripheral ones. Lastly, those constructions and words that do not exhibit the central features to a full extent constitute the periphery of the category.

The core features of modality include:

(i) Semantic characteristics: a modal must express two or more types of modality. For example, verb *móc* [can, be able to] can express capability, permission, or likelihood. As a contrast, the verb *potrafić* [be able to] only expresses capability.

(ii) Morphological characteristics: a modal must express the modal meaning independently, not relying on the construction as a whole.

(iii) Syntactic characteristics: a modal must be a part of the predicate and does not normally occur in other syntactic positions; it is almost always followed by an infinitive.

The peripheral features of modality include:

(iv) Semantic characteristics: a modal must not express any other meanings than the modal ones.

(v) Morphological characteristics: a modal is stripped of some of the characteristics of the category to which it originally belonged, for example, if it is a verb, it cannot form an imperative, or an infinitive. In other words, modal verbs are defective.

According to these criteria, then, such expression as *perhaps* or *probably* are not considered fully-fledged exponents of modality because they do not have all of the core characteristics, for example, they do not express two or more types of modality; they are only able to express epistemicity. Hansen (2004, p. 251) claims that the internal core of the category of Polish modals consists of *mieć* [to have], *móc* [can, be able to], *musieć* [must, have to] and *powinien* [should]. The external core consists of *trzeba* [it is required], *można* [it is permitted], and *należy* [it is necessary], and the periphery consists of *wolno* [it is allowed], *wypada* [to befit] and *nie potrzebować* [it is not necessary].

3.1.4. Summary

In the current thesis, sentences will be regarded as modal when possibility or necessity of the state of affairs is expressed. No other criteria to divide these notions of possibility and necessity into modal types will be applied – here the notions are understood very broadly in line with the results of the study by Divjak, Szymor & Socha-Michalik (2015). Possibility can refer to the state of affairs being possible by virtue of permission, ability, or likelihood. Necessity will refer to the state of affairs

being necessary because of norms, internal and external needs, circumstances, and certainty.

Evidentiality and volition will not be included – the focus here is on legal language and its expression of obligation, permission and possibility, which are not associated with evidentiality and volition. Mood is excluded too because the exponents of modality investigated here only include fully-fledged modal verbs in Hansen's (2004) understanding.

As mentioned in Section 3.1.3, Hansen claims that the internal core of the category of Polish modals consists of *mieć* [to have], *móc* [can, be able to], *musieć* [must, have to] and *powinien* [should], the external core consists of *trzeba* [it is required], *można* [it is permitted], and *należy* [it is necessary], and the periphery consists of *wolno* [it is allowed], *wypada* [to befit] and *nie potrzebować* [it is not necessary] (Hansen, 2004, p. 251). Table 3 shows which of Hansen's modal expressions are included in the current analysis.

category	modal verb / expression	status
internal core	*mieć* [to have]	✖
	móc [can, be able to]	✔
	musieć [must, have to]	✔
	powinien [should]	✔
external core	*trzeba* [it is required]	✔
	można [it is permitted]	✔
	należy [it is necessary]	✔
periphery	*wolno* [it is allowed]	✔
	wypada [to befit]	✖
	nie potrzebować [it is not necessary]	✖

Table 3: Hansen's (2004) list of Polish modal verbs used in the current analysis

Mieć [to have] is excluded because first, it has a very common non-modal meaning of possession, and second, it has strong evidential meaning (Hansen, 2004, p. 247)

that is of no interest in the current analysis of legal language. *Wypada* [to befit] and *nie potrzebować* [it is not necessary] are also excluded. *Wypada* [to befit] is more generally used to express ethical or moral norms, rather than legal norms, which are of concern here, while *nie potrzebować* [it is not necessary] expresses a lack of need for something/to do something, but would not normally be used to express prohibition in legal language. The remaining modals are included in the analysis: *móc* [can, be able to], *można* [it is permitted], *musieć* [must, have to], *należy* [it is necessary], *powinien* [should], *trzeba* [it is required], *wolno* [it is allowed]. Other exponents of modality, such as particles, moods, imperatives, etc. will not be taken into consideration.

3.2. Verbal aspect

Aspect is not an easily defined category. It is concerned with time, but from a different perspective to tense. Tense locates situations in time with reference to another time - usually the present, but also past and future - while aspect describes the 'internal temporal constituency' of a situation (Comrie, 1976, pp. 4-5) or a 'presumption about the way an act proceeds in time' (Bermel, 1997, p. 25). For example, aspect enables speakers to differentiate between repeated as opposed to one-off actions, ongoing as opposed to complete actions, actions whose results endure as opposed to actions that have no results or the results are annulled or interrupted. Various models have been proposed to improve our understanding of aspect - these are briefly discussed in Section 3.2.2. First, however, the way aspect is encoded in different languages is discussed in Section 3.2.1. Finally, Section 3.2.3 looks at the choice of aspect in modal context - the object of the corpus comparison carried out in Part II.

3.2.1. Linguistic encoding of aspect

Different languages express aspect in different ways. In English, aspectual functions are intertwined with tenses - continuous (e.g. *I am mowing the lawn*), generalised (e.g. *I mow the lawn*), and perfect (e.g. *I have mown the lawn*). In Slavic languages, including Polish and Russian, these functions are manifested by pairs of morphologically related verbs - perfective and imperfective (Bermel, 1997, p. 25). For example, the perfective form of the Polish verb *to read* is *przeczytać* (example i), and the imperfective is *czytać* (example ii), the difference being the prefix *-prze*:

(i) PL: *Wczoraj **prze**czytałem książkę.*

EN: *I finished reading a book yesterday.* [completed action, focus on totality]

(ii) PL: *Wczoraj czytałem książkę.*

EN: *I was reading a book yesterday.* [action in progress, focus on duration]

Imperfective verbs are said to take an internal view on how a given situation proceeds, as if looking at it from within (Bermel, 1997, p. 25). They are therefore used for ongoing actions and actions that have no result or the result was interrupted or annulled, but also for repeated actions, general-factual statements, and so on. Perfective verbs are said to take an external view, as if looking at the situation from the outside (ibid.). They are therefore used for single, completed actions, actions that span over short time, actions whose results endure, and so on.

In Polish, the vast majority of verbs exist in the two aspectual forms - the perfective and the imperfective. Some verbs are biaspectual, expressing both aspects at the same time, and some exist in the perfective or the imperfective form only. Polish

speakers are therefore forced to choose the aspectual form of a verb whenever one is used. The type of situation described may guide the speakers' choice of aspectual form, for example:

(iii) PL: *Telefon zadzwonił gdy czytałam książkę*.

EN: *The phone rang when I was reading a book.*

In sentence (iii), the action of reading a book expressed by the imperfective verb suggests that the reading of a book was simultaneous to the phone ringing. That is, the phone rang when the subject was in the process of reading. A perfective verb would change the meaning of the sentence:

(iv) PL: *Telefon zadzwonił gdy przeczytałam książkę*.

EN: *The phone rang when I had finished reading a book.*

In sentence (iv), the phone rang after the action of reading the book was completed. Certain situations can therefore only be described by using one, and not the other, aspectual form, and using the other form would result in changes in the conveyed meaning. Situations that require the use of an imperfective verb will often contain the following adverbials of time: *często* [often], *wciąż/ciągle* [still], *zawsze* [always], *zwykle* [usually], *chwilami* [at times], *czasem/czasami* [sometimes], *nieraz* [many times], *długo* [for a long time], *krótko* [briefly], *całymi dniami/tygodniami/latami* [for days/weeks/months] (Kaleta, 1995, pp. 306-307). Situations that require the use of a perfective verb also often occur with certain adverbials, for example *nagle* [suddenly], *nieoczekiwanie* [unexpectedly], *zaraz/za chwilę* [in a moment], *nareszcie/wreszcie* [finally, at last], *natychmiast* [immediately], *wkrótce* [soon] (ibid.).

However, it is not always the case that the choice of aspectual form is objective - the same situation can often be expressed with either aspectual form, and the choice of form depends on the speaker's point of view (Błaszczyk-Szabat, 2005, p. 95; Comrie, 1976, p. 4). For example, the first action in sentences (iii) and (iv) - the ringing of the phone - could be expressed by either aspectual form. Similarly, sentence (v) is also possible in two aspectual versions - the perfective (v(a)) and imperfective (v(b)). This is because in v(a), the speaker perceives the situation as bounded in time so the perfective aspect is chosen, while in v(b) the speaker perceives the situation as unbounded so the imperfective is chosen (Błaszczyk-Szabat, 2005, p. 95):

(v) PL: *Janek <u>zagrał</u> na pianinie.* (a)

Janek <u>grał</u> na pianinie. (b)

EN: *Janek <u>played</u> the piano*.

A pilot analysis of Russian texts (Janda, et al., 2016) showed that the choice of aspectual form is marked by a contextual trigger (such as the adverbials mentioned above) in only 2% of cases. In 15% of cases the choice of aspect is free - either aspectual form could be used in those instances. Finally, the majority of verb forms (83%) occur in utterances in which only one aspectual form is possible, but there is no clear trigger which would indicate that. Scholars have made numerous attempts at modeling aspectual choice in order to understand the mechanism that guide native speakers, especially in situations where no clear trigger is present. These are briefly discussed in the next section.

3.2.2. Models of aspect

It was mentioned in the previous section that imperfective verbs are used for ongoing or repeated actions, actions with no result, or general-factual statements, among other meanings. Perfective verbs are used for single or completed actions, actions whose results endure, and so on. Scholars have tried to understand why this is the case and have come up with a number of notions that are meant to make it easy to understand why native speakers choose aspectual versions of verb the way they do. Some of these notions are:

(i) boundedness - perfective verbs are used for actions that have clear boundaries (beginning and end), while imperfective verbs are used for actions that are unbounded or the reference to the boundaries of the event is irrelevant;

(ii) totality - perfective actions are viewed as a whole, with all possible sub-actions included in the total event, while imperfective actions refer to only some part of the situation;

(iii) definiteness - perfective actions are single and individuated, representing change, while imperfective actions represent stability.

Janda (2004, p. 482) argues that these accounts are no more than a new layer of labels for the types of situations that aspects are used for, without understanding what underlies those labels. Janda's (2004) own account is cognitively-inspired - it is based on the idealized cognitive model (ICM) of matter. According to the account, 'aspectual choice is at least partially based on the metaphorical association PERFECTIVE IS A DISCRETE SOLID OBJECT and IMPERFECTIVE IS A

FLUID SUBSTANCE' (ibid., p. 490). In other words, when a speaker is forced to choose an aspectual form, they will choose the perfective if the situation triggers affinities with embodied knowledge of discrete solid objects, and imperfective will be chosen if the situation triggers affinities with embodied knowledge of fluid substances (ibid.). Aspectual choices are normally based on what Janda calls "inherent properties of matter". For example, solid objects have clear and firm edges by nature - this corresponds to the clear-cut boundaries of perfective situations. If a situation is clearly delineated in time, a speaker will choose the perfective form of a verb to describe that situation because the clear delineation of the solid object's edges is mapped onto the clear temporal delineation of the situation. In certain circumstances, solid objects are considered more desirable due to their wholeness and manipulability (a characteristic termed *graspability*). This property is mapped onto situations in which the completion of an action is desired. In such situations, Janda claims, the speaker is more likely to choose the perfective form of a verb because, again, the wholeness and manipulability of the solid object is mapped onto the desired wholeness and completeness of the situation. Janda argues that her model accounts for a large part of the variation in aspectual principles across Slavic languages, and it does not simply provide a new set of synonyms for perfective and imperfective, like other accounts. Instead, it demonstrates how the category of aspect is grounded in embodied human experience - a claim that is in line with the principles of cognitive linguistics (Janda, 2004, p. 522).

3.2.3. Aspect and modality

Modal verbs and expressions are almost always followed by an infinitive verb that refers to the action that is being modified by the modal. Previous studies have shown

that there are certain tendencies in terms of the choice of aspectual form when verbs occur in modal contexts. It seems that the perfective form is more likely to occur in dynamic and epistemic contexts (Divjak, 2009, p. 261; Wiemer, 2001, p. 217), whereas the imperfective form is more likely to follow modals that express the deontic meaning (Divjak, 2009, p. 261; Divjak, 2011, p. 76). Other properties have also been shown to correlate with the choice of aspect in modal sentences: polarity and state of affairs applicability (Divjak, 2009; Divjak, 2011).

State of affairs (SoA) applicability

Propositions modified by modal verbs can be generic to everyone everywhere, or can apply to a specific person or a specific situation. The following examples show the difference between specific and generic SoAs:

(i) Generic (the proposition in this case applies to all Member States):

PL: Państwa Członkowskie zapewniają, że produkty określone w art. 1 *mogą* być wywożone ze Wspólnoty jedynie wówczas, gdy są zgodne z niniejszą dyrektywą.

EN: Member States shall ensure that the products referred to in Article 1 may be exported from the Community only if they comply with this Directive.

(extracted from PELCRA)

(ii) Specific (the proposition in this case applies to a specific situation that had already occurred):

PL: W roku 1998 PB stanęło przed ryzykiem załamania i rząd *musiał* interweniować i ocalić bank poprzez dokapitalizowanie, stając się tym samym posiadaczem 99,9 % akcji banku.

EN: In 1998 PB faced risks of collapse and the Government *had to* step in and save the Bank through a recapitalisation, thus becoming a 99.9 % shareholder in the Bank.

(extracted form PELCRA)

Although state of affairs applicability to some extent overlaps with type of modality - generic SoAs are frequent in deontic contexts while specific SoAs are frequent in dynamic contexts - the correlation between modal verbs followed by perfective verbs and specific SoA, and between imperfective verbs and generic SoA, is stronger than the correlation of aspect and modality type (Divjak, 2009, pp. 265-266).

Polarity

Polarity captures the negation added to the modal verb. It has a strong association with aspect - 26% of all imperfective infinitives in modal contexts follow a negated modal, in comparison with only 15% of all perfective infinitives (Divjak, 2009, p. 261). That is, if the modal is negated, the infinitive that follows is more likely to be imperfective than perfective.

3.2.5. Summary

Whenever speakers of Polish use a verb, they are required to choose one of the two aspectual forms. The type of situation that a speaker describes will often guide the choice of aspectual form, for example, if a situation was/is/will be ongoing then the

imperfective has to be used. There are also certain triggers, such as adverbials of time, that require a given aspectual form. However, in many cases it is the viewpoint of the speaker that motivates the choice of aspect. It is still not clear how native speakers decide which form to use, despite a number of different accounts being proposed. There are, however, certain properties that seem to correlate with the choice of aspect, especially in sentences that contain modal verbs - type of modality, state of affairs applicability, and negation.

It will be shown in Chapter 4 that the way translators choose aspectual forms in modal sentences differs from the way authors of texts written originally in Polish do. Before we consider translation universals as an explanation for these differences, we will explore other factors that could underlie them.

PART II. DIFFERENCES BETWEEN TRANSLATED AND NON-TRANSLATED TEXTS

Chapter 4. Corpus analysis

A corpus analysis is performed in order to compare the use of modal verbs in translated and non-translated texts. The Behavioral Profiling approach (Divjak, 2004; Divjak & Gries, 2006) is applied to trace the linguistic 'behaviour' of individual modals at semantic, syntactic and discourse-pragmatic levels. It will be shown in Section 4.4 that differences exist in the distribution of aspectual forms of infinitives that follow the modals in translated and non-translated texts. These differences will be investigated in more detail because they could be an indication of universals translational behaviour. In what follows, the corpus analysis is described in more detail, including sources of data (Section 4.1.), data preparation method and the statistical techniques employed to analyse the data (Section 4.2), and the results (Section 4.3). An interpretation of the results is offered in Chapter 5.

4.1. Source of data

Two corpora are used in this analysis: English-Polish Parallel Corpora (PELCRA)[2] and the National Corpus of Polish (NKJP)[3]. The former will be used as a source of translated data - its features will be analysed to find evidence of unique patterns of translational behaviour. The latter will be used as a source of comparable data, i.e. the non-translated texts, against which the linguistic make-up of translated texts will be analysed. Both corpora are described below.

4.1.1 PELCRA Parallel Corpora

The PELCRA Parallel Corpora are a collection of parallel corpora compiled by the PELCRA research group at the Department of English Language at the University of

[2] http://pelcra.pl/res/parallel/pelcra-par-1/ (last accessed in December 2014)
[3] http://nkjp.pl/ (last accessed in February 2016)

Łódź. The collection contains a multilingual parallel corpus, a Polish-English parallel corpus of literary works, a Polish-Russian Parallel Corpus, and an English-Polish Parallel Corpus. The English-Polish Parallel Corpus contains texts from three EU databases: the CORDIS news database, the RAPID press releases, and the *Acquis communautaire* database. The latter database (henceforth: PELCRA legal corpus) is used in this study - it contains a selection of EU laws, including regulations, directives, and decisions. It contains 23,190 sentence-aligned documents - Polish translations, together with their English originals (a total of 28,571,342 target words). The fact that PELCRA contains the translations' originals is an advantage - it allows for a comparison of source texts and translated texts in order to rule out the influence of the source text and source language on any features of the translated texts that are different from the features of comparable non-translated texts. The corpus can be downloaded from the PELCRA website but there is no software provided to search the corpus. It was therefore necessary to commission a software engineer to create a search engine that could be used to extract the data.

It should be noted that the translation of EU texts into Polish prior to Poland's accession to the European Union in 2004 was surrounded by controversy. A study into the quality of translations completed before 2004 (Szymor, 2011) provides evidence that the translators commissioned with the translation of these texts were often inexperienced trainee translators or students of English philology rather than professional translators with the necessary experience that would enable them to carry out such specialised task as legal translation. Even though Poland's accession to the EU was not officially confirmed until 1997, the country had anticipated this invitation for at least 6 years. Nevertheless, until July 2000 the translation and verification of translated legislation was 'incidental' (Kancelaria Prezesa Rady

Ministrów, 2001, p. 187). The numerous strategies and reports on the advancement of preparation for accession produced in the 1990s hardly mention the need for translating the legislation, and there was no strategy for translation and verification work. Many newspapers and magazines (including *Gazeta Wyborcza*, the country's biggest quality newspaper) discussed the poor quality of the translated legislation. The errors in the translations of EU legislation into Polish were so frequent that the Polish Supreme Administrative Court mentions them in its annual report (NSA, 2010, p. 173). It was argued that these issues were caused by outsourcing the translation and verification work to a translation company that was new to Poland and had little experience translating into Polish. The company won the bid because it offered the lowest price and competitive completion dates, it was argued. As such, the company had to employ students and inexperienced translators because it could not afford to pay professionals. This speculation was confirmed through personal communication with an anonymous translator who was employed by one of the outsourcing companies. The translator confirmed that he was recruited whilst still at university and had no experience in professional legal translation. He argued that he was given the work because he accepted the rate of 12PLN per page of translation, even though the standard rate charged by professionals was at that time around 20-25PLN per page. The fact that inexperienced translators produced pre-accession EU translation is important for the current study because PELCRA consists of EU legislation translated up to 2006, with pre-2004 texts constituting 51.65% and post-2004 texts - 48.35%. That is, it is likely that just over 50% of the texts analysed here were translated by people with little or no experience with legal language. This will become relevant when we discuss the reasons for the differences observed between the features of translated and non-translated texts.

4.1.2. National Corpus of Polish (NKJP)

The National Corpus of Polish (NKJP) is a 239-million-word, balanced, representative, and morpho-syntactically annotated corpus of authentic Polish texts. It is a joint initiative of the Institute of Computer Science at the Polish Academy of Sciences, the Institute of Polish Language at the Polish Academy of Sciences, the Polish Scientific Publishers PWN, and the Department of Computational and Corpus Linguistics at the University of Łódź. NKJP is a freely available resource with online search engine.

NKJP contains literary and journalistic texts, non-fiction, specialist periodicals and journals, other written texts (including administrative texts and instructions), websites, transcripts of conversations, and some other types of texts. This study uses the 7-million-word subcorpus of administrative texts (henceforth: NKJP legal corpus), which matches the normative nature of the EU laws in the PELCRA legal corpus. It contains extracts from the Polish Civil Code (*Kodeks Cywilny*), Penal Code (*Kodeks Karny*), Code of Offences (*Kodeks Wykroczeń*), Labour Code (*Kodeks Pracy*), Commercial Code (*Kodeks Handlowy*), Code of Civil Procedure (*Kodeks Postępowania Administracyjnego*), Family and Guardianship Code (*Kodeks Rodzinny i Opiekuńczy*), and The Journal of Laws of the Republic of Poland (*Dziennik Ustaw Rzeczypospolitej Polskiej*).

The NKJP legal corpus is substantially smaller than the PELCRA legal corpus, which might pose certain problems for a comparison between the two. However, there are no other corpora that would enable a comparison between translated and non-translated Polish texts. The NKJP legal corpus - even though smaller - is still large enough to provide a representative sample of non-translated legal Polish.

4.2. Method

For each of the 7 modal verbs chosen for analysis, two samples of 250 independent observations each were extracted - one from NKJP (the non-translated sample) and one from PECLRA (the translated sample). Only 158 observations of *wolno* [it is allowed] could be found in NKJP and 222 observations in PELCRA. No observations of *trzeba* [it is required] were found in PELCRA so *trzeba* [it is required] was excluded from the analysis. Table 4 lists the extracted samples, including the number of observations.

sample	translated sample size	non-translated sample size
móc [can, be able to]	250	250
można [it is permitted]	250	250
musieć [must, have to]	250	250
należy [it is necessary]	250	250
powinien [should]	250	250
wolno [it is allowed]	222	158
total	1472	1408

Table 4: Size of samples extracted from NKJP and PELCRA

The non-translated samples were extracted via the online search engine provided by NKJP, while the translated samples were extracted with the use of custom built software. While NKJP search engine returns samples of randomly selected observations, the software we built for PELCRA does not have that option. As such, it was necessary to extract all instances of a given modal from PELCRA, manually randomise all observations, and select a random sample of 250. We limited the extraction of observations to one observation from the same text - this was done to

ensure that linguistic preferences of individual authors do not affect the linguistic patterns in the samples. Moreover, all extracted observations were manually validated to ensure that other non-modal meanings of the modals are not mistakenly included in the analysis (e.g. *należy* [it is necessary] as a finite verb means 'belongs', while *wolno* [it is allowed] has an adverbial meaning of 'slowly').

Once extracted and validated, the data was annotated in line with the Behavioral Profiling procedure, described next.

4.2.1. Behavioral Profiling

Behavioral Profiling (BP) is based on the assumption that distributional similarity correlates with functional and conceptual similarity, and that the choice of a lexical item is affected by its broader context (Divjak & Gries, 2006, p. 30). This means that sentences extracted from a corpus are annotated for every clue possible and then analysed statistically to capture the behavioral profiles of individual items. BP has been used in a number of lexical semantic studies (e.g. Berez & Gries, 2009; Dilin, 2012; Divjak, 2004; Divjak & Gries, 2006; Jansegers, Vanderschueren, & Enghels, 2015; to name but a few), but also in studies of learner language (Deshors & Gries, 2014) and mental lexicon (Divjak & Arppe, 2013; Divjak, Dąbrowska, & Arppe, 2016; Divjak & Gries, 2008). The approach has also been applied in contrastive studies (Divjak & Gries, 2009) and translation studies (Halverson, 2017; Szymor, 2015). The approach is typically used to discover differences and/or similarities between meanings of synonymous words or senses of polysemous words. It is also used for cross-linguistic comparisons of word meanings, which enable researchers to establish whether a one-to-one correspondence exists between seemingly equivalent items in different languages. Here, instead of establishing similarities and differences

between words/senses in the same source of data or comparing the meaning of words in two different languages, Behavioral Profiling is applied to compare the use of the same item (various modal verbs) in two different sources of data - translated and non-translated texts. The aim is to compare whether the same item is used similarly by different groups of speakers of the same language - speakers that have been argued to differ in their linguistic choices. Unlike other corpus methods applied in Translation Studies, Behavioral Profiling takes into consideration a number of contextual properties that may correlate with the use of the investigated items, allowing for a fine-grained analysis that is currently unachievable in other ways.

The BP method involves the following steps (Divjak & Gries 2008, p191; Gries & Divjak, 2010, p. 338):

1) Sentence extraction: a random sample of the word's lemma is extracted from a corpus; each instance of the lemma is retrieved with its sentential context to allow for the annotation of contextual properties;
2) Annotation: each instance of the extracted lemma and its context is annotated for a number of relevant properties - morphological (tense, aspect, mood, voice, number, gender, case, etc.); syntactic (clause type, sentence type, etc.); semantic (type of subject, properties of the process denoted by the verb, adverbs, particles, connectors, negation, etc.). Annotation usually starts from the observable characteristics of the lexical item in question, and then extends to the properties of the other elements in the clause/sentence;
3) Statistical analysis: a table of co-occurrences is created and analysed by means of a variety of statistical techniques.

The objective annotation of all instances of the item in question and the subsequent statistical analysis of the features allows for a reliable and replicable investigation.

4.2.2. Annotation procedure

In line with the BP annotation procedure, all 2,880 observations extracted from the two corpora are manually annotated for a number of properties. There is no set list of variables to annotate - it is up to the researcher to decide what to include in the analysis, although it is advised that all possible clues in the sentence are accounted for because we may not know what prompts a choice of a particular lexical item in the speaker's mind.

Apart from the aspectual form of the verb that follows the modal verb, two groups of variables were also considered here: semantic and formal properties. The variables cover the basic elements in the sentence - subject, verb, and object - but also extra-linguistic properties related to the state of affairs. They are annotated for in order to check for any other differences in the use of modal verbs and aspect. Table 5 contains a list of all the variables, followed by a description of each variable.

category	variable	variable level
semantic	modal function	necessity
		permission
		possibility
		probability
	modality type	deontic
		epistemic
		external
		internal
	state of affairs applicability	generic
		specific
	subject semantics	inanimate:abstract
		inanimate:concrete
		animate:human
		none
		animate:organism
	infinitive semantics	physical
		physical other
		physical exchange/transfer
		physical motion
		physical motion other
		speech
		mental/intellectual activity
		perception
formal	voice of the infinitive	active
		passive_byc
		passive_zostac
	polarity	negative
		positive
	aspect of the infinitive	imperfective
		imperfective_only
		perfective
		perfective_only

Table 5: List of annotated properties

(1) Modality type

This property refers to the traditional meanings expressed by modals, as described in Chapter 3 (Section 3.1). For the purposes of the current corpus analysis, Van der Auwera & Plungian's (1998) four-way classification is adopted - each sentence is

annotated as expressing either deontic, epistemic, participant-external or participant-internal modality. The following definitions have been applied when annotating (adopted from Divjak et al., 2015):

(i) Participant-internal - participant's internal abilities (possibility) and needs (necessity):

PL: Mam ten komfort że nie jestem obciążony niechęcią do RM, *mogę* więc o nim pisać obiektywnie.

EN: At least I'm not burdened with animosity towards RM, so I *can* write about it objectively.

(ii) Participant-external - possibilities and necessities influenced by factors external to the participant:

PL: *Możesz* również użyć dowolnie wybranej fotografii dla tła pulpitu.

EN: You *can* also set any chosen picture as a desktop background.

(iii) Deontic - permissions (possibility) and obligations (necessity) imposed on the participant by social /moral / legal norms:

PL: Programy te są darmowe do użytku indywidualnego. Firmy *powinny* zakupić licencje. EN: This software is free for personal use only. Companies *must* buy a license.

(iv) Epistemic - a proposition is judged to be uncertain (possibility) or probable (necessity) relative to some judgement(s):

PL: Matka prawdopodobnie gotuje obiad, bo cóż by *mogła* innego robić w tę letnią słoneczną niedzielę.

EN: Mother is probably cooking dinner because what else *could* she be doing on this sunny Sunday.

In the extracted samples, no instances of participant-internal modality were found.

(2) Modal function

This property is related to modality type. It was mentioned in Section 3.1 that Divjak, Szymor, & Socha-Michalik (2015) look at modality in Polish from the point of view of the functions modal verbs express. They analyse usage data (corpus and experimental) and do not find evidence for the existence of the modal types proposed in traditional accounts of modality. Instead, they find that two modal functions of possibility and necessity make up the core of the Polish modal system, with an extension to more specific meanings of probability, permission and prohibition (ibid., p. 347). The functions are annotated according to the following definitions (examples taken from Divjak, Szymor, & Socha-Michalik, 2015):

(i) necessity - the action stated in the proposition is required, either by circumstances that are internal to the participant or by external circumstances (including legal, social or moral norms), e.g.:

PL: [. . .] studia fotograficzne *musiały* płacić za korzystanie z tego rozwiązania firmie Kodak.

EN: [. . .] photo studios *had to* pay Kodak for using this solution.

(ii) possibility - the action stated in the proposition is possible, either by circumstances that are internal to the participant, or by external circumstances, e.g.:

PL: Nie żartuj, nikt nie *może* przewidzieć, kiedy umrze.

EN: Don't be silly, nobody *can* predict when they'll die.

(iii) permission/prohibition - the action stated in the proposition is possible by virtue of permission, e.g.:

PL: [. . .] nasłani szpiedzy zadali [Jezusowi] to podstępne pytanie: Czy *wolno* nam płacić podatek Cezarowi, czy nie? '

EN: [. . .] the spies asked [Jesus] this tricky question: are we *allowed* to pay tax to Caesar or not?'

(iv) probability - a proposition is judged to be uncertain or probable based on evidence or a judgement, e.g.:

PL: [. . .] gdy to się stało, w Polsce rządził jeszcze Jaruzelski. *Musiało* to więc być dawno.

EN: [. . .] when this happened, Poland was still under Jaruzelski's rule. So it *must* have been a long time ago.'

(3) State of affairs (SoA) applicability

As mentioned in Section 3.2.3, propositions that are modified by modal verbs can be generic to everyone everywhere, or can apply to a specific person or a specific situation. A correlation between state of affairs, aspect, and modality has been

observed - perfective verbs are more frequent in specific SoA and dynamic modality, while imperfective verbs seem more frequent in generic SoA and deontic modality. State of affairs will be annotated either as generic or specific:

(i) Generic (the proposition in this case applies to all Member States):

PL: Państwa Członkowskie zapewniają, że produkty określone w art. 1 *mogą* być wywożone ze Wspólnoty jedynie wówczas, gdy są zgodne z niniejszą dyrektywą.

EN: Member States shall ensure that the products referred to in Article 1 may be exported from the Community only if they comply with this Directive.

(extracted form PELCRA)

(ii) Specific (the proposition in this case applies to a specific situation that had already occurred):

PL: W roku 1998 PB stanęło przed ryzykiem załamania i rząd *musiał* interweniować i ocalić bank poprzez dokapitalizowanie, stając się tym samym posiadaczem 99,9 % akcji banku.

EN: In 1998 PB faced risks of collapse and the Government *had to* step in and save the Bank through a recapitalisation, thus becoming a 99.9 % shareholder in the Bank.

(extracted form PELCRA)

(4) Subject semantics

Following Divjak & Gries (2006) and Divjak, Szymor, & Socha-Michalik (2015), the type of subject in each sentence is annotated according to these labels:

(i) inanimate: abstract

PL: Rynki regulowane *muszą* zapewnić stosowanie właściwych uzgodnień w zakresie nadzoru, umożliwiających kontrolowanie obrotu i rozliczanie takich instrumentów finansowych.

EN: The regulated market *must* ensure that appropriate supervisory arrangements are in place to monitor trading and settlement in such financial instruments.

(extracted from PELCRA)

(ii) inanimate: concrete

PL: Państwa Członkowskie zapewniają, że produkty określone w art. 1 *mogą* być wywożone ze Wspólnoty jedynie wówczas, gdy są zgodne z niniejszą dyrektywą.

EN: Member States shall ensure that the products referred to in Article 1 *may* be exported from the Community only if they comply with this Directive.

(extracted from PELCRA)

(iii) animate: human

PL: Ekspert *musi* być obywatelem Państwa Członkowskiego innego niż kraj wywozu lub kraj przeznaczenia.

EN: The expert must be a national of a Member State other than the exporting country or the country of destination.

(extracted from PELCRA)

(iv) animate: organism

PL: [...] zwierzęta przeznaczone do wywozu *muszą* zostać odizolowane od innych zwierząt do momentu wywozu.

EN: [...] the animals to be exported *must* be segregated from other animals until they are exported.

(extracted from PELCRA)

It should be noted that not all of the modal verbs investigated here take a subject. Two of the modals - *należy* [it is necessary] and *można* [it is permitted] - are subjectless, while *wolno* [it is allowed] allows a subject, but does not require it.

(5) Infinitive semantics

The semantics of the infinitives that follow the modal verbs are also annotated. The classification introduced in Divjak (2004) is used here, with the following categories (each category includes its metaphorical meanings):

(i) physical action - verbs that do not have an accusative slot and refer to actions that involve the subject, e.g.:

PL: Cena ta *może* się zmieniać w zależności od lokalizacji masła oferowanego do sprzedaży.

EN: This price *may* vary according to the location of the quantities of butter offered for sale.

(extracted from PELCRA)

(ii) physical action other - verbs with an accusative slot that is affected by the action, e.g.:

PL: *Należy* zintegrować środki nałożonych w ramach wspólnego stanowiska 2002/829/WPZiB [...]

EN: *It is appropriate* to integrate the measures imposed by Common Position 2002/829/CFSP [...]

(extracted from PELCRA)

(iii) physical exchange/transfer - verbs with an object slot that is transferred but is unaffected by the action, e.g.:

PL: Sprzeciw *powinien* zostać przesłany w ciągu 30 dni od doręczenia pozwanemu nakazu.

EN: The statement of opposition *shall* be sent within 30 days of service of the order on the defendant.

(extracted from PELCRA)

(iv) physical motion - verbs that involve motion of the subject, e.g.:

PL: Zwierzętom nie *wolno* wchodzić na teren zakładu.

EN: No animal *may* enter the premises.

(extracted from PELCRA)

(v) physical motion other - verbs that involve putting another entity into motion, e.g.:

PL: Wyobraź sobie, że *musisz* popchnąć stojący na drodze, unieruchomiony samochód..

EN: Imagine that you *have to* push an immobilised car that's on the side of the road...

(extracted from NKJP)

(vi) speech - verbs involving communication broadly understood, e.g.:

PL: [...] zaświadczenie lekarskie *musi* wskazywać okres jego ważności [...]

EN: [...] medical certificate *must* indicate its period of validity [...]

(extracted from PELCRA)

(vii) mental/intellectual activity, e.g.:

PL: [...] Komisja *powinna* rozważyć jej wprowadzenie w trakcie aktualnie podjętej zmiany rozporządzenia.

EN: [...] the Commission *should* consider its introduction during the revision of the Regulation, currently being undertaken.

(extracted from PELCRA)

(viii) perception - verbs that involve passive looking and active perceiving, e.g.:

PL: [...] nie *można* urządzenia tego bezpośrednio zobaczyć [...]

EN: [...] it is not *possible* to inspect the appliance [...]

(extracted from NKJP)

(6) Voice

Three of the six modal verbs can be followed by passive voice - *musieć* [must, have to], *móc* [can, be able to], *powinien* [should]. Verb that follow the modals are therefore annotated for voice - active or passive. There are three types of passive constructions, which reflects aspectual division - there are two auxiliary verbs to choose from, *zostać* [to become; perfective] and *być* [to be; imperfective] (Kaleta, 1995, p. 304). With the two auxiliaries, speakers can form three passive constructions:

(i) *być* + imperfective passive participle, e.g.:

PL: [...] silnik *musi* **być$_{(IMPF)}$ używany$_{(IMPF)}$** w urządzeniu trzymanym przez operatora podczas wykonywania przypisanych mu funkcji.

EN: [...] the engine *must* **be used** in a piece of equipment that is carried by the operator throughout the performance of its intended function(s)

(extracted from PELCRA)

(ii) *być* + perfective passive participle, e.g.:

PL: Aby produkty mogły **być**$_{(IMPF)}$ **wywiezione**$_{(PF)}$, indyjskie władze celne wystawiają podczas trwania procedury wysyłkowej eksportowy dokument przewozowy.

EN: In order for the goods to **be exported**, the Indian customs authorities issue, during the dispatch procedure, an export shipping bill.

(extracted from PELCRA)

(iii) *zostać* + perfective passive participle, e.g.:

PL: Zmiany programów [...] muszą **zostać$_{(PF)}$ zatwierdzone$_{(PF)}$** przez Komisję.

EN: Amendments to programmes [...] shall **be submitted** to the Commission **for approval**.

(extracted from PELCRA)

The difference between (ii) and (iii) is not clear - factors that would help us understand the mechanisms that guide speakers to choose one or the other way to form a passive voice with a perfective verb have not yet been identified. It should be noted that not all of the modal verbs investigated here allow passive voice. *Należy* [it is necessary]*, można* [it is permitted] and *wolno* [it is allowed] can be followed by an active infinitive only.

(7) Polarity

Polarity captures the negation added to the modal verb and is annotated as either 'negative' (for negated modals) or 'positive' (for non-negated modals). As mentioned in Section 3.2.3, polarity has a strong association with aspect - 26% of all imperfective infinitives in modal contexts follow a negated modal, in comparison with only 15% of all perfective infinitives (Divjak, 2009, p. 261).

(8) Aspect of the infinitive

The category of aspect was introduced in Section 3.2. It was mentioned that some verbs exist in one aspectual form only, or are biaspectual. Such instances are removed because we are interested in situations, in which translators/authors have a choice of verbal form. This is because if differences are to be observed, the speakers must have a choice of more than one form, otherwise we cannot say that the author/translator had a choice that could have been influenced by the process of translation. The remaining instances are annotated as *imperfective* or *perfective*. Table 6 shows how many observations were removed from each sample and how many remained.

modal	removed		retained	
	non-translated	translated	non-translated	translated
móc [can, be able to]	48	43	202	207
można [it is permitted]	17	11	233	239
musieć [must, have to]	71	49	179	201
należy [it is necessary]	4	12	246	238
powinien [should]	58	44	192	206
wolno [it is allowed]	102	46	148	204

Table 6: Removed and retained observations

Once all sentences were annotated for the variables described above, they were analysed statistically in R (Version 3.3.1) to uncover similarities and significant differences between translated and non-translated texts.

4.2.3. Statistical analysis

There are a number of statistical techniques that can be used to analyse BP data. For example, when comparing the behavioral profiles of near-synonyms in order to establish the similarities and differences between them, hierarchical agglomerative clustering (HCA) can be used (Divjak & Gries, 2006, p. 36). The dendrograms that HCA returns feature a number of clusters that exhibit various levels of similarity between the investigated synonyms. T-values can be inspected to identify the properties that discriminate between clusters of synonyms, while z-scores allow to identify properties that discriminate between the synonyms within a cluster (ibid., p. 37). Here, instead of comparing the use of synonymous items in the same source of data, the goal is to compare the use of the same item (modal verbs) in two different sources of data (translated and non-translated texts). That is, the aim here is not to find similarities and differences between the modal verbs in Polish, but to assess whether those verbs are used in the same way by two groups of native speakers of Polish - translators and authors of legal texts. It is therefore not necessary to carry out an HCA. Instead, the distribution of contextual properties of each modal in non-translated texts is compared to the distribution of those properties in translated texts. This is done by cross-tabulating each property with the two sources of data, e.g. polarity in non-translated *można* [it is permitted] with polarity in translated *można* [it is permitted] (see Table 7). The aim is to establish whether the two levels of the property (*negative* and *positive*) are equally distributed, or whether the translated sample differs from the non-translated sample.

polarity	non-translated	translated
negative	71	41
positive	162	198

Table 7: Polarity in translated and non-translated *można* [it is permitted]

The distributions are assessed for significance, i.e. whether they occurred by chance, or whether other reasons - such as the process of translation - influenced them. For example, *można* [it is permitted] is negated 71 times in non-translated texts and 41 times in translated texts. That is, there is less negation in the translated *można* [it is permitted] sample than in the non-translated sample. The aim of significance testing is to establish whether that distribution occurred by chance, or whether we are more likely to find a non-negated *można* [it is permitted] in translated texts due to reasons other than chance, e.g. translation universals. Here, significance is established with the help of a Chi-squared test, which compares whether the observed frequencies differ significantly from expected frequencies (Gries, 2014, pp. 369-370). If the observed frequencies are close to the expected frequencies, then it is likely they occurred as a result of chance and there are no significant differences between non-translated and translated texts in terms of the analysed variable. The bigger the difference between the two, the more chance that what is observed is an actual relationship. This information is provided in the form of a p-value (probability value), with a conventional cut-off point (α) of 0.05 – if the p-value is smaller than 0.05, then we assume a significant relationship; if it is bigger than 0.05, then we consider the relationship occurred by chance (McEnery & Wilson, 2001, pp. 84-85). When more than one test is completed on the same data set, there is a chance that a significant result will be obtained by chance - a false positive (Baayen, 2008, p. 105). In order to mitigate the risk of false positive results, the Bonferroni correction can be applied - the standard α=0.05 is divided by the number of tests. If a test produces a

p-value that is lower than the corrected α then it can be said that the observed relationship is significant (ibid., p. 106). Since eight different tests are performed here on the same data set, the α here was divided by 8, resulting in a corrected α=0.006. Chi-squared test performed on the data from Table 7 resulted in p=0.0009948, i.e. lower than the corrected alpha. This suggests that significant association has been observed. The p-value does not say, however, where the differences lie so we also need to look at Pearson residuals to find out the direction of the differences, i.e. whether a given feature is associated with, or dissociated from the translated or non-translated sample (Gries 2014, p. 370). If the residual is a positive value then there is a positive association between the variables; if it is negative, the correlation is a negative one. Only values larger than absolute 2 point to significant associations/dissociations. Table 8 shows the standardised residuals for the distribution of negation in the translated and non-translated *można* [it is permitted] samples. The table shows that we are significantly more likely to find a non-negated *można* [it is permitted] in the translated sample than in the non-translated sample, and significantly less likely to find a negates *można* [it is permitted].

polarity	non-translated	translated
negative	3.400193	-3.400193
positive	-3.400193	3.400193

Table 8: Polarity in *można* [it is permitted] (standardised residuals)

To sum up, the statistical analysis of the corpus data consists of cross-tabulating each property with the two sources of data and testing for significance (Chi-square test). This is done in order to establish whether the behavioral profile of each modal verb is similar in translated and non-translated texts, or whether authors and translators differ in their use of individual modals. The results are discussed in the next section.

4.3. Results

A number of differences were observed in the distribution of the annotated properties, including differences in aspectual forms of the infinitives that follow modals and the aspectual version of the passive voice. The translated and non-translated texts also significantly differ in the distribution of properties that have been shown to correlate with the choice of aspectual form - polarity, modality type, and state of affairs applicability. Differences observed in the use of individual modals are described in turn, followed by a summary. Only statistically significant differences between translated and non-translated samples are discussed.

4.3.1. *Móc* [can, be able to]

The translated *móc* [can, be able to] sample differs significantly from the non-translated sample in the distribution of modal functions, modality types, subject types, semantic groups of infinitives, and voice of the infinitives. For all results see Table 9: significant differences are highlighted in grey; 'na' stands for variable levels that did not occur in the given sample.

variable	variable level	x^2	p-value	df	standardised residuals	
					non-translated	translated
modal function	necessity	18.74	0.0003095	3	1.013540	-1.013540
	permission				4.012711	-4.012711
	possibility				-1.924017	1.924017
	probability				-3.547735	3.547735
modality type	deontic	17.905	0.0001294	2	4.159659	-4.159659
	epistemic				-3.547735	3.547735
	external				-1.924017	1.924017
SoA	generic	0.22842	0.6327	1	0.9804291	-0.9804291
	specific				-0.9804291	0.9804291
subject type	abstract	45.643	6.755e-10	3	-5.7893479	5.7893479
	concrete				0.7975981	-0.7975981
	human				6.3552687	-6.3552687
	none				na	na
	organism				-0.9890587	0.9890587
infinitive semantics	intellectual	14.339	0.01359	5	-3.4139340	3.4139340
	motion				na	na
	motion_other				-1.3225119	1.3225119
	perception				na	na
	physical				-0.2424940	0.2424940
	physical_other				1.9197911	-1.9197911
	speech				-0.4025367	0.4025367
	transfer				0.5853791	-0.5853791
aspect	imperfective	0.24493	0.6207	1	0.5978141	-0.5978141
	perfective				-0.5978141	0.5978141
voice	active	14.565	0.002229	3	-1.317509	1.317509
	byc_impf				1.112092	-1.112092
	byc_pf				2.289562	-2.289562
	zostac_pf				-2.999843	2.999843
polarity	negative	0.11094	0.7391	1	-0.4769239	0.4769239
	positive				0.4769239	-0.4769239

Table 9: Comparison of translated and non-translated *móc* [can, be able to]

We are significantly more likely to find instances of epistemic modality in translated texts than in non-translated texts, and significantly fewer instances of deontic permission. Moreover, the translated sample contains significantly more subjects that are abstract entities than the non-translated sample, and at the same time, significantly fewer human subjects. There are also significantly more infinitives that convey an intellectual act in the translated than in the non-translated sample. Finally, when forming a passive construction, translators are significantly more likely to

choose the perfective auxiliary *zostać* with a perfective passive participle than authors of non-translated texts, who choose *być* significantly more often.

4.3.2. *Można* [it is permitted]

The translated *można* [it is permitted] sample differs significantly from the non-translated in several aspects - modal function, modality type, state of affairs applicability, infinitive semantics, and polarity. The distributions of other properties do not differ significantly. Table 10 contains the details, with significant differences highlighted in grey. *Można* [it is permitted] does not allow a subject or passive voice so these properties were not included in the analysis.

Translated *można* [it is permitted] contains significantly fewer instances of deontic modality, permission in particular. At the same time, it contains significantly more instances of participant-external modality, especially possibility. We are significantly more likely to come across a specific state of affairs - and significantly less likely to come across a generic one - in the translated sample than in the non-translated sample. In terms of the meaning expressed by the infinitives that follow *można* [it is permitted], there are significantly more of those expressing intellectual act, and significantly fewer expressing a physical act upon another entity, in the translated than in the non-translated sample.

variable	variable level	x^2	p-value	df	standardised residuals	
					non-translated	translated
modal function	necessity	42.564	5.719e-10	2	-1.399317	1.399317
	permission				6.459657	-6.459657
	possibility				-6.284022	6.284022
	probability				na	na
modality type	deontic	40.534	1.932e-10	1	6.459657	-6.459657
	epistemic				na	na
	external				-6.459657	6.459657
SoA	generic	77.325	2.2e-16	1	8.920782	-8.920782
	specific				-8.920782	8.920782
subject type	abstract	na	na	na	na	na
	concrete				na	na
	human				na	na
	none				na	na
	organism				na	na
infinitive semantics	intellectual	21.08	0.001775	6	-3.9992398	3.9992398
	motion				na	na
	motion_other				0.4147169	-0.4147169
	perception				-0.9884155	0.9884155
	physical				1.6763120	-1.6763120
	physical_other				3.2066277	-3.2066277
	speech				-0.8877964	0.8877964
	transfer				0.2562416	-0.2562416
aspect	imperfective	0.0012792	0.9715	1	0.1468277	-0.1468277
	perfective				-0.1468277	0.1468277
voice	active	na	na	na	na	na
	byc_impf				na	na
	byc_pf				na	na
	zostac_pf				na	na
polarity	negative	10.837	0.0009948	1	3.400193	-3.400193
	positive				-3.400193	3.400193

Table 10: Comparison of translated and non-translated *można* [it is permitted]

Finally, the translated sample contains significantly fewer negated instances of *można* [it is permitted] than the non-translated sample, and significantly more non-negated instances. There are no differences in the distribution of aspect.

4.3.3. *Wolno* [it is allowed]

The differences in the use of *wolno* [it is allowed] are similar as those observed for the other possibility modals - *móc* [can, be able to] and *można* [it is permitted]. There

are significant differences in the distribution of modal functions and modal types, subject types, infinitive types, and polarity. Table 11 contains the details, with significant differences highlighted in grey. *Wolno* [it is allowed] does not allow passive voice so it was not included in the analysis.

variable	variable level	x^2	p-value	df	standardised residuals	
					non-translated	translated
modal function	necessity	20.174	4.161e-05	2	-4.3956323	4.3956323
	permission				4.4910482	-4.4910482
	possibility				-0.8529696	0.8529696
	probability				na	na
modality type	deontic	18.47	1.726e-05	1	4.491048	-4.491048
	epistemic				na	na
	external				-4.491048	4.491048
SoA	generic	1.4494	0.2286	1	1.713277	-1.713277
	specific				-1.713277	1.713277
subject type	abstract	22.645	0.0001491	4	-2.5448483	2.5448483
	concrete				-0.5363721	0.5363721
	human				4.2670981	-4.2670981
	none				-1.3811775	1.3811775
	organism				-0.8529696	0.8529696
infinitive semantics	intellectual	17.898	0.006493	6	-3.25231628	3.25231628
	motion				-0.09332074	0.09332074
	motion_other				-1.08168379	1.08168379
	perception				na	na
	physical				-1.91825956	1.91825956
	physical_other				1.07520419	-1.07520419
	speech				1.43652186	-1.43652186
	transfer				0.88670601	0.88670601
aspect	imperfective	2.405	0.1209	1	1.710975	-1.710975
	perfective				-1.710975	1.710975
voice	active	na	na	na	na	na
	byc_impf				na	na
	byc_pf				na	na
	zostac_pf				na	na
polarity	negative	42.097	8.683e-11	1	-6.613415	6.613415
	positive				6.613415	-6.613415

Table 11: Comparison of translated and non-translated *wolno* [it is allowed]

There are significantly more instances of participant-external use of *wolno* [it is allowed] in the translated sample than in the non-translated sample, especially

necessity (not to do something). Translators are less likely to use *wolno* [it is allowed] in its deontic sense (permission in particular) than authors of non-translated texts. The translated sample contains significantly more instances of abstract subjects, and significantly fewer instances of human subjects, than the non-translated sample. Just like translated *móc* [can, be able to] and *można* [it is permitted], translated *wolno* [it is allowed] is used with verbs expressing intellectual acts significantly more frequently than non-translated *wolno* [it is allowed]. Finally, negation is significantly more present in the translated sample, and there are significantly fewer non-negates sentences than in the non-translated sample. No significant differences in state of affairs applicability or aspect have been observed.

4.3.4. *Musieć* [must, have to]

The first of necessity modals is *musieć* [must, have to]. The translated sample differs in almost every respect from the non-translated one, with the exception of modal functions and polarity (see Table 12). *Musieć* [must, have to] only expresses one modal function - necessity - in both samples. Moreover, only non-negated instances of *musieć* [must, have to] have been found.

Musieć [must, have to] only expresses necessity so there is no comparison of modal functions in translated and non-translated samples but the translated sample has significantly more instances of participant-external necessity than the non-translated sample, and significantly fewer instances of deontic necessity. This overlaps with the distribution of state of affairs applicability - the translated sample contains more instances of specific SoAs and fewer instances of generic SoAs than the non-translated sample.

variable	variable level	x^2	p-value	df	standardised residuals	
					non-translated	translated
modal function	necessity	na	na	na	na	na
	permission				na	na
	possibility				na	na
	probability				na	na
modality type	deontic	29.949	4.436e-08	1	5.639983	-5.639983
	epistemic				na	na
	external				-5.639983	5.639983
SoA	generic	8.2351	0.004109	1	3.176168	-3.176168
	specific				-3.176168	3.176168
subject type	abstract	25.061	1.5e-05	3	-2.7584612	2.7584612
	concrete				0.3431806	-0.3431806
	human				4.5768648	-4.5768648
	none				na	na
	organism				-1.8973889	1.8973889
infinitive semantics	intellectual	25.913	0.0002311	6	-1.8625912	1.8625912
	motion				-2.3300225	2.3300225
	motion_other				-1.5056197	1.5056197
	perception				na	na
	physical				-0.2105954	0.2105954
	physical_other				4.2545863	-4.2545863
	speech				-1.2981761	1.2981761
	transfer				-2.8544421	2.8544421
aspect	imperfective	18.423	1.769e-05	1	4.396355	-4.396355
	perfective				-4.396355	4.396355
voice	active	30.584	1.04e-06	3	2.755832	-2.755832
	byc_impf				-1.594586	1.594586
	byc_pf				1.959849	-1.959849
	zostac_pf				-5.032840	5.032840
polarity	negative	na	na	na	na	na
	positive				na	na

Table 12: Comparison of translated and non-translated *musieć* [must, have to]

The distribution of some subject types also differs significantly - we are significantly more likely to come across an abstract subject and significantly less likely to find a human subject, in the translated *musieć* [must, have to] than in non-translated *musieć* [must, have to]. The infinitives that follow *musieć* [must, have to] express a motion or transfer of the subject significantly more frequently in the translated sample than

in the non-translated sample. At the same time, the infinitives express a physical act upon another entity significantly less in the translated sample. The distribution of aspectual forms also differs - translated *musieć* [must, have to] is followed by a perfective infinitive significantly more often than non-translated *musieć* [must, have to], and is followed by an imperfective infinitive significantly less. The aspectual difference can also be observed in the distribution of voice - the translated sample contains significantly fewer instances of active voice than the non-translated sample, and instead, it contains significantly more instances of the perfective *zostać* passive construction. That is, when the passive participle is perfective, translators are more likely to go for the perfective auxiliary *zostać* than authors of non-translated texts.

4.3.5. *Powinien* [should]

Fewer statistically significant differences were observed in the use of *powinien* [should] than in *musieć* [must, have to]. There are no differences in the distribution of types of infinitives, aspect, or polarity. There are still differences in the distribution of modal functions and types, SoA applicability, subject types, and voice. See Table 13 for details.

The translated *musieć* [must, have to] sample contains significantly more instances of epistemic and participant-external modality than the non-translated sample, and at the same time, significantly fewer instance of deontic modality. There are also significantly fewer instances of necessity in the translated sample. As could be expected based on the distribution of modal types, the translated sample contains fewer instances of generic SoAs and more instances of specific SoAs than the non-translated sample. There is also a significant difference in the distribution of subject

types - the translated sample contains more abstract and fewer concrete subjects than the non-translated sample.

variable	variable level	x^2	p-value	df	standardised residuals	
					non-translated	translated
modal function	necessity	3.8856	0.0487	1	2.382819	-2.382819
	permission				na	na
	possibility				na	na
	probability				-2.382819	2.382819
modality type	deontic	77.283	2.2e-16	2	8.790810	-8.790810
	epistemic				-2.382819	2.382819
	external				-8.317684	8.317684
SoA	generic	68.797	2.2e-16	1	8.430926	-8.430926
	specific				-8.430926	8.430926
subject type	abstract	70.912	2.722e-15	3	-8.1982437	8.1982437
	concrete				7.9247611	-7.9247611
	human				1.4558109	-1.4558109
	none				-0.6141948	0.6141948
	organism				na	na
infinitive semantics	intellectual	5.3376	0.5013	6	0.20048970	-0.20048970
	motion				-1.67849719	1.67849719
	motion_other				1.46856082	-1.46856082
	perception				na	na
	physical				-0.43254919	0.43254919
	physical_other				0.45287322	-0.45287322
	speech				-0.05862466	0.05862466
	transfer				-0.36697094	0.36697094
aspect	imperfective	3.1928	0.07396	1	1.890476	-1.890476
	perfective				-1.890476	1.890476
voice	active	101.51	2.2e-16	3	0.8653256	-0.8653256
	byc_impf				-2.1007137	2.1007137
	byc_pf				7.5763522	-7.5763522
	zostac_pf				-8.1071018	8.1071018
polarity	negative	0.001242	0.9719	1	1.037121	-1.037121
	positive				-1.037121	1.037121

Table 13: Comparison of translated and non-translated *powinien* [should]

Finally, when a passive construction is used with *musieć* [must, have to], the imperfective *być* passive and perfective *zostać* passive are significantly more frequent in the translated sample, and the perfective *być* passive is significantly less frequent, than in the non-translated sample. That is, when the passive participle is

perfective, translators are more likely to select the perfective auxiliary *zostać* than authors of non-translated texts, while authors are more likely to go for the imperfective auxiliary *być*.

4.3.6. *Należy* [it is necessary]

The last of the three necessity modals is *należy* [it is necessary]. Like *musieć* [must, have to], *należy* [it is necessary] only expresses necessity so there is no comparison of modal functions in translated and non-translated samples. It also does not allow a subject or passive voice so these two variables were also excluded. Significant differences were observed in the distributions of modal types, SoA, infinitive semantics, and aspect of the infinitive. The distributions of polarity do not differ.

variable	variable level	x^2	p-value	df	standardised residuals	
					non-translated	translated
modal function	necessity	na	na	na	na	na
	permission				na	na
	possibility				na	na
	probability				na	na
modality type	deontic	115.18	2.2e-16	1	10.83948	-10.83948
	epistemic				na	na
	external				-10.83948	10.83948
SoA	generic	154.76	2.2e-16	1	12.66032	-12.66032
	specific				-12.66032	12.66032
subject type	abstract	na	na	na	na	na
	concrete				na	na
	human				na	na
	none				na	na
	organism				na	na
infinitive semantics	intellectual	61.192	2.577e-11	6	3.04206468	-3.04206468
	motion				-1.01126601	1.01126601
	motion_other				-0.03319984	0.03319984
	perception				-2.28517256	2.28517256
	physical				-1.38596645	1.38596645
	physical_other				-6.45556721	6.45556721
	speech				2.32250810	-2.32250810
	transfer				4.49174560	-4.49174560
aspect	imperfective	73.722	2.2e-16	1	8.684303	-8.684303
	perfective				-8.684303	8.684303
voice	active	na	na	na	na	na
	byc_impf				na	na
	byc_pf				na	na
	zostac_pf				na	na
polarity	negative	0.000273	0.9868		-1.01772	1.01772
	positive				1.01772	-1.01772

Table 14: Comparison of translated and non-translated *należy* [it is necessary]

There are significantly fewer instances of deontic necessity in translated texts. Instead, translators often use *należy* [it is necessary] to express necessity that is participant-external but has nothing to do with legal/moral/social norms. Translators are also more likely to use *należy* [it is necessary] in specific SoAs, and less in generic SoAs. The translated sample contains more instances of infinitives that convey perception or a physical act upon another entity, and fewer instances of infinitives that express intellectual activity, speech, and transfer. Translators choose

perfective infinitives significantly more frequently than authors of non-translated texts, and choose imperfective infinitives significantly less.

4.3.7. Summary and conclusions

Overall, statistically significant differences were observed in the distribution of all properties. This did not happen by chance - there are reasons that underlie the differences. Table 15 contains a summary of the distribution of properties in translated and non-translated samples.

The translated samples contain significantly more instances of epistemic and participant-external modality, and significantly fewer instances of deontic modality. In terms of modal functions, there are significantly fewer instances of permission in the translated samples than in the non-translated samples, and significantly more instances of possibility and probability. That is, the range of modal meanings expressed by translators with the use of modal verbs is wider than the range of modal meanings expressed in the non-translated texts. Translated texts also contain significantly more instances of specific - and fewer instances of generic - state of affairs. This is perhaps expected when we recall the distribution of modal types. It was mentioned in Section 3.2.3 that deontic modality is associated with generic SoA, while dynamic - with specific SoA. The translated samples have more instances of dynamic modality (participant-external) and also more instances of specific SoAs.

variable	variable level	x^2	p-value	df	standardised residuals	
					non-translated	translated
modal function	necessity	48.344	1.799e-10	3	-0.03840308	0.03840308
	permission				4.19660740	-4.19660740
	possibility				-5.02286084	5.02286084
	probability				-3.83100035	3.83100035
modality type	deontic	242.15	2.2e-16	2	15.55360	-15.55360
	epistemic				-3.83100	3.83100
	external				-14.77291	14.77291
SoA	generic	270.54	2.2e-16	1	16.51121	-16.51121
	specific				-16.51121	16.51121
subject type	abstract	130.77	2.2e-16	5	-7.7710174	7.7710174
	concrete				5.0620123	-5.0620123
	human				8.1334253	-8.1334253
	none				-0.2376313	0.2376313
	organism				-2.3607728	2.3607728
infinitive semantics	intellectual	19.862	0.005875	7	-1.6429282	1.6429282
	motion				-2.5391144	2.5391144
	motion_other				-1.4874469	1.4874469
	perception				-2.3607728	2.3607728
	physical				-0.1902927	0.1902927
	physical_other				1.0528950	-1.0528950
	speech				0.7243446	-0.7243446
	transfer				1.6510110	-1.6510110
aspect	imperfective	32.918	9.611e-09	1	5.778223	-5.778223
	perfective				-5.778223	5.778223
voice	active	123.6	2.2e-16	3	1.099419	-1.099419
	byc_impf				-1.176459	1.176459
	byc_pf				6.516656	-6.516656
	zostac_pf				-9.442702	9.442702
polarity	negative	7.3772	0.006605	1	-2.76891	2.76891
	positive				2.76891	-2.76891

Table 15: Comparison of properties in translated and non-translated samples

The types of subjects in translated and non-translated samples also differ - translators use modal verbs with inanimate abstract subjects and animate (non-human) subjects significantly more frequently than authors of non-translated texts. Translators also use modal verbs with inanimate concrete subjects and with animate human subjects significantly less frequently. The types of infinitives are also different - overall, significantly more infinitives that express motion of the subject and perception can

be found in the translated samples than in the non-translated samples. Finally, there is significantly more negation in the translated texts than in the non-translated texts.

The above differences could have two interpretations: (1) that the translated samples differ significantly in the type of information they convey, affecting their linguistic features, even though both corpora consist of normative legal texts, or that (2) translators use modal verbs for a wider range of meanings than authors of non-translated texts. On one hand, it may be the case that the translated texts deal with a wider range of modal situations, which modify different types of actions of different types of subjects. The implication of this possibility is that the two corpora are not very comparable after all. On the other hand, however, the translator/author cannot choose not to negate a modal verb if what they are saying requires a negation, or they cannot choose an abstract subject if they are saying what a human subject has to, can, or should not do. They can, however, choose how they encode this information linguistically. It may be the case that authors of non-translated texts use modal verbs to express deontic modality only, and for specific types of subjects and actions, while opting for other modal tools, such as those listed in Section 3.1.3, to express other modal situations. The differences described above could suggest that translators are guided by different factors in their choices of modal tools and that they use modal verbs more flexibly than authors of non-translated texts - the reason for that could lie in the process of translation. In order to explore the two possibilities, other ways of expressing modality in Polish would have to be investigated - this is outside the scope of this study, which investigates the use modal verbs only. Nevertheless, this is an interesting avenue worth pursuing in the future.

There are also differences in the choice of aspect of the infinitives that follow the modal verbs in the sample, and in the choice of aspectual version of the auxiliary verb in the passive construction. We can see from Table 15 that translated samples contain significantly more perfective infinitives and significantly fewer imperfective infinitives than non-translated samples. That is, it would seem that translators are more likely to select a perfective infinitive than authors of non-translated texts. At the same time, if using passive voice with a perfective participle, translators are significantly more likely to choose the perfective auxiliary verb (*zostać*) than authors, who are significantly more likely to select the imperfective auxiliary verb (*być*). As discussed in Section 3.2, the choice of aspectual forms of the infinitives can depend on the type of situation described or the way the speaker views a given situation. In case of the latter, any differences in choices between translators and authors of non-translated texts could be indicative of translation universals.

Behavioral profiling enabled us to capture distributional differences in the use of modal verbs between translated and non-translated texts. By annotating for a number of characteristics of the sentences/clauses in which modal verbs occurred, we were able to observe that translators and authors of non-translated texts differ in the choice of aspectual form of verbs that follow the modals. These differences will be investigated in more detail, but first, a sample validation procedure is performed - its aim is to ensure that the distribution of aspectual forms in all the analysed samples is representative of the general population.

4.4. Validation of results

The results obtained for aspectual choices need to be validated in order to ensure their reliability and replicability. Even though the differences in distribution of

aspectual forms are statistically significant, i.e. they can be said to not have occurred by chance, there is still the risk of obtaining false positive results (type 1 error). By extracting additional samples for each of the modals and checking that the distribution of aspectual forms is the same as observed in the original samples, we ensure that the results were not obtained as a result of type 1 error.

Three 250-sentence validation samples were extracted for each modal verb from each corpus, with the exception of non-translated *musieć* [must, have to], for which only one validation sample could be extracted due to their low frequency in non-translated legal texts. There were not enough instances of translated *wolno* [it is allowed] to extract any validation samples.

4.4.1. Non-translated samples

Table 16 contains the original distributions of aspectual forms in non-translated samples. In the *móc* [can, be able to], *można* [it is permitted] and *powinien* [should] samples, perfective is the significantly more frequent form. For *musieć* [must, have to] and *należy* [it is necessary], both forms are equally frequent. In *wolno* [it is allowed], the imperfective is the significantly more frequent form.

sample	impf	pf	p-value
móc [can, be able to]	76	126	0.00043483
można [it is permitted]	52	181	0.00000000
wolno [it is allowed]	134	14	0.00000000
musieć [must, have to]	96	83	0.33121693
należy [it is necessary]	121	125	0.79869940
powinien [should]	81	111	0.03038282

Table 16: The original distribution of aspectual forms in non-translated samples

The validation of *móc* [can, be able to]*, można* [it is permitted]*, musieć* [must, have to], and *wolno* [it is allowed] confirmed that the original samples contain valid results that replicate when other random samples are extracted. That is, we can be quite certain that for those four modal verbs, the distributions of aspectual choices in the original samples are representative of the population. The results did not replicate for *należy* [it is necessary] and *powinien* [should], however. All three *należy* [it is necessary] validation samples show that, instead of an equal distribution of the two aspectual forms, it is more typical for perfective forms to be significantly more frequent than imperfective forms. In the *powinien* [should] validation samples, both aspectual forms are equally frequent, rather than the perfective being more frequent as the original *powinien* [should] sample suggested. Table 17 (below) contains the details. The significantly more frequent forms, if any, are highlighted in grey.

The original *powinien* [should] and *należy* [it is necessary] sample are not representative of the overall populations and as such need to be replaced by more representative samples. *Należy* [it is necessary] validation sample 1 and *powinien* [should] validation sample 1 were chosen to replace the original samples. The comparison of translated *powinien* [should] and *należy* [it is necessary] with their non-translated counterparts needs to be repeated, which is done in Section 4.4.3.

sample	number	impf frequency	pf frequency	p-value
móc [can, be able to]	sample 1	59	141	0.00000001
	sample 2	53	144	0.00000000
	sample 3	59	145	0.00000000
	average	57	143	0.00000000
można [it is permitted]	sample 1	62	181	0.00000000
	sample 2	42	195	0.00000000
	sample 3	40	180	0.00000000
	average	48	185.3333	0.00000000
musieć [must, have to]	sample 1	99	83	0.23562273
	sample 2	na	na	na
	sample 3	na	na	na
	average	na	na	na
należy [it is necessary]	sample 1	92	151	0.00015380
	sample 2	84	162	0.00000066
	sample 3	92	154	0.00007719
	average	89.33333	155.6667	0.00002256
powinien [should]	sample 1	100	110	0.49015296
	sample 2	99	114	0.30405188
	sample 3	89	113	0.09128957
	average	96	112.3333	0.25779989
wolno [it is allowed]	sample 1	182	41	0.00000000
	sample 2	na	na	na
	sample 3	na	na	na
	average	na	na	na

Table 17: Distribution of aspectual forms in non-translated validation samples

4.4.2. Translated samples

Table 18 contains the original distributions of aspectual forms in translated samples. The significantly more frequent aspectual forms are highlighted in grey. All modals but *wolno* [it is allowed] are followed by a perfective infinitive significantly more frequently than by an imperfective infinitive.

sample	impf	pf	p-value
móc [can, be able to]	72	136	0.00000910
można [it is permitted]	52	187	0.00000000
wolno [it is allowed]	172	32	0.00000000
musieć [must, have to]	63	138	0.00000012
należy [it is necessary]	30	208	0.00000000
powinien [should]	68	138	0.00000108

Table 18: The original distribution of aspectual forms in translated samples

After extracting three validation samples per modal verb from PELCRA, each sentence was manually annotated for aspect. As previously, all instances of infinitives that only exist in one aspectual form were removed - the analysis compares the choice of aspectual form in translated and non-translated texts so all infinitives have to exist in the two aspectual forms for the analysis to be possible. Due to very few instances of *wolno* [it is allowed], it was impossible to validate the results for that modal. The results of validation samples can be seen in Table 19 - it provides the frequencies of imperfective and perfective forms in each validation sample and the average frequencies taken from those three validation samples. The significantly more frequent forms are highlighted in grey.

Without exception, perfective aspect is more frequent in all validation samples, just as was the case in the original samples. This suggests that the original translated samples are representative of the individual populations and there is therefore no need to replace any of them.

sample	number	impf frequency	pf frequency	p-value
móc [can, be able to]	sample1	69	138	0.00000000
	sample2	53	144	0.00000000
	sample3	62	136	0.00000000
	average	61.33	139.33	0.00000000
można [it is permitted]	sample1	49	191	0.00000000
	sample2	55	180	0.00000000
	sample3	45	184	0.00000000
	average	49.66	185	0.00000000
musieć [must, have to]	sample1	71	133	0.00001419
	sample2	59	134	0.00000007
	sample3	64	131	0.00000160
	average	64.66	132.66	0.00000129
należy [it is necessary]	sample1	30	202	0.00000000
	sample2	37	200	0.00000000
	sample3	35	199	0.00000000
	average	34	200.33	0.00000000
powinien [should]	sample1	68	136	0.00000193
	sample2	75	127	0.00025349
	sample3	54	143	0.00000000
	average	65.66	135.33	0.00000089

Table 19: Distribution of aspectual forms in translated validation samples

4.4.3. Re-analysis of *należy* [it is necessary] and *powinien* [should]

The distributions of aspectual forms in two non-translated samples - *powinien* [should] and *należy* [it is necessary] - did not replicate, suggesting that these samples are not representative of the general population. The original *należy* [it is necessary] sample contained an equal number of perfective and imperfective forms, while the original *powinien* [should] sample contained more perfective infinitives than imperfective infinitives. The validation suggested that the opposite is the case - *powinien* [should] occurs with the two aspectual forms equally, while *należy* [it is necessary] occurs with perfective infinitives significantly more frequently than with imperfective infinitives. The two samples are therefore replaced by more representative ones (i.e. ones whose distributions are replicated in any subsequently extracted random samples), and the comparison of translated *powinien* [should] and *należy* [it is necessary] samples with the replaced and validated non-translated *powinien* [should] and *należy* [it is necessary] samples was re-run. Table 20 contains the results for *powinien* [should] and *należy* [it is necessary], plus the original results for other modal verb, for reference. The significant differences in aspectual choices are highlighted in grey.

modal	aspect	chi-squared	p-value[4]	df	non-translated	translated
móc [can, be able to]	imperfective	0.24493	0.6207	1	0.5978141	-0.5978141
	perfective				-0.5978141	0.5978141
można [it is permitted]	imperfective	0.0012792	0.9715	1	0.1468277	-0.1468277
	perfective				-0.1468277	0.1468277
musieć [must, have to]	imperfective	18.423	1.769e-05	1	4.396355	-4.396355
	perfective				-4.396355	4.396355
należy [it is necessary]	imperfective	39.188	3.85e-10	1	6.364805	-6.364805
	perfective				-6.364805	6.364805

[4] Here, the conventional α of 0.05 was used because we are only performing one test per data set.

powinien [should]	imperfective	8.8731	0.002894	1	3.078783	-3.078783
	perfective				-3.078783	3.078783
wolno [it is allowed]	imperfective	2.405	0.1209	1	1.710975	-1.710975
	perfective				-1.710975	1.710975

Table 20: Aspectual forms in translated and non-translated after validation

The results of comparison of the replaced *powinien* [should] and *należy* [it is necessary] do not differ greatly from the comparison of the original samples. There are still significantly more perfective forms in the translated *należy* [it is necessary] sample than in the non-translated sample. In terms of *powinien* [should], the difference in aspectual choice between translated and non-translated samples is now significant, whereas it was only nearing significance in the original sample. That is, it seems that translators choose the perfective form more frequently with *powinien* [should] too, just as they do with *należy* [it is necessary] and *musieć* [must, have to].

The final results are reliable and replicable, and will be used here to investigate the influence of the translation process on aspectual choices made by translators. The results indicate that translators are more inclined than authors of non-translated texts to select the perfective form of an infinitive that follows modal auxiliaries of necessity - *musieć* [must, have to], *należy* [it is necessary], *powinien* [should]. It was shown in the analysis above that these results are highly unlikely to have occurred by chance. The second part of this thesis investigates what causes these unusual linguistic choices exhibited by translators.

Chapter 5. Interpretation of aspectual choices in translated and non-translated texts

It may be the case that the differences in aspectual choices observed between translated and non-translated texts result from translation universals. Two options are considered here - explicitation and normalization (Section 5.4). However, the aim of this investigation is to consider the content of the translated and non-translated corpora first so we will look at the type of information that the two corpora convey, which could affect the choice of aspectual form (i.e. situation types) as well as frequency effects that are likely to guide aspectual choice (i.e. analogical mapping

and chunking). Situation types are explored in Section 5.1, analogical mapping in Section 5.2, and chunking in Section 5.3. That is, instead of immediately attributing the differences to translation universals, a rigorous investigation of other factors should be carried out.

5.1. Situation types

Each observation in the translated and non-translated samples was inspected in order to ensure that the observed differences in the choice of aspect are not caused by differences in the type of situations described in the two corpora. Any observations that impose the aspectual form were removed from the analysis. This way, only sentences where the choice of aspect is truly up to the translator/author are compared to see whether differences can still be observed.

Sentences that would require the translators/author to choose the imperfective form of a verb will refer to actions that are in progress, are developing, or are repeated in the present, past or future. The following adverbials of time may (but do not have to) be present: *często* [often], *wciąż/ciągle* [still], *zawsze* [always], *zwykle* [usually], *chwilami* [at times], *czasem/czasami* [sometimes], *nieraz* [many times], *długo* [for a long time], *krótko* [briefly], *całymi dniami/tygodniami/latami* [for days/weeks/months] (Kaleta, 1995, pp. 306-307). Sentences that would require the translators/author to choose the perfective form of a verb will refer to actions that have been completed in the past or will be completed in the future. By completing the action, a certain result is achieved and it can endure. The following adverbials of time may be present: *nagle* [suddenly], *nieoczekiwanie* [unexpectedly], *zaraz/za chwilę* [in a moment], *nareszcie/wreszcie* [finally, at last], *natychmiast* [immediately], *wkrótce* [soon] (ibid.).

Each observation is inspected manually by the researcher and any observations that contained one of the abovementioned adverbials, as well as those that had no clear marker but still required the use of a particular aspectual form, were removed. Table 21 lists the number of sentences removed per modal verb from the translated and non-translated samples.

modal	non-translated	translated
móc [can, be able to]	54	55
można [it is permitted]	76	125
musieć [must, have to]	73	94
należy [it is necessary]	128	148
powinien [should]	64	124
wolno [it is allowed]	33	79

Table 21: Number of observations removed from each sample

Once the sentences in which the context imposes an aspectual form are removed from the sample, the comparison of aspectual forms in translated and non-translated samples is repeated. The results are given in Table 22.

modal	aspect	chi-squared	p-value[5]	df	non-translated	translated
móc [can, be able to]	imperfective	0.34563	0.5566	1	0.7056402	-0.7056402
	perfective				-0.7056402	0.7056402
można [it is permitted]	imperfective	3.4226	0.06431	1	-1.981042	1.981042
	perfective				1.981042	-1.981042
musieć [must, have to]	imperfective	5.8551	0.01553	1	2.557579	-2.557579
	perfective				-2.557579	2.557579
należy [it is necessary]	imperfective	2.6017	0.1067	1	-1.78299	1.78299
	perfective				1.78299	-1.78299
powinien [should]	imperfective	3.0503	0.08072	1	-1.885285	1.885285
	perfective				1.885285	-1.885285
wolno [it is allowed]	imperfective	1.5447	0.2139	1	1.430476	-1.430476
	perfective				-1.430476	1.430476

[5] Again, an α=0.05 was used as a significance cut-off point.

Table 22: Comparison of aspectual forms after removing sentences that impose aspectual form

The results in Table 22 represent the true differences in choices made by translators and authors because there is nothing in those sentences that would require the translator/author to select one aspectual form over the other form. That is, in these sentences translators and authors could have chosen either form. The only statistically significant difference is between the translated and non-translated *musieć* [must, have to] samples, although *można* [it is permitted] and *powinien* [should] are also nearing significance so there is a chance that the distribution of aspect there is also not accidental. The focus will be on the translated and non-translated *musieć* [must, have to] sample for the remainder of this thesis. The aim is to explain why translators choose the perfective form significantly more often than authors of non-translated *musieć* [must, have to].

5.2. Analogical mapping

There is a possibility that the differences in the choice of aspect in the *musieć* [must, have to] sample, which cannot be explained by the type of situation as discussed in the previous section, could be explained by modality type, state of affairs, or negation. As mentioned in Section 3.2, the perfective form seems to be more frequent in dynamic and epistemic contexts, while the imperfective form is more likely to follow modals that express the deontic meaning, i.e. when legal/moral/social norms determine what is possible or necessary (Divjak, 2009, p. 261; Wiemer, 2001, p. 217). The type of state of affairs - applicable generally to everyone and everywhere, or to a specific person or situation - also correlates with aspect. Although state of affairs applicability to some extent overlaps with type of modality - generic SoAs are frequent in deontic contexts while specific SoAs are

frequent in dynamic contexts - the correlation between modal verbs followed by perfective verbs and specific SoA, and between imperfective verbs and generic SoA, is stronger than the correlation between aspect and modality type (Divjak, 2009, pp. 265-266). Finally, polarity also has a strong association with aspect - 26% of all imperfective infinitives in modal contexts follow a negated modal, in comparison with only 15% of all perfective infinitives (Divjak, 2009, p. 261). That is, we are more likely to find an imperfective infinitive if a modal is negated, than we are to find a perfective infinitive. It is therefore likely that instead of the constraints of the process of translation, translators were influenced by one of these variables in their aspectual choices. Table 23 compares the distributions of the three variables across the translated and non-translated samples.

variable	variable level	non-translated	translated
polarity	positive	106	107
	negative	0	0
modality	deontic	105	98
	epistemic	0	0
	external	1	9
SoA	generic	106	107
	specific	0	0

Table 23: Polarity, modality and SoA applicability in translated and non-translated *musieć* [must, have to] samples

We can see from Table 23 that the translated and non-translated *musieć* [must, have to] samples do not contain any instances of negation, and they both contain only instances of generic instances state of affairs. Polarity and SoA applicability cannot therefore explain the difference in distribution of aspectual forms between the two samples. We can also see that there are more instances of participant-external type of modality in the translated *musieć* [must, have to] sample than there are in the non-

translated sample - the difference is statistically significant (x^2=5.0728, df=1, p=0.0243). It may therefore be the case that there are more perfective verbs in the translated *musieć* [must, have to] sample because there are more instances of participant-external modality. However, out of the 9 instances of participant-external modality in the translated *musieć* [must, have to] sample, the perfective form of a verb was chosen only four times. That is, even though there are more instances of participant-external modality in the translated *musieć* [must, have to] sample, these instances do not correlate with the perfective form of a verb being chosen. This is confirmed with a binomial logistic regression model that was fitted to see how well the aspectual form of a verb is predicted by modality type. Binomial regression calculates the odds of certain outcome occurring, given the predictor (for more information, see Section 6.1.3). Here, we wanted to know whether there are significant odds of the perfective form being chosen in the translated *musieć* [must, have to] samples, if modality type is participant-external. The model suggests there is no significant correlation (estimate: -0.3868, std. error=0.7008, p=0.581). The R code used to fit the regression model is available in Appendix 3; all data sets used in the analyses are available to download online[6].

5.3. Chunking

It is a well-evidenced fact that frequent items are more easily retrieved from memory than less frequent items (see Section 2.1 for more details). Moreover, frequently co-occurring items form chunks and are as such retrieved from memory. Since modal verbs are always followed by infinitives in one or the other aspectual form, it is reasonable to argue that the two items are entrenched in memory as a chunk. Moreover, since the majority of infinitives exist in two forms, it may be the case that

[6] https://drive.google.com/open?id=0Bz9KVHFRWI3NclB4aWRRY2poQ28

each modal chunk exists in two versions - a perfective and an imperfective one. The aspectual versions of modal chunks could have various frequencies, leading to automatic activation of the more frequent version in cases when the situation type does not impose the aspectual form of the infinitive on the speaker and analogical mapping does not influence that choice either. This assumption will henceforth be referred to as the *general chunking hypothesis*. The more frequent aspectual version of a verb or a modal chunk will be referred to as *aspectual preference*. The chunking hypothesis 1 can be formulated as follows:

Modal verbs and infinitives that follow them are entrenched in memory as chunks and the aspectual preferences of those modal chunks facilitate the selection and processing of aspect.

Another effect of frequency of occurrence that will be taken into consideration when explaining the differences between translated and non-translated texts is the usage experience of different speakers. Speakers with different usage experience have been shown to exhibit different frequency effects, suggesting that exposure to different language varieties will result in different language structures in the speakers' minds (Caldwell-Harris, Berant, & Edelman, 2012, p. 182). This is important because of the specialised nature of the texts investigated in this thesis. It is reasonable to argue that legal translators and authors of legal texts will have different type of chunks entrenched in memory than average speakers with limited experience with legal texts. In other words, people with little experience of the legal genre will store 'general' aspectual preferences, as encountered in everyday exposure to language, and rely on them when choosing aspect (preferences of the unigram verb where no modal is present, and chunk preferences where a modal is used). However, speakers

with substantial experience of the legal genre (i.e. the authors and translators of the legal texts investigated here) should in principle store and access aspectual preferences specific to that genre, apart from preferences typical of general language. This will henceforth be referred to as *genre-specific chunking hypothesis*, which can be formulated as follows:

Repeated exposure to legal language results in genre-specific modal chunks being entrenched; the aspectual preferences of those genre-specific modal chunks facilitate the selection and processing of aspect in the given genre.

These genre-specific aspectual preferences may or may not be different from the general aspectual preferences. However, to be able to accurately determine whether aspectual preferences are a significant factor influencing the choice of aspect, it is necessary to consider the option that genre-specific modal chunks exist and may potentially have different aspectual preferences from general modal chunks. If genre-specific chunks exist but we only consider aspectual preferences of general chunks, we may conclude that no frequency effects are found to underlie the differences between translated and non-translated texts simply because we considered the wrong frequency information.

To sum up, it is argued here that aspectual preferences of modal chunks will explain more directly than any hypothesized translation universals why translators choose the perfective form more frequently than authors of non-translated Polish texts. It will be shown that the infinitives that follow *musieć* [must, have to] in the translated sample are more entrenched in the perfective form than in the imperfective form (when in chunk with the modal verb *musieć* [must, have to]) and that is why translators chose the perfective forms of those infinitives. In the non-translated

sample, on the other hand, the verbs that follow *musieć* [must, have to] are more entrenched in the imperfective form, and that is why authors chose the imperfective forms when they encountered those infinitives. In other words, the translated corpus contains verbs that have different aspectual preferences than the verbs in the non-translated corpus, affecting the aspectual choices made by translators.

Because the aspectual preferences of verbs have never been considered as a factor that constrains aspectual choice, we must first explore the psychological plausibility of this idea. This will be done in Part III where the chunking hypothesis is tested in two ways - through a corpus study and an experimental validation. After the psychological plausibility of aspectual preferences is confirmed, the aspectual preferences of modal chunks in the translated and non-translated *musieć* [must, have to] samples are compared. If the verbs contained in translated and non-translated samples have varying aspectual preferences, which align with aspectual choices, explaining the different linguistic behaviour of translators and authors of non-translated texts, we will confirm that comparable corpora lack the necessary comparability to support claims of translation universals.

5.4. Translation universals

If chunking cannot be said to underlie the differences in distribution of aspectual choices in the translated and non-translated *musieć* [must, have to] sample, two translation universals can be considered - explicitation and normalization.

5.4.1. Explicitation

The increased choice of perfective forms in the translated *musieć* [must, have to] sample could be interpreted as explicitation in Faber and Hjort-Pedersen's (2009)

understanding. According to them linguistic explicitation observed in translated texts is a reflection of the mental comprehension process that translators go through when translating a legal text. In other words, when translating a legal text, translators will explicitate certain information in their minds in order to comprehend the information contained in the source text, and this will lead them to subconsciously explicitate this information linguistically in the translated text. This can take the form of more specific lexical items, grammatical forms, or the distribution of meaning components over a number of morphemes or words (Lanstyák & Heltai, 2012, p. 112).

According to Janda's cognitively-inspired account of aspect, described in more detail in Section 3.2.2, perfective verbs are metaphorically associated with solid objects and imperfective verbs are metaphorically associated with fluid substances. Speakers that want to express the desirability for a given action to be completed will choose the perfective verbs, because the wholeness and manipulability of solid objects are mapped onto the desired wholeness and completeness of the act expressed by the verb. It could be argued that translators choose perfective forms because they subconsciously explicitate the desire for a given action to be performed. That is, when translating modal sentences, translators linguistically - by choosing the perfective form of the infinitive - explicitate the requirement to complete the action in question. This would be expected to happen less when authors of non-translated texts choose aspectual forms because authors do not go through the source text comprehension stage that translators go through.

5.4.2. Normalization

An alternative explanation to explicitation can also be proposed - it could be the case that the increased use of perfective forms in the translated *musieć* [must, have to] sample is the result of normalization.

In general texts, *musieć* [must, have to] is the most frequent of all necessity modals, but in legal texts, it is the least frequent - *powinien* [should] and *należy* [it is necessary] 'take over' the role of markers of necessity. It is clear that *musieć* [must, have to] is not a verb that is typically used to express necessity in non-translated legal texts, whereas *powinien* [should] and *należy* [it is necessary] are typical of the legal genre. This is shown in Figure 1 (with frequencies per million words provided).

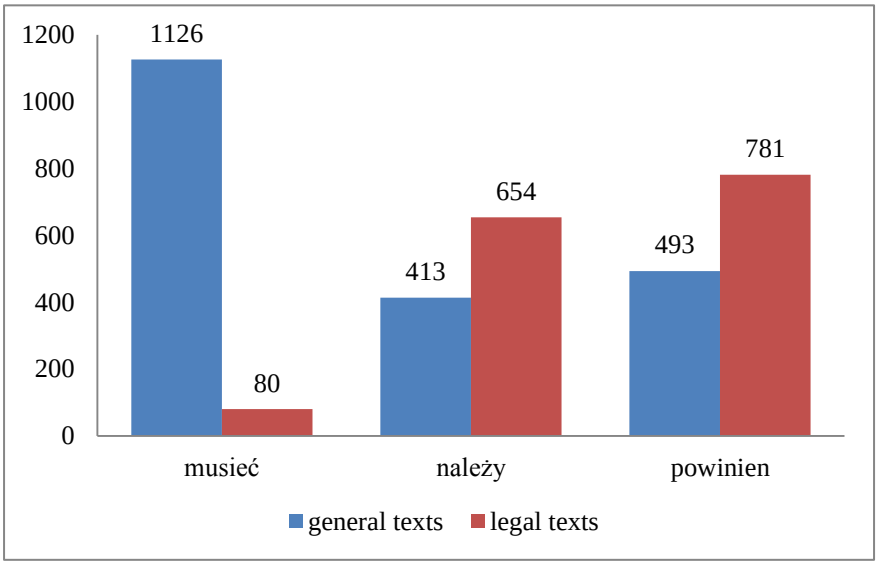

Figure 1: Frequencies per million words in general and legal non-translated texts

Table 24 shows that in the non-translated legal samples, both *należy* [it is necessary] and *powinien* [should] occur with the perfective form more frequently than with the imperfective form, whereas *musieć* [must, have to] is more commonly followed by an imperfective verb. That is, normally authors of non-translated legal texts use the

perfective form of a verb when they express a necessity to complete an action. An imperfective is used only with the modal *musieć* [must, have to], but this modal is atypical for legal texts, i.e. it occurs in them infrequently.

sample	impf frequency	pf frequency	p-value
musieć [must, have to]	68	38	0.00465197
należy [it is necessary]	20	95	0
powinien [should]	59	86	0.05274437

Table 24: Distribution of aspectual forms in non-translated samples

In the translated corpus, *musieć* [must, have to] is used significantly more frequently than in the non-translated corpus - there are 999.463 occurrences per million words, as compared to 75.966 occurrences pmw in the non-translated corpus. Translators are clearly more inclined to use *musieć* [must, have to] than authors of non-translated legal texts, perhaps because of the common occurrence of its English equivalent *must* (which would suggest a phenomenon called *shining through* - see Teich (2003) for more details).

Because legal authors use *musieć* [must, have to] so rarely, they choose the imperfective form of the verb that follows it, in line with the choice of aspect after *musieć* [must, have to] in general Polish - in the general National Corpus of Polish, the imperfective form occurs significantly more often than perfective when following *musieć* [must, have to]. Translators, on the other hand, use *musieć* [must, have to] significantly more frequently, and instead of choosing the imperfective form

as it would normally be done in general Polish, they choose the perfective form, as is done with other necessity modals in the legal genre. That is, they use a modal verb not typically used in legal texts but which is commonly used in everyday contexts, but instead of choosing the aspectual form that would normally be used in everyday contexts (i.e. the imperfective), they choose the other aspectual form, which incidentally is the form used with other necessity modals in legal contexts.

This may be a case of normalization - conforming to 'the patterns and practices that are typical of the target language, even to the point of exaggerating them' (Baker, 1996, pp. 176-177) - but at the level of genre-typical patterns rather than language-typical patterns. The typical pattern for necessity modals in the legal genre is to use the perfective form of the infinitive that follows the modal, so translators, who use *musieć* [must, have to] more frequently than authors, use it with the perfective infinitive like the other necessity modals in the legal genre, *powinien* [should] and *należy* [it is necessary]. That is, translators conform to the aspectual patterns and practices typical of the use of modal verbs in the legal genre - by using the perfective form - but override lexical practices - by using *musieć* [must, have to] to express necessity in legal texts, which is an unusual practice, probably resulting from shining-through of the English *must*.

5.5. Summary

To large extent, situation types explained why translators chose perfective verbs more often than authors of non-translated texts - we saw that the translated samples described situations which required the use of a perfective verb more often than the non-translated samples, influencing the choices made by translators and authors of non-translated texts. This suggests that comparable corpora lack the necessary

comparability. Some differences remained and we checked whether the distribution of polarity, modality types and state of affairs applicability can explain those, since previous studies have shown that the choice of aspect in modal contexts is subject to analogical mapping. No effect was found, however. It was therefore suggested that one more factor - chunking of modal verbs and infinitives that follow them - should be explored first, before turning to translation universals. The chunking hypothesis is explored in Part III.

PART III. CHUNKING HYPOTHESIS - INVESTIGATION

In order to explore whether differences in aspectual preferences in translated and non-translated corpora can explain differences in aspectual choices in translated and non-translated texts, discussed in Section 5.3, we combine corpus and experimental analyses. First, aspectual preferences of modal chunks are extracted from the National Corpus of Polish in order to test the chunking hypothesis. Inferential statistical techniques are used to assess how well aspectual preferences predict the choice of aspectual form in modal contexts in which the choice of aspectual form is not restricted by situation type. The corpus results are then validated by three experimental tasks - judgements of well-formedness, self-paced reading, and forced-choice. The data gathered from the tasks is analysed statistically. The results of the two analyses are compared and used to explain the differences between non-translated and translated texts, reported in Part II.

Methodological pluralism is employed here because language is a multi-faceted phenomenon that requires a multi-methodological approach to be fully understood (Arppe et al., 2010, p. 3). Cognitive linguists increasingly resort to a combination of

different types of linguistic data, typically involving corpus analysis and psycholinguistic experimentation (e.g. Arppe & Järvikivi, 2007; Bermel & Knittl, 2012; Divjak & Gries, 2008; to name but a few). Both methods have their limitations when used in isolation, but if combined, they can give the researcher multiple perspectives on the investigated phenomenon (Gilquin & Gries, 2009, p. 9). Triangulation of data types has become increasingly popular in Translation Process Research (TPR) - scholars combine eye-tracking, think-aloud protocols, keystroke logging, and EEG (e.g. Alves, Pagano, & Silva, 2010; Dragsted, 2010; to name but a few). Unfortunately, investigations of translation universals - the object of the current investigation - still rely solely on corpus-based analyses, with the notable exception of Halverson (2017), who combines corpus evidence with keystroke data as she revisits her Gravitational Pull Hypothesis that was proposed to explain translation universals from a cognitive-linguistic perspective.

Corpus and experimental data are widely used for several reasons. The naturalness of corpus data gives corpus analyses an advantage over the traditional use of artificial, introspective sentences, which can be made up by the researcher to fit his or her purposes and arguments. Introspective data can be abused by playing around with the context in order to make a sentence either grammatical or ungrammatical, whatever suits the researcher's agenda (Pullum, 2007, pp. 38-39). As a result, often questionable sentences are used as data for theoretical claims (Karlsson, 2009, p. 30). By analysing large collections of texts and applying statistical techniques, rather than looking at isolated examples, researchers can investigate complex patterns of language use, and describe and generalise over them in a reliable and consistent manner, not possible otherwise (Biber & Jones, 2009, p. 1287). Moreover, the ever-increasing sizes of corpora enable researchers to study a large number of examples

of a phenomenon, larger than could ever be possible to include in, for example, an experimental setting (Gilquin & Gries, 2009, pp. 8-9). Finally, corpora can be compiled in such a way that they are representative of the proportions of language varieties that an average person experiences, or of a particular variety (or varieties) that are of interest to the researcher[7] (Biber & Jones, 2009, p. 1288). Experimental methods allow the research to tap into two types of linguistic knowledge - implicit and explicit. This can be done through on-line and off-line experiments, respectively. The former, such as self-paced reading, force unconscious and automatic responses to stimuli, while the latter, such as intuition judgements, rely on conscious and controlled decision-making (Marinis, 2010, pp. 139-140). Experimental methods also allow the researcher to control confounding variables - something that is not always possible with corpus data (Gilquin & Gries, 2009, p. 9). Some corpora are annotated for linguistic and meta-linguistic information, but there are also raw corpora, which prevent the researcher from assessing the influence of factors such as the author's gender, age, origin, and many other contextual details that may prove important in interpreting the results. In a careful experimental design, influence of such factors can be controlled and accounted for. Finally, experimental data can be collected for languages and language varieties that for whatever reasons do not have their own corpora of natural texts, and therefore cannot be investigated by applying a corpus-based approach (Schütze & Sprouse, 2013, p. 29).

On their own, the two methods have certain limitations. Many corpus studies are based on the assumption that "frequency in text instantiates entrenchment in the cognitive system" (Schmid, 2000, p. 39). That is, by investigating patterns in language corpora, we can discover the way linguistic knowledge is structured in the

[7] It should be noted that the notion of representativness is a controversial one and it is not easy to define what constitutes a representative corpus, see Biber (1993) for details.

speaker's mind. Some argue, however, that although corpus data enable researchers to identify patterns that would otherwise go unnoticed, those patterns are in need of explanation, rather than having any explanatory power themselves (Biber, Conrad, & Cortes, 2004, p. 376). Therefore, raw frequency in corpora should not be blindly accepted as representative of salience in the mind (Arppe et al., 2010, p. 9). Moreover, significance testing, frequently used in corpus analyses, is sensitive to sample size. If the corpora we use are large enough, we will almost always obtain statistically significant results, but just because a statistical test tells us that the correlation between variables is not random, it does not necessarily imply it is not arbitrary, or that it is motivated or even predictable. Kilgarriff therefore argues that corpus data and statistical analysis can often lead to "unhelpful or misleading results" (Kilgarriff, 2005, p. 272). Experimental methods have also been questioned. Schütze (2005) carried out a survey of a variety of experimental designs in order to evaluate whether participants are likely to understand what researchers ask them to do, whether they are capable of doing it, and whether researchers can be confident that participants are actually doing what the researchers think they do when participants complete the set tasks. The assessed experimental designs include the Wug test, in which participants are asked to inflect nonce verbs, acceptability judgements, which ask participants to assess the acceptability/well-formedness/grammaticality of phrases or utterances, and the cloze test, used to assess proficiency levels of second language learners who are asked to, for example, fill in blanks in passages from which every n^{th} word has been removed. The author concludes that the instructions for participants are susceptible to being misunderstood, and the tasks can generally ask participants to do things, which they are incapable of doing (ibid., p. 477). Instead of asking participants to complete artificial and unfamiliar tasks, the author

argues that researchers should design their experiments to reflect how language is used for everyday purposes as closely as possible (ibid.). Hilpert (in Arppe et al., 2010) also expresses his concerns about the ecological validity of experimental designs, i.e. to what extent does the experimental setting represent a real life linguistic experience.

Combining different types of evidence can mitigate the limitations of using each method in isolation (Gilquin & Gries, 2009, p. 9). For example, by validating corpus results with experimental data, we ensure that the corpus results are not due to the size of the corpus. Also, by testing native speakers, we can evaluate the psychological reality of the corpus analysis and validate the corpus-to-cognition assumption that underlies many corpus-linguistic studies. By statistically analysing naturally occuring corpus data and comparing the results with experimental results, we also mitigate the risks of the experimental study's limited ecological validity and its effect on the speakers' responses. Moreover, corpus and experimental paradigms address the linguistic behaviour and knowledge from different perspectives and reflect different linguistic processes (Arppe et al., 2010, p. 3). Corpora contain naturally occurring language and therefore enable researchers to investigate the process of language production. Corpus analysis can tell us what is probable and improbable in a language (Gilquin & Gries, 2009, p. 15), while off-line experimental data can reveal what speakers find to be an acceptable linguistic output, or in other words, what is possible and impossible in language (Thráinsson et al., 2007, p. 120). Finally, on-line experimental data enable researchers to investigate the processing of linguistic input, as opposed to language production. Combining different types of evidence can therefore contribute to a better understanding of language by giving the

researcher multiple perspectives on the investigated phenomenon (Arppe et al., 2010, p. 3; Gilquin & Gries, 2009, p. 9).

Combining different types of evidence comes with its own problems, however. As mentioned above, the different methods measure different types of linguistic behaviour and knowledge (e.g. production, processing, introspection). Our understanding of these different types of resulting data, what they tell us, and how they relate to each other is still limited (Arppe et al., 2010, pp. 7-8). Researchers should therefore be cautious when comparing the results of corpus and experimental analyses, and take into consideration the different origins and characteristics of each method (Arppe & Järvikivi, 2007, p. 132). Nevertheless, Wasow & Arnold (2005, p. 1495) argue that language researchers should employ multiple types of evidence, as is done in other scientific fields. Methodological pluralism will therefore be employed in the current investigation in order to explore the role of chunking in the aspectual differences observed in Part II. Chapter 6 discusses the corpus study while Chapter 7 - the experimental study.

Chapter 6. Corpus analysis

The role of chunking in aspect assignment in translated and non-translated texts will first be explored by analysing language corpora. The aim of the corpus analysis is therefore to establish to what extent aspectual preferences of modal chunks correlate with the choice of aspect in the samples analysed in Part II. In order to do that, aspectual preferences are extracted from the National Corpus of Polish for all infinitives tied to the modal verbs in the translated and non-translated samples. Then, logistic regression models are fitted to see how well the four variables predict the choice of aspect.

The two hypotheses formulated in Section 5.3 will be tested:

General chunking hypothesis: *Modal verbs and infinitives that follow them are entrenched in memory as chunks and the aspectual preferences of those modal chunks facilitate the selection and processing of aspect.*

Genre-specific chunking hypothesis: *Repeated exposure to legal language results in genre-specific modal chunks being entrenched; the aspectual preferences of those genre-specific modal chunks facilitate the selection and processing of aspect in the given genre.*

6.1. Method

6.1.1. Source of data

Aspectual preferences of the infinitives that follow modal verbs in the translated and non-translated samples are extracted from National Corpus of Polish (NKJP). The legal sub-corpus of NKJP was used to extract samples of modal verbs for the purposes of the corpus analysis performed in Chapter 4. In the current part of the investigation, both the general NKJP corpus and the legal sub-corpus will be used. More information about NKJP can be found in Chapter 4.

6.1.2. Procedure

In order to test the two chunking hypotheses, four types of aspectual preferences are extracted from NKJP:

(i) General unigram - the aspectual preferences of lexemes as they occur on their own in the general sub-corpus of the National Corpus of Polish.

(ii) Legal unigram - the aspectual preferences of lexemes as they occur on their own in the legal sub-corpus of the National Corpus of Polish.

(iii) General chunk - the aspectual preferences of lexemes as they occur in a modal chunk in the general sub-corpus of the National Corpus of Polish.

(iv) Legal chunk - the aspectual preferences of lexemes as they occur in a modal chunk in the legal sub-corpus of the National Corpus of Polish.

Comparing the predictive power of general unigram and general chunk aspectual preferences by fitting binary logistic regression to the data will enable us to confirm or disprove the general chunking hypothesis: if the general chunk preferences perform significantly better than unigram preferences, we will have found some support for the chunking of modal verbs and the infinitives that follow them. Comparing the predictive power of general chunks and legal chunks will allow us to confirm or disprove the genre-specific chunking hypothesis. Unigram aspectual preferences are checked to ensure that they do not predict aspectual choice better than legal chunk aspectual preferences. If the legal chunk preferences perform significantly better than the general chunk and legal unigram preferences, we will have found some support for genre-specific chunking. The aspectual preference that is shown to predict aspectual choices best will then be used to explain the differences in aspectual choices between translators and authors of non-translated texts.

The general aspectual preferences (unigram and chunk) are extracted from the general portion of NKJP, while the legal aspectual preferences are extracted from the legal sub-corpus of NKJP. For the unigram preferences, all instances of the verb's aspectual forms are retrieved (conjugated and infinitival). Because these instances include instances in which the verb follows a modal, we subtract the number of times a given verb occurs in a modal chunk from the number of its overall occurrence, arriving at the unigram frequencies.

The difference between the number of perfective instances and imperfective instances of the verb are checked for significance by means of a Chi-squared test - the observed corpus frequencies are compared to expected frequencies. If the p-value is lower than α=0.05 then the more frequent aspectual form is considered to be the preferred one for that verb in the given context (unigram/chunk/general/legal). If the p-value is higher than α=0.05 then the verb is considered not to have an aspectual preference in the given context[8].

Each of the four aspectual preferences is annotated as a separate variable in the data set and has four levels:

(i) PF if the preferred form is perfective;

(ii) IMPF if the preferred form is imperfective;

(iii) IMPF/PF if a given verb has no preferred form; i.e. perfective and imperfective forms are equally frequent;

(iv) NA if there are no attested forms in the corpus, i.e. the corpus does not contain any instances of a given verb.

[8] Some researchers would argue that a Bonferroni correction should be applied here due to the large number of tests being performed. There seems to be no consensus as to what constitutes a single data set for the purposes of statistical testing. In the case of aspectual preferences, the α can remain at its conventional level. This is because each comparison between imperfective and perfective forms is done on a different set of sentences extracted from the corpus - for example, general unigram aspectual preferences are calculated on all instances of unigram verb extracted from the general section of NKJP, while chunk general aspectual preferences are calculated on all instances of the verb in modal sentences, extracted from NKJP separately, and not overlapping. That is, each test is performed on a different data set, unless one considers the corpus to be a single data set. In that case, a Bonferroni correction would have to be applied. However, it would be difficult to decide how to calculate the Bonferroni correction - should the standard α be divided by the total number of tests performed in all samples (translated and non-translated), or perhaps by the number of tests for each verb? In case of the former, there would be thousands of tests (3,832 to be precise), resulting is such a low α=0.00001305 that hardly anything would be considered significant. In case of the latter, the α would be more reasonable at 0.0125, although it would mean that we treat all occurrences of a single verb (unigram+chunk+general+legal) - rather than the whole corpus - as one data set. To avoid setting an arbitrary and overly harsh α, and because it is reasonable to treat each of the aspectual preference tests as calculated on a separate data set, the conventional α=0.05 is used

Once the preferred form for every verb encountered in all translated and non-translated samples are annotated, logistic regression is fitted to the data to establish which of the aspectual preferences is the best predictor of aspectual choices. The general chunking hypothesis will be tested by fitting models to all data at once (translated and non-translated); the genre-specific chunking hypothesis will be tested by fitting separate models to the translated and non-translated data. This is because the aims of the two hypotheses are different. The general chunking hypothesis is to assess whether unigram aspectual preferences or chunk aspectual preference predict aspectual choice best. If the latter, we will be justified in claiming that modals and verbs that follow them form chunks and are entrenched in the speaker's memory as such. We assume all speakers will process modals and verbs that follow them as chunks, regardless of their level of experience with the legal genre so we can test this hypothesis on all data jointly. For the genre-specific chunking hypothesis, however, experience with the legal genre plays a crucial role. As mentioned in sections 4.1.1 and 5.3, we suspect that translators and non-translators have different experience with the legal genre: it is likely that translators were less experienced than authors of non-translated texts so we expect their aspectual choices to be predicted best by different aspectual preferences. As such, we separate the data from the translated and non-translated corpora to see whether genre-specific and general chunks provide different predictions for the different sets of data, thus supporting the genre-specific chunking hypothesis.

Before we move on to the regression results, a caveat is in order. Unigram aspectual preference of a verb was established based on the number of occurrences of one aspectual form of that verb in comparison with the other aspectual form. This is the simplest way to operationalize frequency, and may not necessarily be the most

psychologically plausible - it only reflects the frequency with which a stimulus is repeated in the environment and does not reflect the way brain makes use of frequency of occurrence (Divjak & Caldwell-Harris, 2015, p. 57). Contextual diversity, semantic distinctiveness, conditional probabilities, have proven to be more powerful than simple frequency of repetition (ibid., pp. 58-60). Nevertheless, frequency of repetition, i.e. frequency with which a stimulus is repeated in the environment, is still a strong predictor of linguistic behaviour (ibid.) and will therefore be used here. Moreover, chunk aspectual preferences can be treated as a form of a contextualised frequency because they take into consideration the frequency of each aspectual form of the verb, given the modal verb it follows (i.e. its linguistic context).

It should also be mentioned that a number of verbs in the analysed samples have no preference for any aspectual form; others have occurrence so rare that frequencies could not be extracted. If aspectual preferences indeed guide the speakers' choice of aspectual form and we explain translators' aspectual choice with reference to aspectual preferences, then we need to consider what happens when speakers come across verbs that have no aspectual preference, or verbs so rare that frequency information is most likely not entrenched enough to facilitate processing. In those cases, we may hypothesize that abstract schemas guide aspectual choice; schemas are the commonalities abstracted from repeated exposure to patterns that may on the surface seem different but nevertheless share certain organisational features (Langacker, 2000, p. 4). Schemas that are acquired through exposure to language use are used as templates when novel expressions are encountered (Langacker, 2008, p. 168). It is therefore reasonable to assume that through exposure to modal chunks, a MODAL(ASPECT) schema is abstracted, since modal verbs in Polish are almost always

followed by an infinitive in one or the other aspectual form. For example, analysis of the National Corpus of Polish shows that *musieć* [must, have to] occurs with imperfective infinitives significantly more frequently than with perfective infinitives ($p<0.0001$). This may suggest that a general MUSIEĆ(IMPF) schema is entrenched in the speakers' memory and guides the speakers' choice of aspectual form when no other clues are available. In other words, it may be the case that the imperfective form is the default form chosen for *musieć* [must, have to] chunks with no aspectual preference. We will later fit models to the *musieć* samples in order to establish whether this is likely to be the case, i.e. whether there are significant odds that the imperfective form is selected for verbs in modal chunks with *musieć* that have no aspectual preferences.

6.1.3. Statistical analysis

The two chunking hypotheses are tested by fitting binary logistic regression in R (version 3.3.1), using the {MASS} (Venables & Ripley, 2002) and {rms} (Harrell, 2001) packages. Binary logistic regression allows the researcher to statistically model the effect of one or several predictors - the independent variables - on a binary response variable - the dependent variable (Speelman, 2014, p. 488). The response is categorical and has two levels - in our case, imperfective or perfective - while predictors can be either numerical or categorical. The method assumes that both levels of the response variable are possible, regardless of the configuration of the variables, and that they are mutually exclusive, i.e. they perform the same linguistic function and there is no third option that could be used instead (ibid., pp. 489-490). The technique calculates the probability of the response - here, the choice of imperfective or perfective - given the linguistic context - here, the aspectual

preferences of infinitives that follow modal verbs. In other words, we are calculating the probability of a given aspectual form of an infinitive being chosen, given the aspectual preference of that infinitive.

The aim of regression analysis is to find a model that fits the data best by producing the least unexplained variation (Crawley, 2015, p. 5). That is, in fitting models we are looking for predictor variables that will most accurately predict the response variable. Fitting models (or model building) involves variable selection and model comparison, which can be done in many ways, partly depending on the type of model and the purpose of the modeling. There are two types of models - nested and non-nested. Two models are nested if one model can be reduced to the other model by removing variables, for example:

1. (a) aspect ~ general unigram aspectual preference + modality_type + soa

   (b) aspect ~ general unigram aspectual preference + modality_type

Model 1b is a reduced version of model 1a because all of the predictor variables are the same, with the last variable - soa - removed. Two non-nested models cannot be reduced to one another by removing variables. For example:

2. (a) aspect ~ general unigram aspectual preference + modality_type + polarity

   (b) aspect ~ general chunk aspectual preference + modality_type + polarity

The two models above have the same number of predictors, but one of the predictors differs - model 2a has unigram aspectual preference, while model 2b has chunk aspectual preference as a predictor. Nested models are used to find the best combination of variables that can explain a response, while non-nested models are

used to test rival theories that are supposed to explain the same phenomenon (Pesaran & Weeks, 1999). In other words, with nested models we are looking for a combination of factors that jointly explain the response variable. With non-nested models, we are testing alternative explanations for the response variable, with the aim of choosing one over the other.

In order to test the general chunking hypothesis, we compare how well general chunk aspectual preferences and general unigram aspectual preferences predict the choice of aspect. We will therefore fit two non-nested models because we are testing two alternative explanations for the response variable. In order to test the genre-specific chunking hypothesis, we compare how well general chunk aspectual preferences and legal chunk aspectual preferences predict the choice of aspect, but we also check the performance of legal unigram aspectual preferences to rule out their role in aspect assignment. Again, non-nested models are fitted.

There are various ways to compare the predictive powers of models. We can look at the AIC (Akaike's Information Criterion) scores - the lower the AIC, the better the model is at predicting the outcome. If the difference between the two models is less than 2, then both models are by and large equivalent. If the difference is between 4 and 7, then there is considerably less support for the higher AIC model. Finally, if the difference between models is more than 10, then there is essentially no empirical support for the model with the higher AIC (Burnham & Anderson, 2002, p. 70). In other words, if the AIC of one model is lower than the AIC of another model by at least 10, then we can say with certainty that the lower AIC model is a significant improvement over the higher AIC model.

R^2 can also be used to compare models - the measure tells us how much of the variation in the response variable can be explained by taking into account the predictors. If the predictors do not correlate with the response variable, R^2 will be close to zero; the better the model, the less unexplained variation there is and the closer to 1 the R^2 is (Baayen, 2008, pp. 88-89). Here, it can be used to assess how much of the variation in the choice of aspectual form can be explained by taking into account the four types of aspectual preferences of the infinitives. By comparing the R^2 scores of all four types of aspectual preferences, we will be able to choose the one that significantly explains the most variation and can therefore be considered the best predictor of aspectual choice. It should be noted that the use of R^2 measures with logistic regression models - and the use of R^2 as a goodness-of-fit measure in general - has been disputed. This is because R^2 measures are based on a comparison of the values predicted from the fitted models with the values from a null model - Hosmer and Lemeshow (2000, p. 164) argue that a true measure of fit should be instead based on a comparison of the observed values with the predicted values from the fitted model. Nevertheless, R^2 values may prove helpful when evaluating competing models fitted to the same data (Hosmer & Lemeshow, 2000, p. 167) and will therefore be reported in the current investigation. It should be noted that compared to R^2 values from linear regression models, R^2 values from logistic regression models are comparatively smaller and may mislead an audience that is used to linear R^2 values into thinking that the logistic model explains little of the variation (ibid.). That is, the low R^2 values reported here should not be interpreted as indicating little explained variation.

Harrell's C is considered more reliable than R^2. It is commonly reported with the results of logistic regression - unlike R^2, it calculates the index of concordance

between the predicted probability and the observed response, giving an estimate of the predictive power of the model. If the C value is below 0.5 then we know that the predictions are random, while values of 1 mean perfect predictions; anything above 0.8 indicates that the model may have a real predictive capacity (Baayen, 2008, p. 204).

Finally, the p-value of the model can be looked at to assess whether the predictor variables together are explanatory (ibid.). The model with predictor variables is compared to a null model with intercept only, and if the p-value is lower than $\alpha=0.05$ then we know that the variables contribute significantly in explaining the variability in the data.

All of the measures will be used to compare the models. P-values will be consulted to establish whether a given variable significantly correlates with the choice of aspect. Then, AIC scores and C scores will be looked at - the model with the lowest AIC and highest C will be chosen as the best performing. R^2 will also be looked at to ensure there are no discrepancies, but decisions about best performance will be made based on AIC and C scores.

6.2. Corpus results

The aim of the corpus analysis was to investigate the role of chunking in aspect assignment in modal contexts that do not impose which aspectual form is required: we wanted to establish to what extent aspectual preferences of modal chunks correlate with the choice of aspect. We did that by extracting aspectual preferences from the National Corpus of Polish for all infinitives tied to the modal verbs in the translated and non-translated samples. Then, we fitted logistic regression models to see how well the aspectual preferences as attested in the corpus predict the choice of

aspect in the translated and non-translated samples. The results - discussed below - support both the general chunking hypothesis and the genre-specific chunking hypothesis.

The R code used to fit the regression model is available in Appendix 3; all data sets used in the analyses are available to download online[9].

6.2.1. General chunking hypothesis

The general chunking hypothesis states that 'modal verbs and infinitives that follow them are entrenched in memory as chunks and the aspectual preferences of those modal chunks facilitate the selection and processing of aspect.' In order to test the hypothesis, two models were fitted, each with a different predictor - chunk aspectual preferences and unigram aspectual preferences:

(1) aspect ~ general chunk aspectual preference

(2) aspect ~ general unigram aspectual preference

The models were compared by looking at the measures described in Section 6.1.3 - AIC, R^2, Harrell's C. Table 25 contains the results.

aspectual preference	AIC	R2	C	P-value
general chunk	1639	0.309	0.755	<0.0001
general unigram	1986.6	0.030	0.564	<0.0001

Table 25: Comparison of general chunk and general unigram aspectual preferences

The model with the lowest AIC and p-value, and highest C and R^2 is considered best. The p-values of both models are very low, suggesting that both types of aspectual preferences make a significant contribution to explaining the choice of aspect in the

[9] https://drive.google.com/open?id=0Bz9KVHFRWI3NclB4aWRRY2poQ28

samples. However, it seems that chunk aspectual preferences predict the aspectual version of the modal chunk more accurately than unigram aspectual preferences - the AIC of the chunk model is lower than the unigram model by 347.6 scores. The chunk model also has higher Harrell's C (0.755), which is very close to the crucial value of 0.8 that implies some real predictive capacity. Finally, the R^2 of the chunk model is ten times higher than that of the unigram model, although it remains relatively low at 0.309, suggesting that legal chunk preferences on their own explain only around 30% of the variance in the data. The models coefficients suggest that if chunk aspectual preference is 'perfective' then all other things being equal, the probability of the perfective form being chosen increases significantly (estimate=2.5224, std. error=0.1527, p<0.0001). The probability of perfective being chosen also significantly increases if the infinitive has no aspectual preference (estimate=0.8709, std. error=0.1981, p<0.0001).

To sum up, chunk aspectual preferences perform much better as a predictor of aspectual choice than unigram preferences. This result supports the general chunking hypothesis, according to which modal verbs and the infinitives that follow them form chunks and that the aspectual preferences of those chunks predict (to some extent) the aspectual form of the infinitive in a modal chunk.

6.2.2. Genre-specific chunking hypothesis

The genre-specific chunking hypothesis stated that 'repeated exposure to legal language results in genre-specific modal chunks being entrenched; the aspectual preferences of those genre-specific modal chunks facilitate the selection and processing of aspect in the given genre.' In order to test the hypothesis, three models were fitted, each with a different predictor - general chunk aspectual preferences,

legal chunk aspectual preferences, and legal unigram aspectual preferences (general unigram preferences were already ruled out as a suitable predictor in the previous section):

(1) aspect ~ general chunk aspectual preference

(2) aspect ~ legal chunk aspectual preference

(3) aspect ~ legal unigram aspectual preference

Here, separate models are fitted to the translated and the non-translated data - as mentioned in Section 4.1.1, we suspect that translators of the texts in PELCRA had little experience with legal genre, as compared with the authors of the texts in NKJP. We therefore expect the aspectual choices in the translated samples will be more accurately predicted by different type of aspectual preferences (general chunk) than the aspectual choices in the non-translated samples (legal chunk).

Table 26 contains the results for the non-translated data and Table 27 - for the translated data.

aspectual preference	AIC	R^2	C	P-value
general chunk	824.93	0.384	0.786	<0.0001
legal chunk	779.34	0.437	0.827	<0.0001
legal unigram	1043.8	0.074	0.621	<0.0001

Table 26: Models fitted to the non-translated data

We can see that the AIC of the non-translated legal chunk model (779.34) is much lower than that of the general chunk model (824.93) and that of the legal unigram model (1043.8), suggesting that legal chunk preferences predict the choice of aspect much better than general chunk and legal unigram preferences. R^2 suggests that legal chunk aspectual preferences predict 43% of all aspectual choices - more than the

other two models. Finally, the C value is higher than the crucial 0.8 so the model has a real predictive power. This supports the genre-specific chunking hypothesis and suggests that speakers with experience in legal language have legal chunks entrenched in memory and those chunks - rather than the more general ones - are retrieved when producing legal language. More specifically, if legal chunk aspectual preference is 'perfective', then all other things being equal, the probability of the perfective form being chosen increases significantly (estimate=3.6821, std. error=0.2585, p<0.0001). There are also increased odds of the perfective being chosen if the infinitive has no aspectual preference (estimate=1.8794, std. error=0.2507, p<0.0001).

aspectual preference	AIC	R^2	C	P-value
general chunk	811.82	0.228	0.717	<0.0001
legal chunk	864.87	0.137	0.675	<0.0001
legal unigram	898.31	0.076	0.622	<0.0001

Table 27: Models fitted to the translated data

The AIC of the translated legal chunk model (864.87) is higher than the AIC of the general chunk model (811.82) but lower than that of legal unigram preferences (898.31). The general chunk model has the highest R^2 and C of the three models too. This suggests that aspectual choices of translators are better predicted by general chunk aspectual preferences than legal chunk aspectual preferences and legal unigram aspectual preferences. This seems to confirm translators of the texts in PELCRA had less experience with legal language that would probably be desired. It also seems to confirm the genre-specific chunking hypothesis, especially when we compare these results to those obtained from the non-translated samples.

All in all, the corpus data seems to provide solid support for genre-specific chunking hypothesis, according to which, speakers with various usage histories have various

types of chunks entrenched in memory and activate different frequency information depending on the context of use.

6.2.3. Bootstrap validation

Bootstrap validation of the models that best predict aspectual choice in the translated and non-translated samples is necessary to ensure that the models are accurate and not overfitted (Baayen, 2008, p. 205). The procedure estimates the performance of a given model on new data by drawing repeated samples from the original data and re-calculating the model statistics, including R^2. If the bootstrapped R^2 is considerably lower than the original model's R^2 then it is likely that the original model is overfitted, i.e. the calculated coefficients are too extreme (ibid.). The validate() function in R was used to perform this procedure, as suggested in Baayen (ibid.). For both models, the bootstrapped R^2 is very close to the original models' R^2, which means that both original models were accurate (see Table 28).

data	variable	original R^2	bootstrapped R^2
non-translated	legal chunk preference	0.437	0.432
translated	general chunk preference	0.228	0.220

Table 28: Original and bootstrapped R^2

6.3. Conclusions

The choice of aspect in translated and non-translated texts seems to be predicted by different types of aspectual preferences. In non-translated samples, the legal chunk aspectual preference is the best predictor, while in translated samples, it is general chunk aspectual preferences that perform best. It is clear that in both cases, aspectual preferences of verbs in chunk, rather than their unigram aspectual preferences, correlate with the choice of aspect more closely. That is, general chunking

hypothesis that modal verbs and the infinitives that follow them form chunks, and that the aspectual preferences of those chunks correlate with the choice of aspect, seems to be confirmed. Moreover, the genre-specific chunking hypothesis that genre-specific chunks are formed in those speakers that have substantial experience with a given genre also seems to be confirmed.

This has two important implications. Since aspectual preferences of modal chunks predict choice of aspect better than aspectual preferences of the unigram verb, it can be concluded that modal chunks exist as prefabricated units in the minds of speakers. That is, when exposed to modal utterances, it is likely that speakers process the modal verb and the infinitive that follows it as a non-compositional, prefabricated unit, rather than composing the phrase on the spot by separately retrieving the modal verb and adding the verb in its relevant aspectual form separately. Similar process would take place in production - when building a modal utterance, it is likely that speakers retrieve the modal verb together with the required infinitive, rather than retrieving the two separately. Moreover, the results would suggest that the aspectual preferences of those chunks are also entrenched, either as part of information associated with the modal chunk for a given verb, or even as two separate chunks for each aspectual form of the verb. These aspectual preferences seem to guide the choice of aspect to some extent.

Finally, the results seem to suggest that repeated exposure to legal language could result in genre-specific modal chunks being formed and guiding the choice of aspect in legal contexts. It was mentioned in Section 4.1.1 that translators of the texts from PELCRA are likely to have had little experience translating legal language. The fact that aspectual choices in the translated samples are better predicted by general chunk

aspectual preferences than legal chunk aspectual preferences, while the choices of authors of non-translated texts are better predicted by legal chunk aspectual preferences, would seem to fit in well with this argument and suggest that linguistic experience indeed results in different linguistic structures being accessed when processing linguistic input or producing utterances.

One of the limitations of corpus analyses is the lack of certainty that the patterns observed in corpora reflect how language is structured in the speakers' minds. Some argue that linguistic preferences, cognitive functions, and processes associated with language can be investigated using corpora (e.g. Schmid, 2000). Others, on the other hand, doubt that corpus data can explain how linguistic structures are represented in the speakers' minds (e.g. Biber, Conrad & Cortes, 2004). In order to mitigate any risk that the results of the corpus analysis do not reflect how speakers really make aspectual choices, the two chunking hypotheses are also validated with a series of psycholinguistic experiments.

Chapter 7. Experimental analysis[10]

Experimental validation was performed to confirm the plausibility of corpus results. Three experimental tasks were performed by 45 native speakers of Polish. The data gathered from the tasks was analysed statistically. Two hypotheses were tested:

General chunking hypothesis: *Modal verbs and infinitives that follow them are entrenched in memory as chunks and the aspectual preferences of those modal chunks facilitate the processing of aspect.*

Genre-specific chunking hypothesis: *Repeated exposure to legal language results in genre-specific modal chunks being entrenched; the aspectual preferences of those genre-specific modal chunks facilitate the processing of aspect in the given genre.*

In order to (dis)prove the two hypotheses, participants were presented with imperfective and perfective versions of infinitives in modal chunks and their responses to those forms were measured in three different ways: perceived well-formedness of the aspectual version (a judgement task), the time it takes to read them (a self-paced reading task), and which form is chosen more frequently (a forced-

[10] The experimental studies received ethical approval from the School of Languages and Cultures at the University of Sheffield.

choice task). This way, if participants' responses to one aspectual version of the infinitive in a chunk are more favourable than the responses to the other aspectual version of the same infinitive in the same chunk, then we can establish which aspectual preference of the infinitive facilitated that response. Section 7.1 discusses the methodology in more detail. Section 7.2 contains the results.

7.1. Method

The three tasks performed by participants were: a judgement task, a self-paced reading task, and a forced choice task[11]. The tasks belong to two different groups of psycholinguistic experiments - on-line and off-line - which tap into different types of linguistic knowledge. The judgement task and the forced-choice task are both considered off-line experiments because they rely on conscious and controlled decision-making (Marinis, 2010, pp. 139-140). As such, they tap into the speaker's explicit linguistic knowledge and metalinguistic abilities. Judgement tasks tell us which of the linguistic variants available to speakers are found more appropriate in a given context. In the current investigation, participants were asked to rate the well-formedness of modal chunks embedded in sentences. The ratings given to the different versions of modal chunks were then compared to see if aspectual preferences affected the perceived well-formedness of modal chunks. Forced-choice tasks tell us which of the linguistic variants available to speakers are preferred in language production. In the current investigation, participants were asked to choose one of two modal chunks provided to them in order to fill gaps in sentences. The choices were then compared to see if aspectual preferences play a role in the choice of linguistic variants. The self-paced reading is on-line because it forces unconscious

[11] Initially, the three tasks were: judgement task, self-paced reading, and proofreading. Due to a glitch in the keystroke logging software, the proofreading task had to be replaced by the forced-choice task.

and automatic responses to stimuli, tapping into the speaker's implicit linguistic knowledge (Marinis, 2010, p. 140). Self-paced reading tasks tell us which of the linguistic variants available to speakers causes greater cognitive load in the processing of linguistic input. In the current investigation, participants read sentences divided up into chunks, while themselves controlling the rate of presentation of the next chunk in the sentence. The reading times were then compared to see if aspectual preferences facilitate the processing of modal chunks.

7.1.1. Participants

Three groups of participants performed the three tasks, with fifteen participants per group[12]. The groups differ in terms of the usage histories of their members, but there are also some differences in terms of the gender make-up and educational background.

(1) Naive native speakers (NS) - native speakers of Polish with no knowledge of linguistics or translation, and no formal training or experience in the fields. Participants were recruited through personal contacts and via announcements on social media. Eight participants are female and seven are male. Five participants are based in the UK and ten in Poland. They are between 28 and 58 years old, the average age being 38.9. Nine participants have completed higher education and six finished secondary education.

(2) Trainee translators (TT) - native speakers of Polish who are current or very recent students on university-level courses in translation with no professional experience in translation (except experience gained as part of the course). Participants were

[12] Participants were offered a small financial incentive due to the time commitment required to complete all tasks. The funding was kindly provided by the University of Sheffield's Prokhorov Foundation research fund.

recruited at British and Polish universities that provide translation courses, including the University of Sheffield, University College London, and the University of Rzeszów. Twelve females and three males participated. Seven of them are based in the UK and eight in Poland. Ages range between 23 and 34, with an average age of 26. Each participant either already has a higher education diploma or is working towards one.

(3) Professional translators (PT) - native speakers of Polish who are English-into-Polish translators with formal university-level qualifications in translation and at least five years' part-time experience translating legal and administrative texts. Participants were recruited in the UK and Poland by directly contacting translators listed on publicly available registers (Chartered Institute of Linguists and Institute of Translation and Interpreting in the UK, and the register of sworn translators available on the Polish Ministry of Justice website in Poland). Fourteen of the recruited translators are female and only one is male. Seven translators are based in the UK and eight are in Poland. They are between 27 and 60 years old, with an average age of 41. They are therefore the oldest group of participants, followed by naive native speakers and trainee translators.

It should also be noted that the majority of participants are bilingual and it could therefore be argued that their knowledge of a second language may affect how they use their first language (see Section 1.2.1.1 for more details) and that this should be reflected in the experimental design in order to prevent any unwanted influence on the responses. However, whether a speaker is bilingual or not, the frequency with which a given aspectual form occurs in its natural environment (here, the National Corpus of Polish) does not change. That is, the perfective form of verb X is still

more frequent than that verb's imperfective form, even if the speaker has knowledge of another language. Aspectual form of modal chunks is the only variable manipulated in the experimental stimuli so if one group of participants sees a modal chunk with the less frequent aspectual form while another group sees the same modal chunk with the more frequent aspectual form, and the responses are consistently different, there will be no doubt that frequency was the variable affecting the processing of that chunk. Moreover, there seems to be no evidence in the literature on aspect to suggest that bilingualism affects the way it is processed. It is therefore argued that the experimental design leaves no room for interference from the participants' bilingual knowledge and it therefore does not have to be addressed further.

7.1.2. Stimuli[13]

Authentic sentences from the corpus analysis performed in Chapter 4 and additional authentic sentences extracted from the National Corpus of Polish were used. The two hypotheses were tested using different experimental sentences - the general chunking hypothesis compares general chunk aspectual preferences with general unigram aspectual preference, while the genre-specific chunking hypothesis compares general chunk aspectual preferences with legal chunk aspectual preferences. The aspectual preferences of modal chunks used for the general chunking hypothesis must therefore differ from the aspectual preferences of modal chunks used for the genre-specific chunking hypothesis.

[13] All selected stimuli can be found in Appendix 2.

7.1.2.1. General chunking hypothesis

The general chunking hypothesis can be confirmed by showing that chunk aspectual preferences of infinitives in modal chunks explain participants' responses to modal chunks better than the unigram aspectual preferences of the infinitives. We therefore need to select modal chunks in which the infinitives have chunk aspectual preferences different from unigram aspectual preferences. For example, when *nadawać/nadać* [to grant] occurs in a modal chunk with *móc* [can, be able to], it is more frequent in the perfective form. At the same time, in its unigram form (i.e. without the modal), *nadawać/nadać* is more frequent in the imperfective form (see Table 29). Henceforth, the more frequent aspectual form of an infinitive will be called its *preferred* form, while the less frequent aspectual form of that infinitive will be called *dispreferred*. For example, the chunk-preferred aspectual version of *nadawać/nadać* is perfective while its chunk-dispreferred aspectual version is imperfective. At the same time, the unigram-preferred aspectual version of *nadawać/nadać* is imperfective while its unigram-dispreferred aspectual version is imperfective.

aspectual preference	pf frequency	impf frequency	p-value	preferred aspect	dispreferred aspect
unigram	13713	20517	0	IMPF	PF
general chunk	238	177	0.00275	PF	IMPF

Table 29: Aspectual preferences of *nadawać/nadać*

Infinitives like *nadawać/nadać* will enable us to test the general chunking hypothesis: if the chunk-preferred version of *móc nadawać/nadać* (i.e. the perfective) is rated more favourably, read more quickly, and chosen more frequently

than the unigram-preferred version (i.e. the imperfective), it may suggest that chunk aspectual preferences facilitated the processing of the modal chunk, thus supporting the general chunking hypothesis. In other words, we are testing whether responses to an infinitive in a modal chunk differ depending on the aspectual form in which participants see it: the chunk-preferred form or the unigram-preferred form.

Only sentences that match the aspectual preferences listed in Table 30 are selected for experimental validation of the general chunking hypothesis. That is, sentences that have general chunk and legal chunk preference for imperfective must have general unigram and legal unigram preference for perfective. Sentences that have general chunk and legal chunk preference for perfective must have unigram general and unigram legal preference for imperfective.

general chunk	legal chunk	general unigram	legal unigram
impf	impf	pf	pf
pf	pf	impf	impf

Table 30: Aspectual preferences of stimuli - general chunking hypothesis

A total of sixteen utterances matching the criteria in Table 30 were selected from the corpus samples extracted for the analysis in Chapter 4; these were used in the judgement task and in the self-paced reading task. A total of twenty-six utterances matching the criteria were extracted from the National Corpus of Polish - these were used in the forced-choice task[14].

[14] Initially, the three tasks were: judgement task, self-paced reading, and proofreading. Due to a glitch in the keystroke logging software, the proofreading task had to be replaced with the forced-choice task. New set of experimental sentences had to be selected - all sentences selected from the corpus samples extracted for the analysis in Chapter 4 were already seen by participants in the judgement task, self-paced reading, and proofreading, and could not be re-used in the forced-choice task.

7.1.2.2. Genre-specific chunking hypothesis

The genre-specific chunking hypothesis can be confirmed by showing a difference between the responses of professional translators and other native speakers of Polish. It was shown in the corpus analysis that legal chunk aspectual preferences explained the choice of aspectual form in legal translation more accurately than general chunk aspectual preferences. This suggested that in legal language production, translators are more likely to select an aspectual version of a modal chunk that is typical of legal texts than the version that is typical of general language. Speakers that are not experienced with legal texts will selected the aspectual version typical of general language in all contexts because that is the only version they have entrenched in memory. In comprehension, we expect that the responses of the experienced and inexperienced speakers will differ. Naive native speakers and trainee translators have little or no experience with legal genre so aspectual preferences of modal chunks they have entrenched in memory will be typical of general language. This means that they will only be able to rely on general aspectual preferences when processing all types of linguistic input, general or genre-specific. This will result in the processing of modal chunks in legal context being hindered (e.g. slower reading or lower perceived well-formedness) if the legal aspectual preference of a chunk is different from the general aspectual preference available to inexperienced speakers, resulting in 'unusual' linguistic input that is more difficult to process. Based on what we know about frequency of occurrence and its effects (see Chapter 2 for more details), we do not expect translators to have such difficulties because they will have access to both the legal and general preferences of that chunk so the linguistic input will not be 'unusual'.

For the purposes of testing the genre-specific hypothesis, the above has different implications for the different tasks used here. The self-paced reading task and the judgement task test the participants' comprehension of modal chunks and how aspectual preferences facilitate that. In the self-paced reading task, naive native speakers and trainee translators will be expected to respond more quickly to the aspectual version of a modal chunk typical of general language because that is the version with which they are familiar. Professional translators are in turn expected to read both aspectual versions equally fast because they are familiar with them both. That is, professional translators are expected to exhibit less processing difficulties relative to the other participants. In the judgement task, we expect that the versions of modal chunks typical of general language will be rated by naive native speakers and trainee translators as more acceptable than the versions typical of legal language, while professional translators will rate both equally. In terms of language production (i.e. the forced-choice task), we expect naive native speakers and trainee translators to exhibit a tendency to select the version of a modal chunk typical of general language, while professional translators will select the form typical of legal language because of the legal nature of the stimuli to which they are exposed in the task. Therefore, in order to test the genre-specific hypothesis, we will have to expose participants with varying usage histories to modal chunks whose general aspectual preferences and legal aspectual preferences differ.

Ideally, all participants would be presented with the same stimuli. However, in order to show that the two groups (professional translators and other native speakers) respond to different aspectual preferences, stimuli with different combinations of aspectual preferences will be required. The infinitives in modal chunks seen by trainee translators and naive native speakers have general chunk aspectual

preferences different from all other types of aspectual preferences (legal chunk, general unigram, legal unigram). This way, if the general chunk-preferred version of a modal chunk is rated more favourably, read more quickly, and chosen more frequently, we will know with certainty that general chunk aspectual preferences - rather than any other aspectual preferences - affected the processing of the modal chunk, thus supporting genre-specific chunking hypothesis. Table 31 contains the possible combinations of aspectual preferences of stimuli selected for trainee translators and naive native speakers.

general chunk	legal chunk	general unigram	legal unigram
impf	pf	pf	pf
pf	impf	impf	impf

Table 31: Aspectual preferences of stimuli for TTs and NSs - genre-specific chunking

The infinitives in modal chunks seen by professional translators have legal chunk aspectual preferences different from all other types of aspectual preferences (general chunk, general unigram, legal unigram). This way, if the legal chunk-preferred version of a modal chunk is chosen more frequently than the other aspectual versions, we will know with certainty that legal chunk aspectual preferences - rather than any other aspectual preferences - affected the choice of aspectual versions of modal chunks, thus supporting the genre-specific chunking hypothesis. Table 32 contains the two possible combinations of aspectual preferences of stimuli selected for professional translators.

general chunk	legal chunk	general unigram	legal unigram
pf	impf	pf	pf
impf	pf	impf	impf

Table 32: Aspectual preferences of stimuli for PTs - genre-specific chunking hypothesis

Only two sentences from the corpus samples extracted for the analysis in Chapter 4 were found to match the criteria listed in Table 31 (stimuli for trainee translators and naive native speakers), and only six matched the criteria listed in Table 32 (stimuli for professional translators). These are used in the judgement task and in the self-paced reading task. Twelve additional utterances were extracted from the National Corpus of Polish - these were used in the forced-choice task.

7.1.2.3. Using decontextualized sentences

It should be noted that the sentences selected for the experimental tasks are decontextualized, i.e. the wider context in which they naturally occur is removed. It may be argued that this reduces the ecological validity of the experimental set-up: the wider context may be necessary for the processing and interpretation of a sentence so by removing it, we are preventing participants from processing that sentence in a way they would if the context was available. This may in turn lead to conclusions that are not necessarily generalizable to all real-world situations. It is argued here, however, that knowledge of the genre of wider context of use of the experimental stimuli is irrelevant to the processing of modal chunks investigated in the experimental study. There are several reasons for that.

First, to ensure that no other factors than aspectual preference affect the processing of aspectual versions of modal chunks, certain variables are carefully controlled. It was mentioned earlier that the type of situation can sometimes affect the choice of aspectual form (see sections 3.2.1 and 5.1 for more details). For that reason, only sentences that allow both aspectual versions were selected, substantially reducing the role of context in the processing of the different aspectual versions of modal chunks. Moreover, modality type, state of affairs applicability and polarity can also

sometimes exert influence over the speaker's choice of aspect (see sections 3.2.3 and 5.2). The role of modality type can be ruled out because all stimuli are deontic in nature. The other two variables are contained within the sentential context so any potential information that they carry will be available to the participants and can also be factored in during analysis of the results.

Second, as mentioned above, naive native speakers and trainee translators have little or no experience with the legal genre so the aspectual preferences of modal chunks they have entrenched in memory will be typical of general language. That is, even had they been clearly instructed that the context of use is legal, or if they were provided with the wider context that would enable them to deduct this themselves, it would not have changed the type of knowledge to which they have access; they would still only be able to process the modal chunks by relying on general aspectual preferences because due to their language experience they have no other types of chunks entrenched. This may result in the processing of modal chunks in legal context being hindered (e.g. slower reading or lower perceived well-formedness) because the aspectual preference of a chunk in a genre-specific text (such as the experimental stimuli) may be different to what it would be in general language, resulting in 'unusual' linguistic input. Their awareness of the genre or context of use of the experimental stimuli will not have any effect on these processing difficulties. That is, the availability of information about the genre or context of use will not change the results of the experimental tasks completed by naive native speakers and trainee translators.

Finally, professional translators are assumed to have substantial experience with the legal genre so they are expected to have two types of aspectual preferences

entrenched in their memory: those typical of general language and those typical of the legal genre. Based on what we know about frequency of occurrence and its effects (see Chapter 2 for more details), it is assumed that both types of preferences will be available to professional translators at all times rather than context-specific knowledge being activated depending on the context of use. This means that professional translators will encounter less processing difficulties: even if the aspectual form of a chunk in a given context is the opposite of the preferred form in that context, they also have that chunk's other aspectual preferences to rely on and facilitate processing. For example, if a given modal chunk has general aspectual preference for imperfective and legal aspectual preference for perfective, whatever form that chunk takes, it will not be 'unusual' to the translator and will therefore be processed more efficiently. Again, whether they are aware of the genre or content of use or not, this knowledge of different types of aspectual preferences is available to them. That is, awareness of the context of use or genre will have no effect on the experimental results.

To sum up, the extent to which the wider context of use can affect the processing of modal chunks is controlled by choosing only those sentences that allow both aspectual forms of the chunks contained in them. Other factors that have been previously shown to affect aspectual choice (polarity, modality and state of affairs applicability) are contained within the sentential context, which is provided as part of the experimental stimuli. The nature of frequency information allows us to rule out the role of wider context of use or genre in the processing of modal chunks included in this experiment and it is therefore not necessary to present the selected experimental stimuli together with their wider context for the two hypotheses to be tested reliably.

7.1.3. Experimental sets[15]

The judgement task and the self-paced reading task use the same stimuli - the experimental sets used in the two tasks were compiled in the same way so the set-up procedure for both tasks is described jointly in section 7.1.3.1. The forced-choice task was set up after the other two tasks were already completed and therefore contains different stimuli - the procedure is described separately in Section 7.1.3.2.

7.1.3.1. Judgement task and self-paced reading task

The set-up procedure involved three steps. First, separate experimental sets were created for the three participant groups. Sentences that test the general chunking hypothesis are the same in each set, although their numbers differ: naive native speakers saw more sentences than trainee translators and professional translators[16]. The number of sentences that test the genre-specific chunking hypothesis also differs in the three sets[17] but additionally, professional translators saw different sentences than trainee translators and naive native speakers because of the different ways the genre-specific chunking hypothesis has to be tested on participants with different usage histories (see Section 7.1.2.2). Table 33 contains the number of experimental sentences included in each set.

set	participant group	general chunking hypothesis	genre-specific chunking hypothesis	total
1	naive native speakers	16	2	18
2	trainee translators	11	1	12
3	professional translators	10	4	14

Table 33: Number of experimental stimuli in experimental sets

[15] All experimental sets can be found in Appendix 2
[16] Trainee and professional translators participated in a fourth task (not reported here), in which several of the experimental sentences were used. As a result, their judgement task set and the self-paced reading task set contain fewer experimental items than the sets used for naive native speakers.
[17] As above.

Apart from experimental sentences, each set also contains two filler sentences per experimental sentence. The purpose of the filler sentences is to distract the participants from the aims of the task. Each set begins with four practice sentences to allow participants to get used to the task, and ends with four practice sentences. In total, the sets include from 52 to 70 sentences (see details in Table 34).

set	participant group	experimental sentences	filler sentences	practice sentences	total
1	naive native speakers	18	36	16	70
2	trainee translators	12	24	16	52
3	professional translators	14	28	16	58

Table 34: Number of all stimuli in in experimental sets

Second, the sentences in each of the three sets were divided into two, creating two experimental subsets per participant group (see Table 35). The two subsets contain different experimental sentences that do not overlap - if subset 1 contains sentences A, B, C, then subset 2 contains sentences D, E, F.

set	participant group	experimental sentences	subset	
			subset no.	stimuli
1	naive native speakers	18	1	9
			2	9
2	trainee translators	12	1	6
			2	6
3	professional translators	14	1	7
			2	7

Table 35: Experimental subsets per participant group

Half of the sentences in each subset contain modal chunks in the dispreferred aspectual form and the other half in the preferred aspectual form. This was done to ensure that each participant rates not only modal chunks that we expected to be rated highly because they occurred in the preferred form, but also modal chunks that we

expected to be rated less favourably because they occurred in the dispreferred aspectual form.

Finally, an aspectual mirror subset was created for each of the experimental subsets. In each mirror set, the aspectual forms of modal chunks were reversed. That is, if sentence A of subset 1 contains a modal chunk with a chunk-preferred aspectual form of the infinitive, that same sentence in the aspectual mirror subset will contain the infinitive in the unigram-preferred aspectual form. This enables us to directly compare ratings given by different participants to two aspectual versions of the same chunk.

This set-up procedure is visualised in Figure 2 (based on set 1) - it resulted in four experimental subsets per participant group. In each subset, the sentences were ordered randomly.

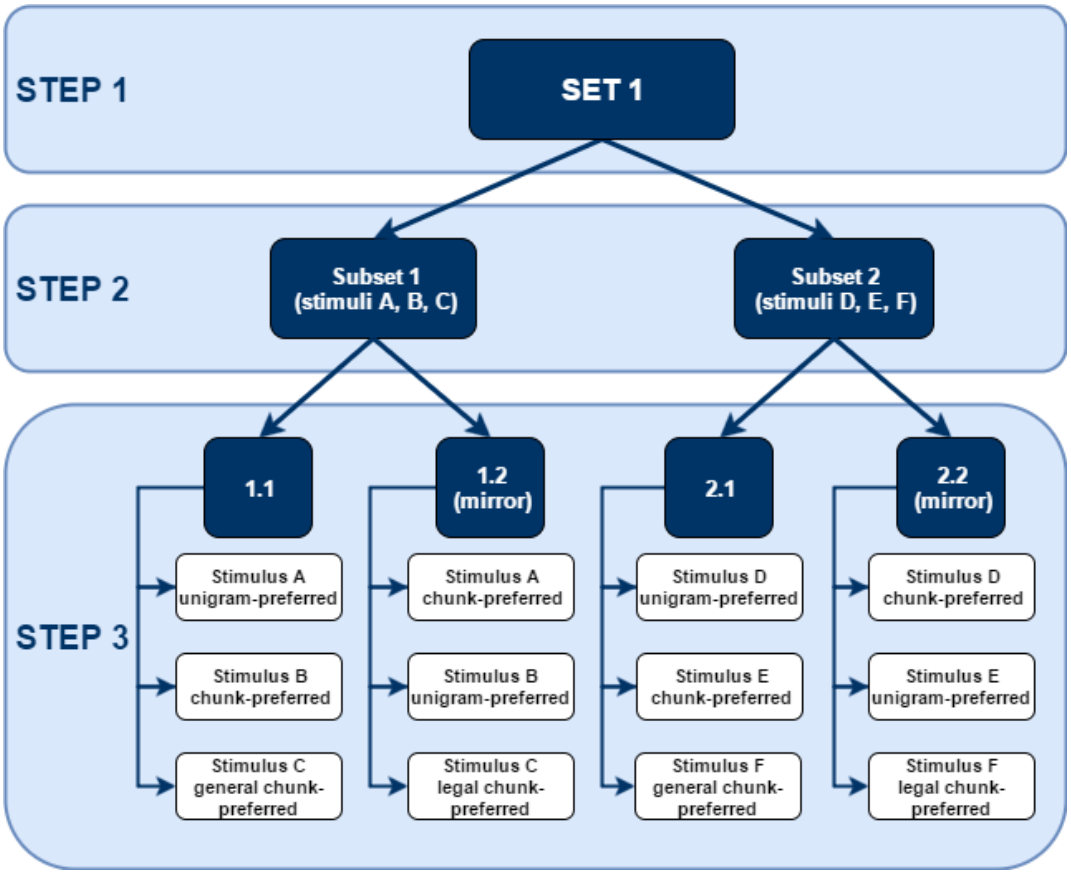

Figure 2: Judgement task and self-paced reading task - set-up procedure

The subsets were used in the judgement task and in the self-paced reading task but the same participant was never assigned the same subset in both tasks. If a participant was assigned subset 1.1 or 1.2 for the judgement task, then that participant could only be assigned subset 2.1 or 2.2 for the self-paced reading task. The subsets were assigned based on participant's date of birth, as shown in Table 36.

date of birth (month)	judgement task subset	date of birth (year)	SPR task subset
Jan/Feb/Mar	1.1	even (1984, 1986...)	2.1
		odd (1985, 1987...)	2.2
Apr/May/Jun	1.2	even (1984, 1986...)	2.1
		odd (1985, 1987...)	2.2
Jul/Aug/Sept	2.1	even (1984, 1986...)	1.1
		odd (1985, 1987...)	1.2
Oct/Nov/Dec	2.2	even (1984, 1986...)	1.1
		odd (1985, 1987...)	1.2

Table 36: Assignment of judgement task and SPR task sets

7.1.3.2. Forced-choice task

The set-up procedure involved two steps. First, the 24 sentences extracted from the National Corpus of Polish for the purposes of the forced-choice task (see Section 7.1.2) were split in half to form two sets, with 12 experimental sentences in each set. Each set also contains 8 practice sentences and 24 filler sentences (see Table 37).

sentence type	number of sentences
experimental	12
filler	24
practice	8
total	44

Table 37: Stimuli in forced-choice task sets

Second, four subsets were created for each of the two sets. The subsets only differ in the ordering of sentences to ensure that the order of presentation of stimuli did not affect participants' choices. Each participant was only assigned one set, which is done at random based on the participant's date of birth - set 1 was assigned to participants born in odd years (1983, 1985, etc.) while set 2 was assigned to those born in even years (1984, 1986, etc.). Subsequent subsets were assigned based on the participant's month of birth, as shown in Table 38.

date of birth (year)	set	date of birth (month)	set version
odd (1985, 1987...)	1	Jan/Feb/Mar	1.1
		Apr/May/Jun	1.2
		Jul/Aug/Sept	1.3
		Oct/Nov/Dec	1.4
even (1984, 1986...)	2	Jan/Feb/Mar	2.1
		Apr/May/Jun	2.2
		Jul/Aug/Sept	2.3
		Oct/Nov/Dec	2.4

Table 38: Assignment of forced-choice task subsets

7.1.4. Set-up

The set-up of each task involved certain methodological decisions, which are described in this section, together with the administration procedure of each task. Before commencing the tasks, each participant was asked to read and sign a consent form (available in Appendix 1), which provided a summary of the study and explanation of all tasks. The consent form informed participants of the right to withdraw at any moment without giving a reason, and contained data protection information. Participants were asked to first complete the judgement task, followed by the self-paced reading task. The forced-choice task was last.

7.1.4.1. Judgement task

Ratings of well-formedness can be collected in several ways, including Likert scales, magnitude estimation, and thermometer task. The thermometer task (Featherston, 2008) combines the advantages and shortcomings of Likert scales and magnitude estimation, and will therefore be used here, albeit in a slightly modified form. In a Likert scale task, participants are asked to rate an utterance on a numerical scale (1 to 5, 1 to 7, or similar) with the endpoints defined as acceptable or unacceptable. In a magnitude estimation task, participants are first asked to rate a standard sentence with a specific numerical value of their choice, e.g. 100. Any additional sentences are then rated with reference to the standard sentence. For example, if a sentence is considered twice as acceptable as the standard sentence, it is rated 200 (ibid., pp. 33-34). The Likert scale is simpler to understand and more natural than magnitude estimation because it is easier to decide whether a sentence is closer to the 'acceptable' or 'unacceptable' end of the scale than deciding how proportionately good or bad it is compared to the standard sentence (Dąbrowska, 2010, p. 8). Recent

studies have even shown that participants struggle to make ratio comparisons of two sentences, meaning that the primary assumption of magnitude estimation is not met (Schütze & Sprouse, 2013, p. 35). There are, however, certain advantages of magnitude estimation tasks - they are more likely to pick up on subtle differences between sentences than a Likert scale with a small number of fixed values. In a thermometer task, participants are provided with two reference sentences and their associated ratings, for example 20 and 40. They are then asked to rate other sentences with reference to those two reference sentences. This way, the sensitivity of magnitude estimation and the intuitive nature of Likert scales are combined (ibid.).

The thermometer task was modified for the purposes of the current investigation. In addition to providing participants with two reference sentences, we argue that the number of possible values should also be fixed (e.g. from 1 to 100): this gives participants clear endpoints to the scale of acceptability and unacceptability, creating a more natural and intuitive task. By asking participants to rate utterances on a linear scale from 1 to 100, we do not have to rely on participants' ability to make ratio comparisons, but at the same time we allow for even very subtle differences in well-formedness of sentences to be captured.

The task was set up online using Qualtrics software of the Qualtrics Research Suite (Qualtrics, Provo, UT)[18]. A slider scale with grid lines and numerical values at the increments of 10 was used (see Figure 3). Participants were informed that '0' means "the phrase sounds particularly bad" and that 100 means "the phrase sounds very good". The two reference sentences are set at 10 and 90 rather than 0 and 100. This way, participants can give a rating of 0-9 to sentences that are perceived as worse

[18] Version 59627 of the Qualtrics Research Suite. Copyright© 2014 Qualtrics.

than the 'bad' reference sentence, or a rating of 91-100 to sentences that are perceived as better than the 'good' reference sentence.

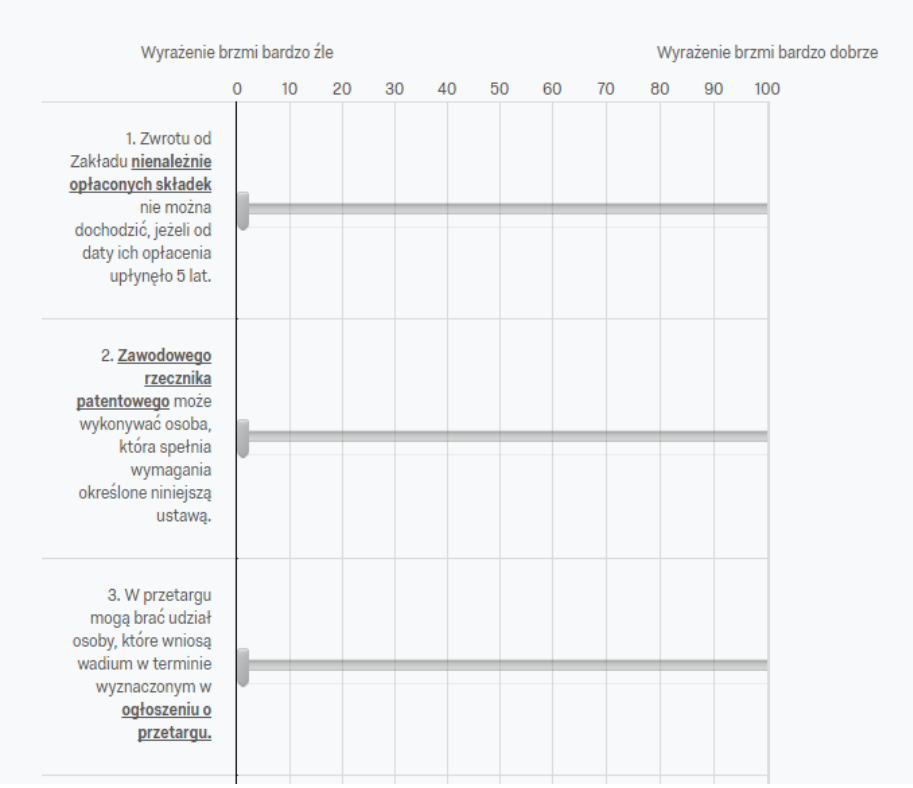

Figure 3: A screenshot of the scale slider used in the judgement task

Each participant was given a link to their assigned experimental subset. The task begun with a summary of the study, an explanation of the task, and information about the participants' right to withdraw. This was followed by questions about the participant's background: gender, age, place of birth and residence, education, profession, and reading habits. This information was collected for statistical purposes. After filling in the personal information, participants were given the following instructions in Polish:

PL: Poniżej znajdą Państwo (XX) zdań. Proszę ocenić podkreślone wyrażenia (biorąc pod uwagę kontekst całego zdania) w skali od 0 do 100, gdzie 0

oznacza: “wyrażenie brzmi bardzo źle po polsku” zaś 100 oznacza: “wyrażenie brzmi bardzo dobrze po polsku”. Proszę kierować się intuicją i swoim wyczuciem językowym. Nie ma poprawnych ani niepoprawnych odpowiedzi.

Przykład zdania ocenionego na 10: Ostateczna decyzja musi *być wydawana* jutro.

Przykład zdania ocenionego na 90: Opakowanie *musi umożliwiać* identyfikację towaru.

EN: Below you will find (XX) sentences. Please rate the underlined phrase on a scale from 0 to 100, taking into account the context of the sentence. A rating of 0 means "the phrase sounds particularly bad in Polish" while a rating of 100 means "the phrase sounds very good in Polish". Please rely on your intuition when rating the sentences. There are no incorrect answers.

An example of a sentence rated 10: Ostateczna decyzja musi *być wydawana*[19] jutro.

An example of a sentence rated 90: Opakowanie *musi umożliwiać*[20] identyfikację towaru.

[19] Translation: *The final decision should be issued tomorrow.* Here, the aspectual form is imperfective but it should be perfective because the adverbial of time - *tomorrow* - binds this situation temporally. Also, the state of affairs is specific - it applies to one instance of issuing the final decision. As such, a perfective form is required. Therefore, this sentence is rated as 10.

[20] Translation: *The packaging must enable identification of the product*. Here, the aspectual form is imperfective and it is the correct form, given the context. The state of affairs applies generally to all packaging on all products at all times. Therefore, this sentence is rated as 90.

Participants were asked to rate the well-formedness of the underlined phrase rather than the entire sentences to ensure that they rate the modal chunk that is of interest for the purposes of the current study. In experimental sentences, the modal chunk was underlined. In filler and training sentences, other elements were underlined to ensure that participants did not realise that modal chunks and aspect of the verb were the items under investigation. Half of the filler and training phrases were manipulated to make them ungrammatical - this was done to ensure that participants do not develop a strategy of rating all items highly because the majority of the sentences that they saw were formed correctly.

The ratings begun with four practice sentences, followed by experimental and filler sentences (with each experimental sentence separated from the next experimental sentence by at least one filler sentence), and ended with the remaining practice sentences. The sentences were distributed over 6 or 7 screens (depending on the experimental set), with no more than 6 sentences on each screen. Each new screen begun with the same examples of a sentence rated as 10 and a sentence rated as 90, and with slider guidelines. The questionnaire ended with a thank you message and the researcher's contact details.

7.1.4.2. Self-paced reading task

Stimuli in a self-paced reading task can be presented to participants in three ways: cumulative, linear non-cumulative, and centre non-cumulative (Marinis, 2010, p. 150). The cumulative presentation resembles the natural process of reading but enables the reader to form expectations and predictions about the next word(s): the utterance is displayed on the computer screen as a series of dashes, with each dash corresponding to each letter in the utterance. When the button is pressed, the first

word is displayed and remains on the screen when the button is pressed again and the next word is displayed. The linear non-cumulative presentation is similar, but each displayed word is replaced with dashes when participant presses the button to display the next word. The centre non-cumulative presentation prevents any expectations from being formed but does not resemble the reading process as closely as the cumulative or linear methods - words are displayed in the centre of the screen one by one, without any indication as to the length of the words or sentences.

The centre non-cumulative presentation is used here. We are interested in the influence of frequency on the processing of modal chunks; if the elements preceding the modal chunks are displayed, the participants are likely to re-read the other elements, increasing the reading times of the modal chunks and preventing us from drawing conclusions about the role of frequency. The same applies to the display of dashes: participants might form expectations based on the length of the sentence, influencing the reading time of the modal chunks. The centre non-cumulative presentation prevents such unwanted influence.

The task was set up in PsychoPy v1.78.00 (Peirce, 2007). A Cedrus response pad was connected to a Windows 8 laptop. A response pad was used instead of the laptop's keyboard as it allows for more accurate recording of reading times (Kaiser, 2013, p. 141). Sentences were divided into chunks and displayed on the screen one chunk at a time:

PL: Każdy wierzyciel :: może wnosić :: swoje roszczenie :: w języku urzędowym :: Państwa Członkowskiego.

EN: Any creditor :: may lodge :: his claim :: in the official language :: of a Member State.

The task was administered at various locations, in quiet rooms with minimal distractions. All participants completed the task on the same device. Each participant was presented with their assigned experimental subset. The task begun with brief instructions from the researcher. Participants were reminded of their right to withdraw at any time. The following instructions were then repeated on the computer screen, giving the participants an opportunity to familiarise themselves with the response pad before commencing the task:

PL: Proszę o przeczytanie 35 zdań podzielonych na fragmenty. Po pojawieniu się pierwszego fragmentu zdania, proszę nacisnąć środkowy guzik, aby przejść do kolejnego fragmentu. Po przeczytaniu wszystkich fragmentów każdego zdania pojawi się symbol '+' - w tym momencie można zrobić przerwę w czytaniu. Proszę czytać zdania ze zrozumieniem, ponieważ po przeczytaniu niektórych zdań zostaną zadane pytania dotyczące ich treści. Na pytania proszę odpowiedzieć 'Tak' (zielony guzik) lub 'Nie' (czerwony guzik). Proszę czytać zdania we własnym tempie i bez pośpiechu. Aby przejść do zadania, proszę nacisnąć środkowy guzik.

EN: Please read the following 35 sentences, which have been divided into chunks. After reading the first chunk, press the middle button and the next chunk will appear. After reading the whole sentence, the symbol '+' will appear - this is when you can take a break from reading. Read the sentences carefully as you will be asked simple yes (green button) / no (red button) questions about some of them. Read the sentences at your own pace and without rushing. Press the middle button to begin the task.

After reading the instructions, which were divided up over two screens, participants proceed to the task. Each participant saw the sentences in a different order, starting with the four practice sentences (randomly ordered), followed by randomly ordered experimental and filler sentences (with each experimental sentence separated by at least one filler sentence), and ending with four randomly ordered practice sentences.

Comprehension questions were included in this task to ensure that participants are focused on reading the sentences and understanding the content, rather than pressing the button mechanically (Marinis, 2010, p. 153). These were simple yes/no questions that appeared immediately after selected sentences and referred to the sentence presented immediately before the question. There were 8-12 questions in each set, depending on the number of experimental and filler sentences. If a participant failed to provide a correct answer to more than two questions, their reading times were removed from the analysis as it indicates they did not focus on reading the sentences.

7.1.4.3. Forced-choice task

Forced-choice tasks take various forms - participants can be asked to fill in gaps in sentences (e.g. Bermel, Knittl, & Russell, 2015) or to choose the more acceptable sentence from a list of two (Schütze & Sprouse, 2013, p. 32). The aim is to establish whether there is a qualitative difference between two or more linguistic variants. The assumption is that two items that do not differ will result in a random 50/50 split in the two forms being chosen (ibid., pp. 31-32). Here, participants were presented with sentences from which elements were removed and they were asked to select their preferred element from a list of two to fill the gap. In experimental sentences, the modal chunks were removed and three dots in square brackets were inserted in their place, for example:

(1) PL: Notatki, o których mowaw ust. 6, **[...]** omówienie wyniku ponownego rozpatrzenia sprawy, wyniku skargi lub rewizji.

(removed part: **powinny zawierać**)

Below the sentence, a list containing two aspectual versions of the modal chunk removed was provided: the chunk-preferred and unigram-preferred aspectual version for the general chunking hypothesis, and the general chunk-preferred and legal chunk-preferred for the genre-specific chunking hypothesis. For each filler and training sentence, other non-specific elements were removed, for example:

(2) PL: Notatki, **[...]** w ust. 6, powinny zawierać omówienie wyniku ponownego rozpatrzenia sprawy, wyniku skargi lub rewizji.

(removed part: **o których mowa**)

Below the filler and training sentences, lists containing two phrases were provided: one of them was the original element removed from the sentence, and the other one was a synonymous phrase. Participants were instructed to choose the one they found more appropriate. For sentence (2) above, the following two options were given:

(a) o których mowa (EN: which are discussed)

(b) wspomniane (EN: mentioned)

The order in which the two choices were presented to participants was randomised for each participant, for example, one participant saw the chunk-preferred version of a modal chunk first and the unigram-preferred version second, while for another participant the order is reversed. This was done to ensure that the order in which the two options are presented does not affect the responses.

The task was set up online using Qualtrics software of the Qualtrics Research Suite (Qualtrics, Provo, UT)[21]. Each participant was given a link to their assigned experimental subset. The task begun with the following instructions in Polish:

PL: Dziękuję za udział w ostatniej części badania. Zadanie polega na uzupełnieniu brakujących fragmentów w 44 zdaniach. W miejscu brakującego fragmentu znajduje się ten symbol: [...]. Pod każdym ze zdań znajdą Państwo dwa fragmenty do wyboru - proszę kliknąć na ten, który według Państwa najlepiej pasuje w danym zdaniu (ignorując literówki). Jak w poprzednich zadaniach, proszę kierować się swoją intuicją i wyczuciem językowym - wszystkie odpowiedzi są poprawne.

EN: Thank you for participating in the final task. You will be asked to complete 44 sentences. The missing part in each sentence will be marked with this symbol: [...]. Under each sentence, you will be given two fragments to choose from: please click on the one you think fits the given sentence best (please ignore any typos). As in previous tasks, please trust your linguistic intuition; all answers are correct.

The sentences were distributed over 6 screens, with 7-9 sentences on each screen. The questionnaire ended with a thank-you message and the researcher's contact details.

7.1.5. Statistical analysis

Two types of regression are applied here to analyse the data obtained from the three experimental tasks: binary mixed effects logistic regression (judgement task and

[21] Version 59627 of the Qualtrics Research Suite. Copyright© 2014 Qualtrics.

forced-choice task) and linear mixed effects regression (self-paced reading task). In the experimental analyses, we deal with repeated measures designs, in which multiple participants respond to multiple items (Baayen, 2008, p. 242). That is, each item (here: modal chunk) selected for the tasks is seen by several participants, and each participant is exposed to several items. Individual preferences of each participant, and the individual characteristics of each sentence, may influence participants' responses. For example, some participants may be slower readers while others might be faster, affecting the reading times recorded in the self-paced reading task, which might mistakenly be interpreted as the effect of a given stimulus. Similarly, some participants might generally be more generous in their ratings of stimuli than others, potentially resulting in misleading conclusions. At the same time, certain sentences or modal chunks may be more difficult to process, making a given item longer or more demanding to read, resulting in a longer reading time or lower rating of well-formedness: such responses might have nothing to do with aspectual preferences, but could be mistakenly interpreted as being influenced by aspectual preferences. It is therefore very important to account for individual differences when a regression model is fitted to experimental data. This allows the experimenter to ensure that the results obtained are truly influenced by the predictor variables, and not by the individual preferences of participants, or characteristics of individual stimuli.

We can neutralise the unwanted influence of the abovementioned factors by adding *random effects* to the fitted models (Winter, 2013). Adding random effects for subject and item tells the model to adjust the intercepts for individual subjects and individual items. When we add random effects to the fitted model, on top of the

fixed variables that we investigate (here: aspectual preference), we obtain a *mixed effects* model.

Mixed models are fitted to all of the experimental data obtained here. Mixed effects binomial logistic regression and mixed effects linear regression are fitted in R using {lme4} package (Bates et al., 2015). Packages {lattice} (Sarkar, 2008), {car} (Fox i Weisberg, 2011), and {effects} (Fox, 2003) were used for plotting. The two random effects added to each model will be *participant* and *chunk*. We first fit a model with *participant* as a random effect, and then a separate model with *participant* and *chunk*. The two models are compared using the anova() function to see which random effects improve the performance of a given model. The better performing model will be used to interpret the results of the experiments.

7.1.5.1. Judgement task - data preparation

A thermometer scale with a fixed number of possible values (1-100) was used to enable participants to account for subtle differences in well-formedness of utterances, while keeping the task natural and intuitive. The data was to be analysed by fitting mixed effects linear regression. The technique assumes that data are normally distributed, but Figure 4 shows that the distribution of ratings is strongly bimodal so the basic assumption of linear regression is not met. The participants' responses do not cluster around a single mode, like they would in a normal distribution. Instead, one group of responses clusters around one mode (31.16041), and another group of responses clusters around another mode (88.62227). That is, the responses are at either one endpoint of the 100-point scale, or at the other endpoint, suggesting that the scale of well-formedness is seen as a matter of good or

bad, rather than a gradual progression. This finding could be seen as support for using a simpler scale with fewer values.

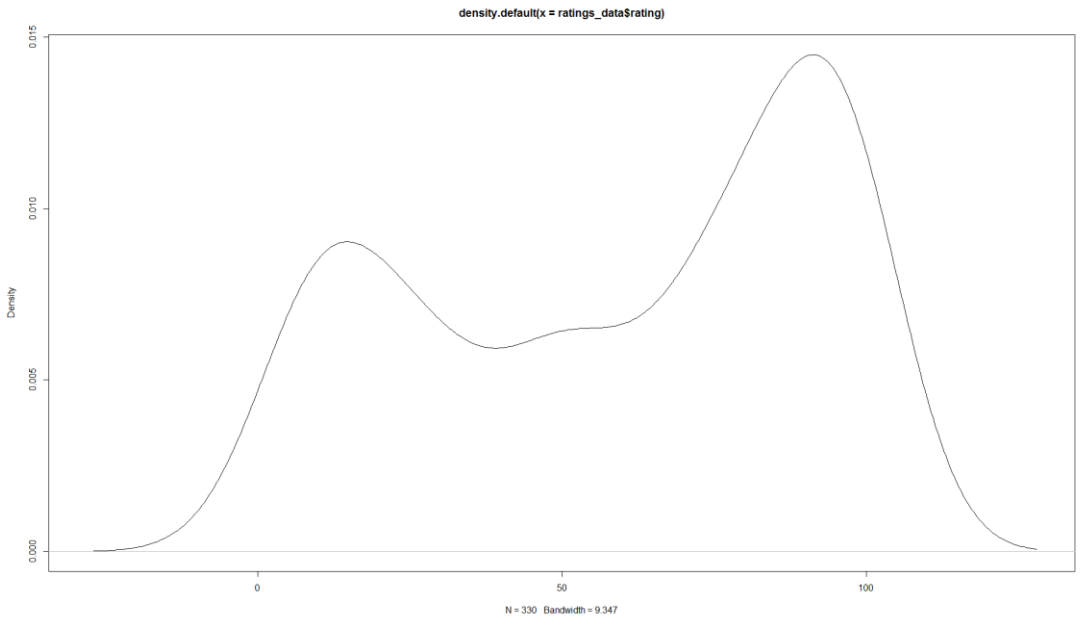

Figure 4: Distribution of judgement data

Whether a distribution of data is significantly bimodal can be established by calculating the dip statistic (Hartigan & Hartigan, 1985). The higher the dip statistic and the lower the p-value, the more bimodal the distribution. The dip statistic for the acceptability data was calculated with the function dip.test() in R - it returned a D value of 0.050124 and a p-value of 6.647e-06, confirming that the distribution of data is significantly bimodal. One of the assumptions of linear regression is that data is normally distributed; since our data is not, linear regression cannot be performed. One way to deal with the strongly bimodal data is to binarise the ratings and to fit binomial regression instead of linear regression. Ratings below a given threshold (the choice of which is discussed below) are therefore coded as 0, and those above that threshold - as 1.

In order to find a threshold along which to split the data, a two-component mixture model is fitted in R using the {mixtools} package (Benaglia et al., 2009) - the procedure estimates the parameters of the two modes in a bimodal distribution. Then, based on the estimated parameters of a mixture model, an algorithm[22] calculates the weight of the smaller mode and the weight of the pit between the two modes, and then finds a threshold across which the two modes can be split. The procedure provides two cut-off points (see Figure 5). The cut-off point of 67.57894 (red line) ensures that the probability of making an error in favour of one group is exactly the same as making an error in favour of the other group. The cut-off point of 64.04857 (blue line) increases the chance of wrong classification - by choosing this value to split the data we are three times as likely to make an error in favour of the lower group, when it should be in the higher group, compared with the higher group, when it should be the lower group.

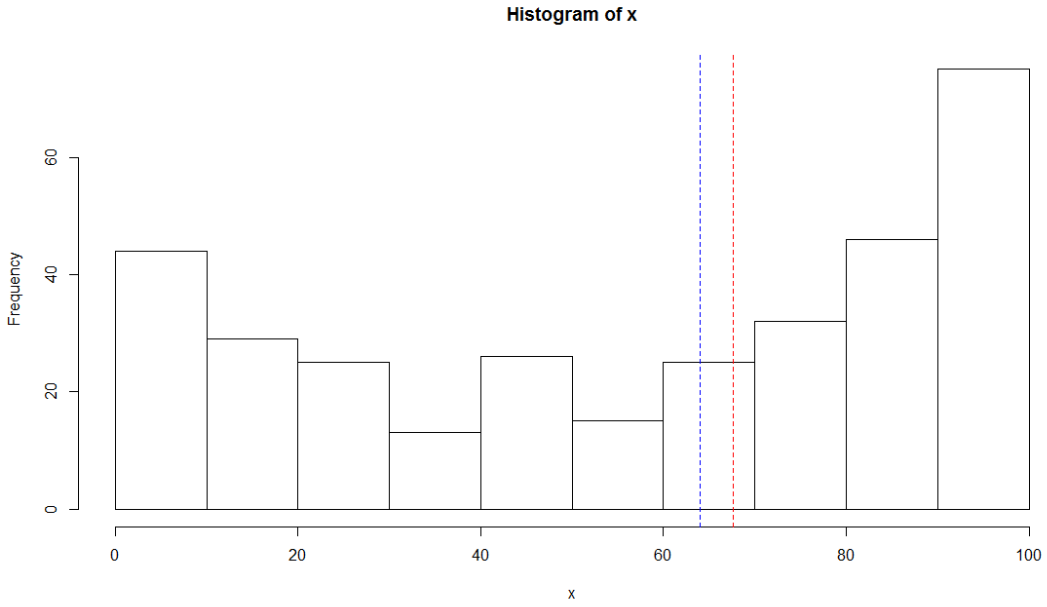

Figure 5: Cut-off point for splitting the binarised judgement data

[22] As described on http://stats.stackexchange.com/questions/57993/how-to-explain-how-i-divided-a-bimodal-distribution-based-on-kernel-density-esti/78397#78397 (accessed on 15/04/2016).

We therefore use the cut-off point of 67.57894 to split the bimodal data to ensure that the risk of a wrong classification is equal for both groups. Logistic regression estimates the probability of the event occurring so the binary response variable is coded as either 1 (the event occurring) or 0 (the even not occurring). Therefore, all ratings above or equal to 67.57894 are coded as 1, indicating that a sentence was given a high rating (a rating above 67.57894). All ratings below 67.57894 are coded as 0, indicating that a rating above 67.57894 was not given. Mixed effects logistic regression is then fitted to the binarised data in R using the function glmer() from package {lme4} (Bates et al., 2015).

7.1.5.2. Self-paced reading - data preparation

One of the assumptions of linear regression is a normal distribution of data. Hartigan's Dip Test for unimodality confirms that the distribution of the reading time data is unimodal (D=0.016436, p-value=0.833). However, plotting the data (Figure 6) indicates that we are dealing with a highly skewed unimodal distribution. Outliers are removed but Shapiro-Wilk test for normality confirms that data is not normally distributed (W=0.95093, p=5,347e-09). The data is therefore logarithmically transformed in order to remove some of the skewedness (Baayen, 2008, p. 71).

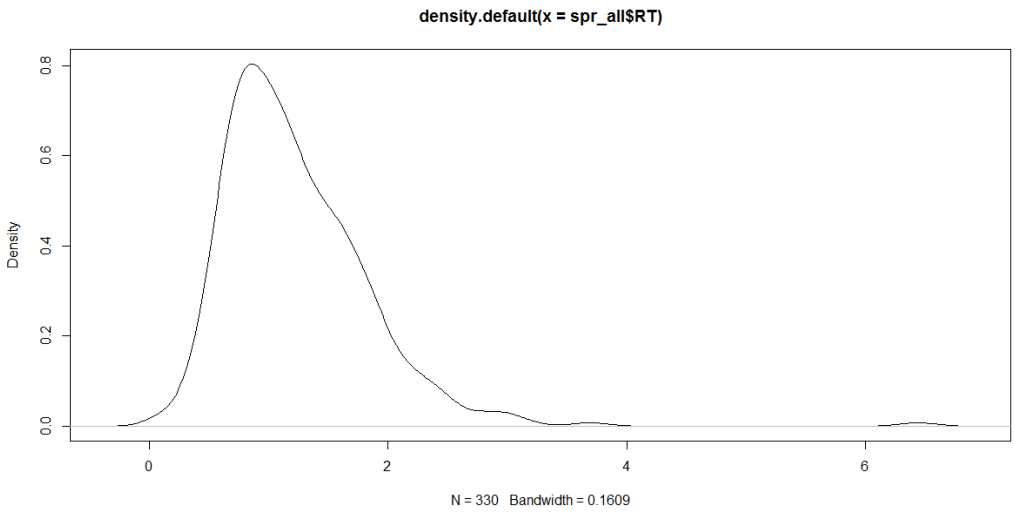

Figure 6: Distribution of the reading time data

The distribution of the log-transformed data (with outliers removed) resembles the bell-curve shape of a normal distribution more closely (Figure 7). The Shapiro-Wilk test confirms that the data now follows a normal distribution (W=0.99367, p=0.1918), and can as such be used to fit a linear regression model.

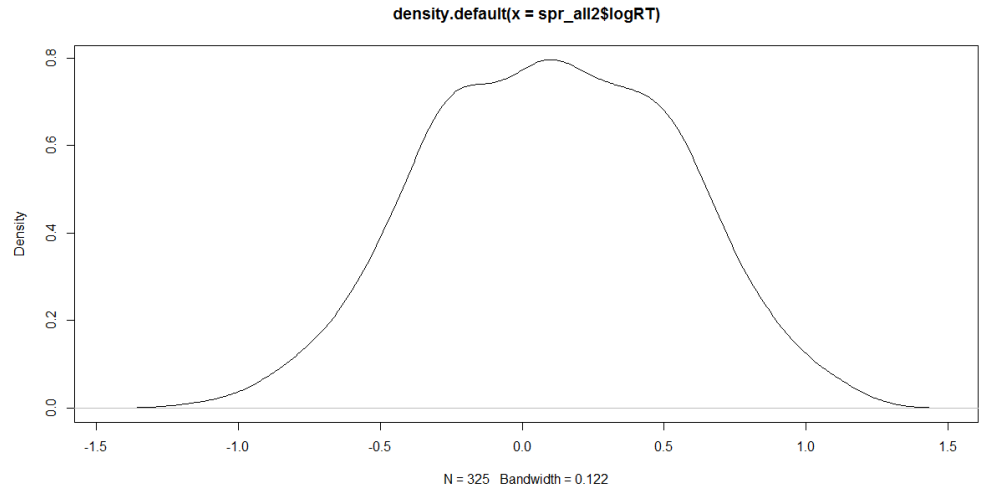

Figure 7: Distribution of log-transformed reading time data

The same procedure is performed on the reading times of the post-modal chunks, which will be checked for a spillover effect (see Section 7.2.1.2 for more details).

Hartigan's Dip Test for unimodality confirms that the distribution of the post-modal chunk reading time data is unimodal (D=0.017893, p=0.7082). However, Figure 8 suggests that data is skewed; outliers are removed but data is still not normally distributed (W=0.92127, p=4.819e-12) and is therefore log-transformed.

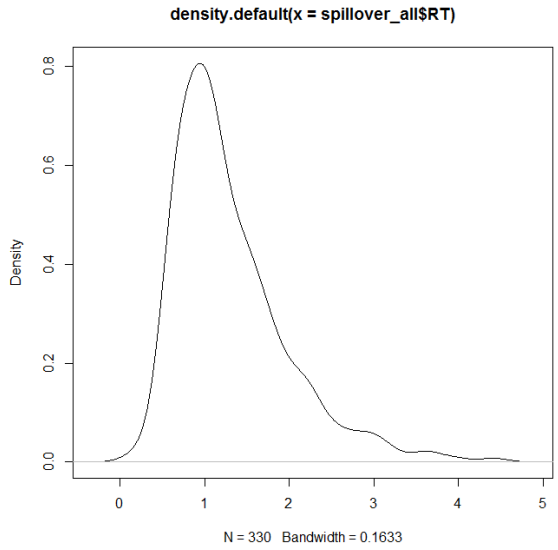

Figure 8: Distribution of the post-modal chunk reading time data

After removing some of the outliers, the data is now close to a normal distribution (W=0.98914, p=0.01661).

The function lmer() from package {lme4} (Bates et al., 2015) is used to fit the mixed effects models to the log-transformed data.

7.2. Experimental results

The corpus analysis of aspectual preferences supported both chunking hypotheses - the aim of the experimental analysis was to validate those results. This was done with a series of three experimental tasks, completed by native speakers of Polish.

The three tasks measured different aspects of the speakers' linguistic knowledge. The aim of the judgement task was to establish whether aspectual preferences of infinitives in modal chunks affect the perceived well-formedness of those modal chunks. The aim of the self-paced reading task was to establish whether aspectual preferences of infinitives in modal chunks affect the speed of processing of those chunks. Finally, the aim of the forced-choice task was to establish whether aspectual preferences of infinitives in modal chunks affect the choice of the aspectual versions of those chunks. The collected data was analysed using mixed effects logistic and linear regression. The results - discussed below - support the general chunking hypothesis but not the genre-specific chunking hypothesis.

7.2.1. General chunking hypothesis

The general chunking hypothesis stated that 'modal verbs and infinitives that follow them are entrenched in memory as chunks and the aspectual preferences of those modal chunks facilitate the selection and processing of aspect.' The results provide clear support for the hypothesis. Judgement task is discussed in Section 7.2.1.1, self-paced reading task in Section 7.2.1.2, and forced-choice task in Section 7.2.1.3.

The R code used to fit the regression model is available in Appendix 3; all data sets used in the analyses are available to download online[23].

7.2.1.1. Judgement task

For the purposes of the judgement task, the general chunking hypothesis can be operationalized as follows:

[23] https://drive.google.com/open?id=0Bz9KVHFRWI3NclB4aWRRY2poQ28

Ratings given to chunk-preferred versions of modal chunks will be significantly higher than ratings given to unigram-preferred aspectual versions of modal chunks.

Rating is the response variable being predicted. One fixed predictor variable called *version* was added, which specifies whether a modal chunk was presented to a participant in the infinitive's chunk-preferred aspectual form (the *preferred* form) or in the infinitive's unigram-preferred aspectual form (the *dispreferred*). That is, the purpose of the statistical analysis was to see if there are any correlations between the ratings given by participants and the aspectual form of the modal chunk which a given participant saw. Two random effects were also added to the model - one for participant (to rule out the influence of the characteristics of individual participants on the results) and one for modal chunk (to rule out the influence of the characteristics of individual modal chunks on the results).

The mixed effects model was fitted using the glmer() function from the package {lme4} (Bates et al., 2015). Two models were fitted and compared using the function anova():

(1) rating ~ version + (1|participant)

(2) rating ~ version + (1|participant) + (1|chunk)

Anova shows that model 2 performs significantly better than model 1 (p=0.01217). The AIC of model 2 (361.2) is lower than that of model 1 (365.5) by over 4 points, which suggests that model 2 is considerably better than model 1. We are therefore reporting the results of model 2 with two random effects.

The general trend is for the chunk-preferred aspectual version of the modal chunk to correlate with increased odds of a rating above the cut-off point of 67.57894

(estimate=1.2323, standard error=0.3788, p-value=0.00114). That is, when participants see a modal chunk with the chunk-preferred aspectual version of the infinitive, the odds of a higher rating being given to such a modal chunk increase significantly by 1.2323.

When an interaction between aspectual preferences and participant groups is added to the model, differences in how the three groups rate can be observed. Figure 9 visualises the interaction - the red dotted line is for ratings of modal chunks with the chunk-preferred aspectual version of the infinitive, while the black line is for chunks with the unigram-preferred aspectual version. It should be noted that the connecting lines do not suggest that a linear relationship exists between the three variables.

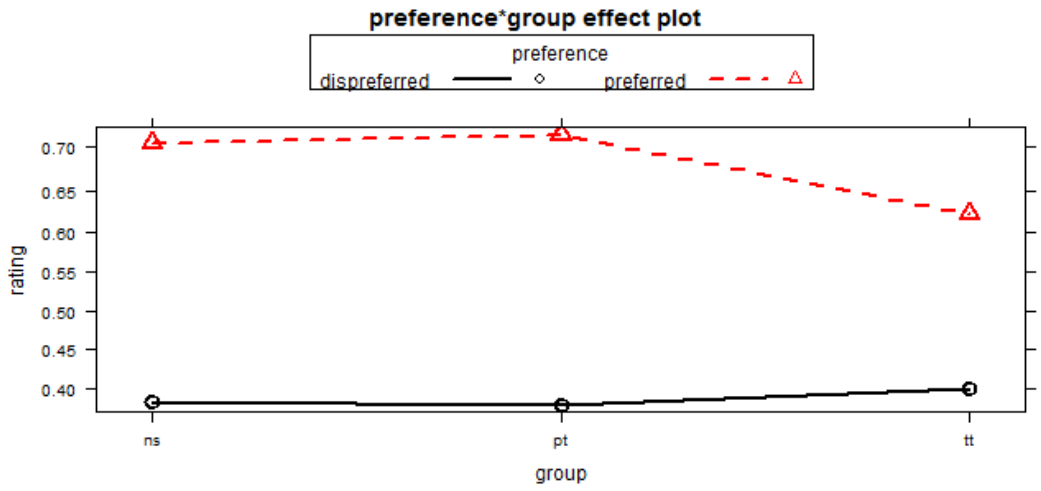

Figure 9: Differences in ratings between the three groups of participants

The plot suggests that the difference between the ratings of chunk-preferred and unigram-preferred versions of modal chunks given by trainee translators is not as pronounced as between the ratings given by naive native speakers and professional translators. Separate models fitted to the ratings given by individual groups support this - there are significant correlations between chunk-preferred aspectual versions of

modal chunks and high ratings given by naive native speakers (estimate=1.3875, standard error=0.6648, p-value=0.0369) and professional translators (estimate=1.2770, standard error=0.5148, p-value=0.0131), while the correlation for trainee translators is not significant (estimate=0.9424, std. error= 0.5465, p=0.0847). In order to avoid a type 1 error, however, we consider the result for trainee translators as insignificant. That is, we conclude that the ratings given by trainee translators are not affected by aspectual preferences of modal chunks. Table 39 contains the model parameters for the three groups of participants; significant results are highlighted in grey.

group	aspectual form	estimate	standard error	p-value
NS	preferred	1.3875	0.6648	0.0369
	dispreferred	-0.4957	0.4612	0.2824
TT	preferred	0.9424	0.5465	0.0847
	dispreferred	-0.4822	0.4090	0.2384
PT	preferred	1.2770	0.5148	0.0131
	dispreferred	-0.5353	0.3639	0.1413

Table 39: Parameters of the models fitted to ratings

The results obtained for naive native speakers and professional translators support the general chunking hypothesis: the ratings given to chunk-preferred and unigram-preferred versions of modal chunks differ significantly, with the former more likely to be rated favourably. There is no significant interaction with the variable *age*. That is, the age of participants does not seem to affect the ratings.

7.2.1.2. Self-paced reading task

For the purposes of the self-paced reading task, the general chunking hypothesis can be operationalized as follows:

The reading times of chunk-preferred aspectual versions of modal chunks will be significantly shorter than the reading times of unigram-preferred aspectual versions of modal chunks.

LogRT (log-transformed reading times - see Section 7.1.5.2 for details) is the response variable being predicted. One fixed predictor variable called *preference* is added, which specifies whether a modal chunk was presented to a participant in the infinitive's chunk-preferred aspectual form or in the infinitive's unigram-preferred aspectual form. That is, the purpose of the statistical analysis was to see if there are any correlations between the reading times and the aspectual version of the modal chunk which a given participant saw. Moreover, three additional variables were added to account for the length of the modal chunk (some chunks are longer than others and will naturally require more time to read), the position of the modal chunk in the sentence (some modal chunks may be located towards the beginning of the sentence while others towards the end, which may affect the processing of that chunk), and the position of the sentence in the experiment (each participant was presented with a different random order of sentences, which may also affect processing). Finally, two random effects were added - one for participant (to rule out the influence of the characteristics of individual participants on the results) and one for modal chunk (to rule out the influence of the characteristics of individual modal chunks on the results).

The mixed effects model is fitted using the lmer() function from the package {lme4} (Bates et al., 2015). Two models are fitted and compared using the function anova():

(1) logRT ~ version + sentence_position + chunk_position + chunk_length + (1|participant)

(2) logRT ~ version + sentence_position + chunk_position + chunk_length + (1|participant) + (1|chunk)

Anova shows there is no difference between the performance of model 1 and performance of model 2 (p=0.317). That is, the type of modal chunk participants read does not affect the reading times. We are therefore reporting the results of the mixed effects model with only one random effect. The model's summary informs us that there is a negative correlation between aspectual preference and reading times, with the log reading times decreasing by 0.069698 when participant reads the chunk-preferred aspectual version of a modal chunk (std. error=0.038578, t=-1.807). The package {lme4} does not provide a p-value because it is not clear how degrees of freedom should be calculated in mixed effects linear models (Baayen, 2008, pp. 247-248). However, the absolute value of the t-statistic does not exceed 2, which means that the relationship is not significant (ibid.).

When an interaction between aspectual preferences and participant groups is added to the model, differences in how quickly the three groups read can be observed. Figure 10 visualises the interaction - the red dotted line is for ratings of modal chunks with the chunk-preferred aspectual version of the infinitive, while the black line is for chunks with the unigram-preferred aspectual version. It should be noted that the connecting lines do not suggest that a linear relationship exists between the three variables.

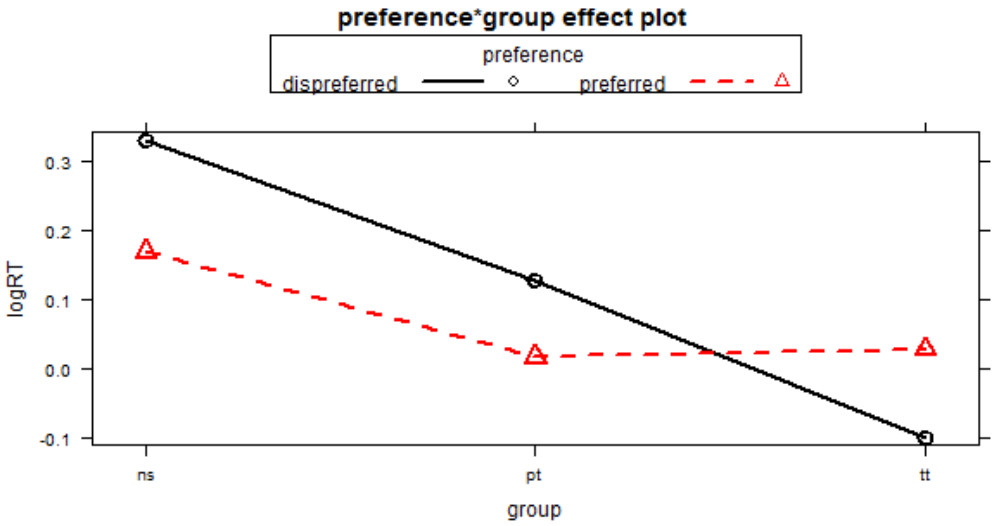

Figure 10: Self-paced reading - group differences

It is clear from the plot that the reading times of trainee translators are in direct opposition to the reading times of professional translators and naive native speakers. Trainee translators seem to be slightly quicker at reading unigram-preferred versions of modal chunks, while the other two groups seem quicker at reading the chunk-preferred versions. Regression models fitted to the data for individual groups confirm that there is no significant correlation between aspectual preference and reading times for trainee translators (estimate=0.147838, std. error=0.081586, t=1.812). That is, there is no difference in how trainee translators process chunk-preferred and unigram-preferred versions of modal chunks, suggesting that perhaps these modal chunks are not yet entrenched in their memory deeply enough to improve processing. Naive native speakers and professional translators, on the other hand, process the chunk-preferred versions of modal chunks significantly more quickly than the unigram-preferred versions:

- Naive native speakers: estimate= -0.149296, std. error=0.055816, t= -2.675)
- Professional translators: estimate= -0.1425370, std. error=0.0656853, t= -2.170)

Because trainee translators are the youngest group of participants (average age of 26, as compared to 38.9 for naive native speakers and 41 for professional translators), we also add an interaction between aspectual preferences and age of participants. Figure 11 shows that there is a clear relationship between the reading times and the age of participants - at around the age of 25, participants start off with reading the unigram-preferred aspectual versions of modal chunks slightly faster than the chunk-preferred versions, but the difference seems relatively small. At around the age of 33, participants seem to have been exposed to modal chunks enough for them to be entrenched and activated as prefabricated expressions, with the chunk-preferred aspectual versions of modal chunks requiring less cognitive load and being processed more quickly than the unigram-preferred versions. The older the participants, the larger the difference in the processing time of chunk-preferred and unigram-preferred versions of chunks, with the chunk-preferred forms taking significantly less time. The regression model confirms what the plot suggests: with age, the reading times of chunk-preferred versions of modal chunks decrease significantly by 0.012842 (std. error=0.003805, t= -3.375).

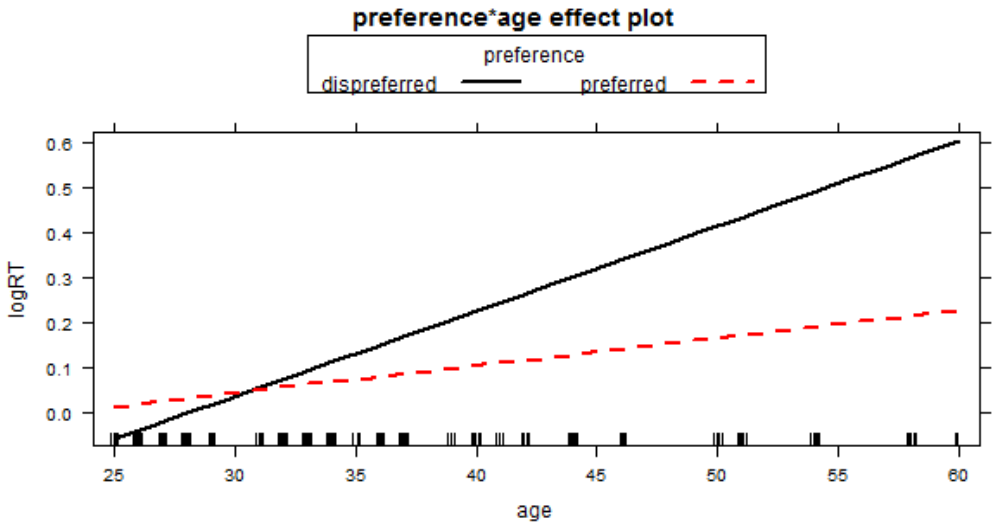

Figure 11: Self-paced reading - age differences

There clearly is support for the general chunking hypothesis from the self-paced reading task. Professional translators and naive native speakers seem to react significantly more quickly to chunk-preferred than to unigram-preferred aspectual versions of modal chunks. Trainee translators do not, but this may be influenced by their age.

We are also interested in the reading times of the chunks that follow the modal chunks - one of the challenges of analysing self-paced reading data is that the impact of an item on processing might only be noticeable when we look at the next item or two (Kaiser, 2013, p. 141). It therefore important to look for spillover effects in other parts of the sentence, not only the modal chunk. No such effects were found (estimate= -0.036831, std. error=0.039241, t= -0.939). There are also no significant effects when interactions with *group* and *age* are added.

7.2.1.3. Forced-choice task

For the purposes of the forced-choice task, general chunking hypothesis can be operationalized as follows:

The aspectual versions of modal chunks chosen by participants will correlate with chunk aspectual preferences.

The analysis of the forced-choice data differs from the analyses performed in the two previous tasks. In the judgement task and in the self-paced reading task, some participants saw a given modal chunk in either the chunk-preferred aspectual version or the unigram-preferred aspectual version - this allowed us to compare participants' reactions to the two different aspectual versions of the same modal chunk. In the forced-choice task, each participant saw both aspectual versions of the modal chunk, and their task was to choose the form they felt is more appropriate for a given context. As such, the aspectual version of a chunk that a participant sees cannot be used as a predictor of aspectual choice. Instead, we need to establish which type of aspectual preference correlates better with the choices made by participants. In this respect, the analysis of the forced-choice data will resemble the analysis of the corpus data, performed in Chapter 6 (cf. Divjak, Dąbrowska and Arppe 2016).

The response variable is the aspectual form chosen by participants - a binary variable called *choice*, which has two levels, *impf* or *pf*. The predictor variable will be the *chunk aspectual preference* of the given chunk, again with two levels - *impf* or *pf*. We are looking for a correlation between the chunk-preferred aspectual form and the choice of aspect. Given the hypothesis and the results of the previous tasks, we expect that the perfective versions of chunks will be chosen if the chunk aspectual preferences are for perfective. If a given modal chunk has chunk aspectual

preference for perfective then we know that this chunk's unigram aspectual preference is for imperfective. That is, because of the way the experiment is set up, if we find a positive correlation between chunk aspectual preference and choice of aspectual form, it will automatically imply a negative correlation between unigram aspectual preference and choice of aspectual forms. We therefore only fit models with chunk aspectual preferences as predictor - the results obtained for chunk aspectual preferences will imply the opposite tendency for unigram aspectual preferences.

Like in previous sections, random effects are added to control for the preferences of individual participants and characteristics of individual sentences. We fit two models using the glmer() function from the package {lme4} (Bates et al., 2015), which are then compared using the function anova():

(1) choice ~ chunk aspectual preference + (1|participant)

(2) choice ~ chunk aspectual preference + (1|participant) + (1|sentence)

The model with two random effects proved to be significantly better than the model with just one random effect (x^2=61.183, df=1, p=5.201e-15). The AIC of model 2 (397.6) is lower than that of the model 1 (456.8) by 59 points, which suggests that model 2 is significantly better than model 1. We are therefore reporting the results of model 2.

The model summary suggests a significant positive correlation (estimate=2.4076, std. error=0.777, p=0.00194) between the perfective aspectual form and the choice of the perfective version of a given modal chunk. That is, participants are significantly more likely to choose the perfective version of a modal chunk if the infinitive's

chunk aspectual preference is for perfective. This means that participants are less likely to choose the imperfective if the preference is for perfective. It would therefore suggest that choices of aspectual versions of modal chunks correlate with chunk aspectual preferences of modal chunks. When an interaction between aspectual preferences and participant groups is added to the model, no significant differences are observed. That is, all three groups of participants seem to choose aspectual versions of modal chunks similarly. There is also no significant correlation with the age of participants.

The forced choice task provides clear support for the general chunking hypothesis as there seems to be a clear positive correlation between chunk aspectual preference and the choice of an aspectual version of a modal chunk, which implies a negative correlation between unigram aspectual preference and the choice of aspectual version of a modal chunk.

7.2.1.4. Summary

All three tasks provide support for the general chunking hypothesis. First, we saw that ratings given to modal chunks by naive native speakers and professional translators correlate with chunk aspectual preferences. Participants that saw chunk-preferred versions of modal chunks were more likely to give those chunks a high rating as compared to participants who saw the same chunks but in the unigram-preferred aspectual version. The ratings given by trainee translators, however, do not correlate with either type of aspectual preferences. Second, we saw that reading times were shorter for chunk-preferred versions of modal chunks. Again, significant effects were observed for professional translators and naive native speakers, but not for trainee translators. We also observed a significant effect of age in the self-paced

reading task - the older the participants, the quicker the reading times of chunk-preferred versions of modal chunks. Finally, we saw a significant correlation between chunk aspectual preferences and aspectual choices for all three groups of participants.

The results would seem suggest that the modal verbs and the infinitives that follow them are entrenched in the speaker's memory as a prefabricated phrase, rather than as two separate items that are then combined into a phrase when required. In other words, it seems that the two items form a modal chunk, and are as such retrieved from memory and processed when a speaker comes across a modalized sentence.

Another interesting observation was made, which relates to the age of participants. The reading times of modal chunks change with age - around the age of 25, participants still read the unigram-preferred aspectual versions of the chunks slightly faster than the chunk-preferred versions, but the difference seems relatively small. At around the age of 33, participants start reading the chunk-preferred version faster, and the older they get, the more pronounced that difference becomes. It would therefore seem that with age, as participants become more exposed to modal chunks, the chunks become more entrenched in memory as prefabricated expressions, and are accessed and processed as such. This tendency was only observed in the self-paced reading task; no correlation with age was found in the judgement task or in the forced choice task. Self-paced reading is an online task that forces unconscious and automatic responses and therefore taps into implicit knowledge, while the other two tasks allow participants time to think about their responses and to use their explicit linguistic and metalinguistic knowledge. It may therefore be the case that the self-paced reading task is more sensitive picking up subtle differences in the way

chunking affects processing of linguistic input at an implicit level, than the other two tasks.

The role of age may help us understand why trainee translators - the youngest age group of all three participant groups - performed differently in the self-paced reading task and in the judgement task. It may be the case that they have not acquired sufficient linguistic experience to have these modal chunks entrenched deeply enough for the frequency information to affect processing. The other two participant groups - professional translators and naive native speakers - are on average at least 12.9 years older than trainee translators which means that they have had substantially more language exposure, perhaps leading to deeper entrenchment of modal chunks.

7.2.2. Genre-specific chunking hypothesis

The genre-specific chunking hypothesis stated that 'modal verbs and infinitives that follow them are entrenched in memory as chunks and the aspectual preferences of those modal chunks facilitate the selection and processing of aspect.' The three tasks provide little support for the hypothesis. It should be noted, however, that the experimental design placed great restrictions on the selection of stimuli. As a result, very few sentences from the translated and non-translated samples matched the selection criteria and we were able to gather very limited data from the judgement task (54 observations) and the self-paced reading task (51 observations). Additional stimuli were extracted from the National Corpus of Polish for the forced-choice task, but the data gathered was nevertheless limited (180 observations). It is therefore difficult to draw firm conclusions about the (lack of) existence of genre-specific chunking.

Results of the judgement task are discussed in Section 7.2.2.1, self-paced reading task in Section 7.2.2.2, and forced-choice task in Section 7.2.2.3.

7.2.2.1. Judgement task

For the purposes of the judgement task, the genre-specific chunking hypothesis can be operationalized as follows:

Ratings given to genre-specific aspectual versions of modal chunks by NSs and TTs will be significantly lower than ratings given by them to general aspectual versions of modal chunks. Ratings given to genre-specific aspectual versions of modal chunks by PTs will be similar to ratings given by them to general aspectual versions of modal chunks.

Rating is the response variable and *version* (of the chunk; legal or general) is the predictor variable. Like before, two models are fitted and compared, with the more accurate model selected for further analysis:

(1) rating ~ version + (1|participant)

(2) rating ~ version + (1|participant) + (1|chunk)

Interactions with *group* and *age* are also added. Anova results suggest that there is no difference between the two models so the results of model 1 are reported. The model summary shows that the odds of a rating above the cut-off point of 67.57894 increase significantly (estimate=2.4967, std. error=0.6646, p-value=0.000172) if participants rate modal chunks that occur in their legal aspectual form.

When an interaction between aspectual preferences and participant groups is added to the model, differences in how the three groups rate can be observed. Figure 12

visualises the interaction - the red dotted line is for ratings of the legal aspectual versions of modal chunks, while the black line is for the general aspectual versions of chunks. It should be noted that the connecting lines do not suggest that a linear relationship exists between the three variables.

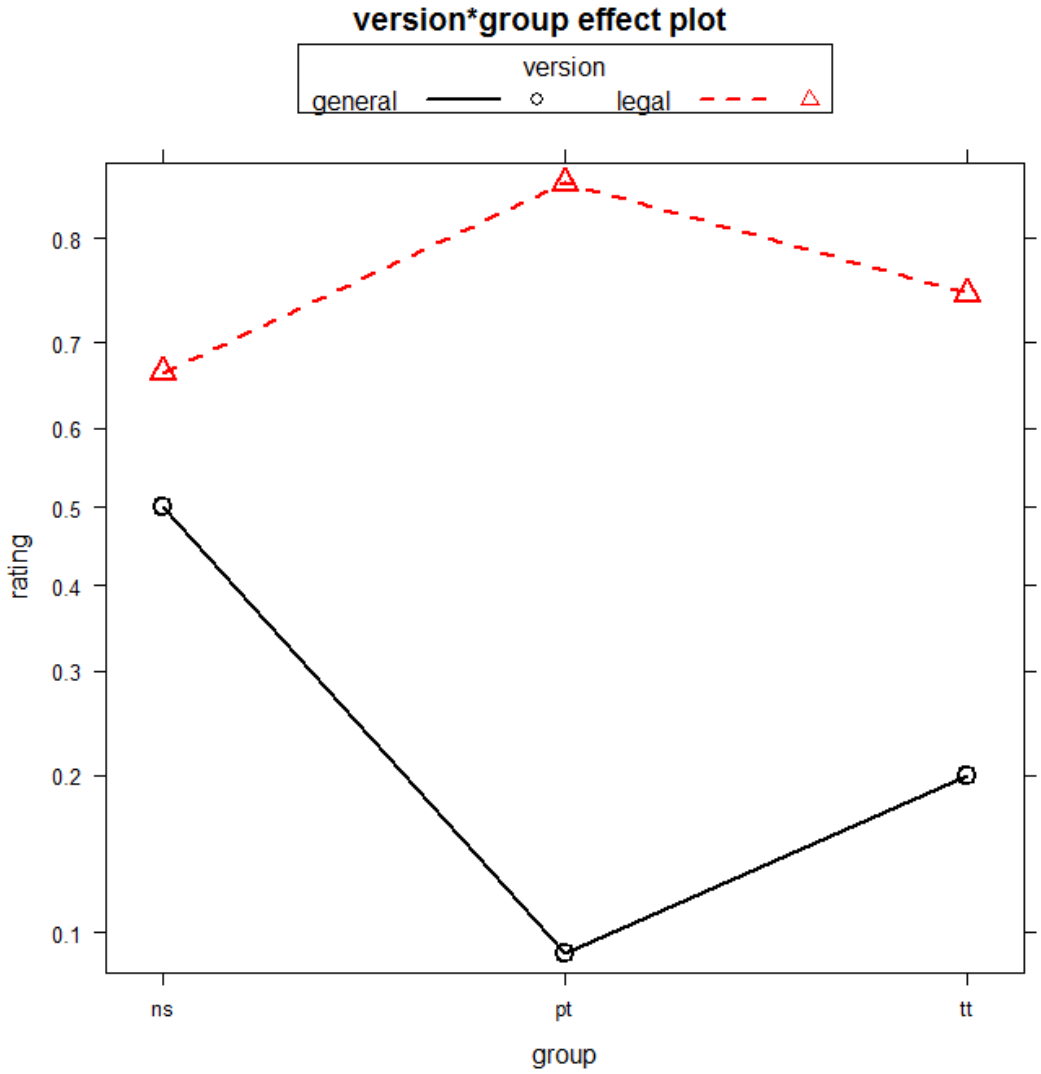

Figure 12: Judgement task - group differences

The ratings given to the two aspectual versions by naive native speakers do not differ significantly, with the legal chunk being higher up on the plot, suggesting higher (but not significantly) ratings (estimate=6.931e-01, std. error=1.082e+00, p-value=0.522). This suggests that the ratings given to the different versions of modal

chunks do not differ. The ratings given to legal and general chunks by trainee translators differ more substantially but again, not significantly (estimate=2.458, std. error=1.801, p-value=0.168). Surprisingly, in both cases the ratings given to the legal and general chunks do not differ, which is opposite to what we expected, given the genre-specific chunking hypothesis. The ratings given to legal chunks by professional translators are significantly higher than the ratings given by them to general chunks (estimate=3.977, std. error=1.223, p-value=0.00115). This would suggest that when exposed to modal chunks in legal contexts, professional translators find the genre-specific versions of modal chunks more acceptable than the general versions. This again contradicts the genre-specific chunking hypothesis, according to which these ratings should be similar.

7.2.2.2. Self-paced reading task

For the purposes of the self-paced reading task, genre-specific chunking hypothesis can be operationalized as follows:

The reading times of the genre-specific aspectual versions of modal chunks by NSs and TTs will be significantly slower than the reading times of the general aspectual versions of modal chunks The reading times of the genre-specific aspectual versions of modal chunks by PTs will be similar to the reading times of the general aspectual versions of modal chunks.

We therefore expect that the time it takes naive native speakers and trainee translators to read general modal chunks will be shorter than the time it takes them to read genre-specific modal chunks. Conversely, we expect that it will take professional translators the same time to read genre-specific and general modal chunks.

LogRT is the response variable and *version* is the predictor variable. A model with one random effect is fitted first (*participant*), then another random effect is added (*chunk*) - the more accurate model is selected for further analysis. An interaction with *group* is also added.

Both models perform similarly so the model with only one random effect is reported here. The model does not show any significant correlations between aspectual version and reading times (estimate=0.0127534, std. error=0.1359271, t=0.094). There are no significant correlations if we consider group and age differences. No spillover effect was observed in general (estimate=-0.0410355, std. error=0.1232737, t=-0.333) but with age, the reading of chunks that follow legal chunks tends to take longer (estimate=0.032802, std. error=0.012663, t=2.590).

7.2.2.3. Forced-choice task

For the purposes of the forced-choice task, the genre-specific chunking hypothesis can be operationalized as follows:

The aspectual versions of modal chunks chosen by participants will correlate with legal chunk aspectual preferences (for professional translators) or general chunk aspectual preferences (for trainee translators and naive native speakers).

We therefore expect that translators will choose the legal chunk-preferred version of the modal chunk significantly more frequently than the general chunk-preferred version of the modal chunk. Conversely, we expect that trainee translators and naive native speakers will choose the general chunk-preferred version of the modal chunk significantly more frequently than the legal chunk-preferred version of the modal chunk.

Like previously, the response variable is the aspectual form chosen by participants - a binary variable called *choice*, which has two levels, *impf* or *pf*. The predictor variable is also binary - the *legal chunk aspectual preference* of a given chunk, either *impf* or *pf*. We expect to observe a significant positive correlation for professional translators and a significant negative correlation for trainee translators and naive native speakers. This would mean that the odds of the perfective form being chosen by professional translators increases if the legal chunk aspectual preference is for perfective. It would also mean that the odds of the perfective being chosen by trainee translators and naive native speakers decrease significantly. General chunk aspectual preference are opposite to what legal chunk aspectual preferences are so a negative correlation with legal chunk aspectual preference would automatically imply a positive correlation with general chunk aspectual preference.

Two models are fitted - the model with two random effects (*participant* and *sentence*) performs significantly better than the model with one random effect (x^2=15.597, df=1, p=7.838e-05). There is, however, no significant effect of general chunk aspectual preference on the choice of aspect (estimate=-1.6509, std. error=1.0003, p=0.0988). That is, the model tells us that there are no correlations - positive or negative. There are no differences between the groups; the age of participants also plays no role. Therefore, it seems that no support for genre-specific chunking hypothesis was found in the forced-choice task.

7.2.2.4. Summary

Only one of the tasks - the judgement task - provides some support for experience-based differences in linguistic behaviour, but not how we expected on the basis of the genre-specific chunking hypothesis. We saw that professional translators gave

higher ratings to legal modal chunks than to general modal chunks even though we expected them to rate both similarly because they are familiar with both types of chunks equally. However, these results seem to suggest that modal chunks typical of a specific genre are perceived as more acceptable than general chunks if they occur in that genre, which confirms that genre-specific modal chunks are a psychologically plausible phenomenon that affects processing of aspect. No differences were observed in the ratings given by naive native speakers and trainee translators, which is surprising because the expectation was that they would rate the general chunks more favourably than the legal chunks. However, naive native speakers and trainee translators only saw one stimulus each so it is difficult to draw firm conclusions due to scarcity of data. No significant correlations were observed in the self-paced reading task and the forced-choice task, apart from the age correlation observed for the spillover effect.

All in all, the experimental analysis provided clear support for the general chunking hypothesis, but only incidental support for the genre specific chunking hypothesis was found in the ratings given to stimuli by professional translators. This is most likely due to the scarcity of data - in comparison with the data gathered for the general chunking hypothesis, the data for the genre-specific chunking hypothesis was very limited (see Table 40). The results obtained for the genre-specific chunking hypothesis are therefore tentative and the hypothesis requires further testing.

task	general chunking hypothesis	genre-specific chunking hypothesis
judgement task	276	54
self-paced reading task	274	51
forced-choice task	360	180

Table 40: Observations per task

7.3. Chunking and aspectual choices in *musieć* [must, have to]

Chunking of modal verbs and infinitives that follow them seems to be a plausible factor in aspect assignment in modal contexts, in which both aspectual forms are possible. Based on the results of the corpus and experimental investigations, we can be confident in saying that modal verbs and the infinitives that follow them are processed as non-compositional, prefabricated units, instead of being composed on the spot by separately retrieving the modal verb and separately adding the infinitive in the relevant aspectual form. Aspectual preferences of modal chunks seem to guide the choice of aspect to some extent when other factors, such as situation types, do not impose the aspectual form of the verb and the speaker is free to choose which form to use. When processing linguistic input, aspectual preferences of modal chunks seem to facilitate the processing of the chunks, which is reflected in shorter reading times (and therefore lower cognitive load) and higher perceived well-formedness of the more frequent aspectual versions of modal chunks. It is therefore safe to consider aspectual preferences as a factor that underlies the differences observed in the distribution of aspectual forms in the translated and non-translated *musieć* [must, have to] samples, as hypothesized in Section 5.3 of Part II. In this section, we check whether the differences in aspectual choices that cannot be explained with situation types and analogical mapping, can be explained with aspectual preferences for which evidence was found during the empirical investigation.

We first look at the aspectual preferences and aspectual choices in the non-translated *musieć* [must, have to] sample (Section 7.3.1) and in the translated *musieć* [must, have to] sample (Section 7.3.2). We then compare how the aspectual preferences in

the two samples differ from each other, and how well they can explain the different choices made in the translated and non-translated *musieć* [must, have to] (Section 7.3.3). If significant differences still remain between the translated and non-translated sample, we will be justified in looking at translation universals for explanation because we have considered other factors that we knew may have played a role.

7.3.1. Aspectual preferences vs. aspectual choices in non-translated *musieć* [must, have to]

The choice of aspect in non-translated samples is best predicted by legal chunk aspectual preferences, as shown in Part II. However, because the experimental analysis was not successful in confirming the psychological plausibility of legal chunks, we look at general chunk aspectual preferences to see whether they can explain the aspectual choices in the non-translated *musieć* [must, have to] sample.

Table 41 compares the aspectual forms of modal chunks observed in the sample, with aspectual preference of those modal chunks. We can see that the majority of the imperfective chunks have aspectual preference for imperfective (71%), and the majority of perfective chunks have preference for perfective (68%). In total, the majority of aspectual choices in the non-translated sample are in line with aspectual preferences of modal chunks (73%). Some modal chunks occur in the dispreferred aspectual form (21%) and some have no aspectual preferences at all (6%).

aspectual form of a modal chunk	number of observations	aspectual preferences	
imperfective	68	impf	48 (71%)
		impf/pf	6 (9%)
		pf	14 (20%)
perfective	38	impf	8 (21%)

		impf/pf	4 (11%)
		pf	26 (68%)

Table 41: General chunk aspectual preferences of modal chunks vs. aspectual versions chosen

A chi-squared test suggests that the difference between the aspectual preferences of imperfective modal chunks and aspectual preferences of perfective modal chunks is statistically significant (x^2=26.178, df=2, p=0.00000207). That is, significantly more modal chunks which occurred in the imperfective form have aspectual preference for imperfective (48) than modal chunks which occurred in the perfective form (8). At the same time, significantly more modal chunks which occurred in the perfective have aspectual preference for perfective (26) than modal chunks which occurred in the imperfective form (14). There is therefore a clear correlation between the aspectual versions of modal chunks chosen by authors and the aspectual preferences of those modal chunks.

Some modal chunks do not have any aspectual preferences. As mentioned in Section 6.1.2, it may be the case that speakers faced with the choice of aspectual version of a modal chunk that has no aspectual preferences are guided by the overall aspectual preference of that modal, in this case *musieć* [must, have to]. In other words, it may be the case that an abstracted MUSIEĆ(IMPF) schema is entrenched in the speakers' memory and guides the speakers' choice of aspectual form when no other clues are available. The results of the corpus analysis in Section 6.2 suggests that overall, when a verb has no aspectual preferences, there are significant odds that a perfective, rather than imperfective form, will occur. A model fitted to *musieć* data confirms that: when the verb in a *musieć* chunk has no aspectual preference, there are significant odds that this verb will occur in a perfective form (estimate=1.4888, std. error=0.5484, p=0.00663). We will therefore consider perfective the 'preferred' form

for chunks that had no aspectual preference in the corpus; this changes the proportions of explained and unexplained choices - 64% of choices are now explained, and 36% remain unexplained.

7.3.2. Aspectual preferences vs. aspectual choices in translated *musieć* [must, have to]

The choice of aspectual form in translated texts is best predicted by general chunk aspectual preferences, and the experimental analysis confirmed the psychological plausibility of general chunks (as shown in Part II). We are therefore comparing the aspectual choices made in the translated sample with the general chunk aspectual preferences of the modal chunks that occurred in this sample.

Table 42 shows that less than half of the imperfective chunks have aspectual preference for imperfective (34%), while the majority of perfective chunks have preference for perfective (86%). In total, almost two thirds of aspectual choices in the translated sample are in line with aspectual preferences of modal chunks (62%). Some modal chunks occur in the dispreferred aspectual form (27%) and some have no aspectual preferences at all (11%).

aspectual form of a modal chunk	number of observations	aspectual preferences	
imperfective	50	impf	17 (34%)
		impf/pf	7 (14%)
		pf	26 (52%)
perfective	57	impf	3 (5%)
		impf/pf	5 (9%)
		pf	49 (86%)

Table 42: General chunk aspectual preferences of modal chunks vs. aspectual versions chosen

Like with the non-translated sample, a chi-squared test performed on the translated sample also suggests that the difference between the aspectual preferences of imperfective modal chunks and aspectual preferences of perfective modal chunks is statistically significant (x^2=16.801, df=2, p=0.00022475). That is, significantly more modal chunks which occurred in the imperfective form have aspectual preference for imperfective (17) than modal chunks which occurred in the perfective form (3). At the same time, significantly more modal chunks which occurred in the perfective have aspectual preference for perfective (49) than modal chunks which occurred in the imperfective form (26). There is therefore a clear correlation between the aspectual versions of modal chunks chosen by translators and the aspectual preferences of those modal chunks.

Here, some modal chunks do not have any aspectual preferences either and like earlier, we will consider the perfective form to be the default form for this type of chunks. The proportions of explained and unexplained aspectual choices changes - 66% of choices are now explained, and 34% remain unexplained.

7.3.3. Comparison of translated and non-translated *musieć* [must, have to]

It is clear from the above that in approximately two thirds of the observations, aspectual choices made by authors and translators can be explained with the aspectual preferences of modal chunks. These observations can be taken out of the equation and we can compare the remaining unexplained aspectual choices to see if these differ significantly. If significant differences still remain between the translated and non-translated sample, we will be justified in looking at translation universals for explanation because we have considered other factors that we knew may have influence. Table 43 lists the unexplained aspectual choices in the two samples - the

non-translated sample contains 20 imperfective chunks that do not have aspectual preference for imperfective, and 8 perfective chunks that do not have preference for perfective. The translated sample contains 33 imperfective chunks that do not have preference for imperfective and 3 perfective chunks in the translated sample that do not have aspectual preference for perfective.

aspectual form observed	non-translated	translated
imperfective	20	33
perfective	8	3
total	28	36

Table 43: Unexplained aspectual choices in translated and non-translated *musieć* [must, have to]

The difference in distribution of aspectual forms chosen against their aspectual preferences in the two samples is above the critical $\alpha=0.05$ so we consider the difference statistically insignificant[24] ($x^2=3.22$, df=1, p=0.07265481). That is, even though there are choices of aspectual forms that cannot be explained with aspectual preferences, the distribution of these forms in the two *musieć* [must, have to] samples does not differ significantly and it is therefore safe to assume this difference is due to chance rather than influence of any factor in particular.

7.3.4. Conclusions

The current investigation of differences between translated and non-translated texts shows the importance of re-considering the usefulness of comparable corpora. Rather than immediately attributing the differences observed in such corpora to the effect of translation universals, it is important to consider the differences in the contents of the translated and non-translated components of comparable corpora. Three such factors were considered here - situation types, analogical mapping, and

[24] It should be noted that Yates correction had to be applied because at least 20% of expected frequencies are less than 5.

chunking. Two of the factors (situation types and chunking) helped explain the majority of the observed differences; the remaining differences are no longer statistically significant. By looking at differences in the situation types described in the translated and non-translated texts and at frequency effects that underlie the different verbs contained in them, we were able to explain the differences in aspect assignment. That is, we have demonstrated that comparable corpora lack sufficient comparability that would enable us to draw conclusions about the nature of translated language, answering the research questions stated in Section 1.3.1. This has important implications for Translation Studies, which are discussed in detail in Chapter 8. There are, however, certain limitations of the current investigation that need to be born in mind.

First, the phenomenon that formed the basis of the argument against translation universals is necessarily limited in scope - the evidence comes from one language (Polish), one genre (legal texts), and one linguistic feature (modality). As such, the results cannot be automatically generalised to other languages and other linguistic features, or treated as the ultimate proof for non-existence of translation universals. The investigation does, however, provide an interesting line of inquiry for future studies, with all of its implications relating to comparable corpora and to the role that language-general cognitive processes have in translation (see Chapter 8). Hopefully, similarly rigorous studies will follow, and the notion of translation universals will be tested in a scientifically sound manner by taking into consideration findings from neighbouring disciplines, such as linguistics, psychology, and cognitive sciences.

Second, the type of frequency information extracted here to demonstrate that chunking and entrenchment are likely to play a role in the differences between

translated and non-translated texts may be less psychologically plausible than other types of frequency that have been used, such as contextualised frequencies or probabilities. The aspectual preference of a verb was established based on the number of occurrences of one aspectual form of that verb in comparison with the other aspectual form in the National Corpus of Polish. This is a standard, but rather crude way to operationalize frequency - it only reflects the frequency with which a stimulus is repeated in the environment and may not reflect the way brain makes use of frequency of occurrence (Divjak & Caldwell-Harris, 2015, p. 57). Contextual diversity, semantic distinctiveness, attraction and reliance, or conditional probabilities, have proven more powerful than pure repetition, at least for some phenomena (ibid., pp. 58-60). However, although unigram aspectual preferences are decontextualized, chunk aspectual preferences can be treated as a form of a contextualised frequency because they take into consideration the frequency of each aspectual form of the verb, given the modal verb it follows. While chunk aspectual preferences predict aspectual choice in 45% of cases, it may be the case that these predictions would be more accurate if a more psychologically plausible operationalization of frequency was applied.

Finally, the experimental design has placed certain limitations on the selection of stimuli which resulted in little data being collected to test the genre-specific chunking hypothesis. As a result, the psychological plausibility of legal chunk aspectual preferences could not be confirmed, even though the corpus analysis showed that legal chunk aspectual preferences predict more variation in the non-translated data than general chunk aspectual preferences. Nevertheless, there is some indication that history of experience with legal genre may play a role in the processing of modal chunks, and this line of inquiry should therefore be investigated

further, with an experimental design that allows for more data to be obtained. This could involve less strict selection criteria or a more extensive search for stimuli that match the current criteria by using the entire corpus, rather than the twelve samples that were used here.

PART IV. CONCLUSIONS

Chapter 8. Implications

The aim of this thesis was to re-visit the notion of translation universals in order to assess its psychological plausibility from a usage-based perspective. The usage-based approach was adopted because its assumptions about the role of frequency of occurrence in language acquisition and use have been supported at almost all levels of linguistic analysis. The approach confirmed that chunking and entrenchment underlie some of the differences between translated and non-translated texts, which combined with the analysis of situation types described in the two sources of data, enabled us to rule out the role of translation universals in the choice of aspect in

modal contexts. Although explicitation and normalisation seemed like plausible explanations for the observed differences, differences in the type of information conveyed in translated and non-translated texts provided a more psychologically realistic account. The results should encourage re-consideration of comparable corpora as a method suitable for investigations of the nature of translated language and the development of more psychologically and linguistically realistic models of translational behaviour.

The results obtained here are important from both the perspective of Cognitive Linguistics and Translation Studies. For the former, the thesis supports the very basic assumption of the usage-based approach to language, i.e. that experience and general cognitive abilities inform the speaker's linguistic system. For the latter, it shows the importance of considering the differences in corpora of translated and non-translated texts in the investigation of the translation process. In the next sections, I discuss these implications in more detail.

8.1. Cognitive Linguistics

It was argued in Section 5.3 that based on what we know about the structure of implicit linguistic knowledge, and how it is shaped by usage and rooted in general cognitive abilities, it would be logical to assume that the choice of aspectual form, at least in situations in which either form is possible, is subject to frequency effects. The argument was based on the choice of aspectual form in modal context, i.e. when the verb in question (in its infinitive form) follows a modal verb, e.g. *musieć czytać* [must read]. Chunking of items that co-occur frequently has been shown to underlie many prefabricated expressions, constructions, conventionalised collocations, and even multi-word compositional phrases (Arnon & Snider, 2010; Bybee, 2010;

Divjak & Caldwell-Harris, 2015; Langacker, 2000). Since modal verbs are always followed by infinitives (with very few exceptions), it is possible that the two items will become entrenched in memory as a chunk. Moreover, since the majority of Polish verbs occurs in two aspectual forms, it is also possible that each form will be entrenched with the modal verb separately. In other words, it was argued that a perfective verb will form a separate chunk with the modal verb, and its imperfective equivalent will form a separate chunk. The more frequent aspectual version will be retrieved from memory more quickly, and will be chosen by the speaker when no other factors, such as the communicative intentions or the type of situation, are available to guide the speaker's choice of aspect. The corpus and experimental investigations carried out in Part II supported this hypothesis, which has importance from the point of view of two disciplines - it supports the very basic assumption of the usage-based approach to language that frequently co-occurring items form chunks and that they become entrenched as such in the speaker's memory, and it also suggests that assignment of aspect in Polish (and probably other Slavic languages too) may be subject to frequency effects, a possibility not investigated before. These two implications are discussed in the following two sections.

8.1.1. Frequency, chunking, entrenchment

The corpus and experimental analyses in Part II show that phrases that have traditionally been considered as computed - modal verb + infinitive - are subject to effects of chunking and entrenchment in the same way as idiomatic or conventionalised expressions. Aspectual preferences of verbs in modal chunks have been shown to predict the choice of aspect in modal context better than aspectual preferences of unigram verbs. That is, when choosing an aspectual form of a verb in

modal context, speakers seem to select the aspectual form that occurs in the modal chunk more frequently, rather than the form that occurs more frequently in non-modal contexts. The predictions made on the basis of corpus data were confirmed with experimental validation. Participants who saw the chunk-preferred aspectual version of a modal chunk rated it as more acceptable and read it more quickly than participants who saw the unigram-preferred aspectual version of the same modal chunk. Moreover, of the two aspectual versions of the same modal chunk, the chunk-preferred one was more likely to be chosen in a forced-choice task. All of this would seem to suggest that the processing of modal verbs and the infinitives that follow them can be understood more clearly if we consider the two items to form a chunk that is retrieved from memory as a prefabricated expression. That is, instead of the modal verbs and the infinitive being retrieved from memory separately and being computed as and when needed, the two seem to be entrenched in memory as a chunk. This prefabricated modal chunk is retrieved and processed more efficiently than would be the case if the modal verb and the infinitive that follows it were retrieved separately and then computed into a phrase. This provides strong support for chunking of frequently co-occurring items, including phrases that have traditionally been thought of as computed. This in turn supports the single-system theories, including the usage-based approach, which argues that mental grammar and mental lexicon are not two separate systems but a continuum (see Langacker, 2008, p. 15).

The corpus analysis in Part II also suggested that experience with different genres may play a role in what type of frequency information is retrieved from memory and applied when aspectual choices are made. It was shown that the aspectual choices made by authors of non-translated legal texts were predicted by legal chunk

aspectual preferences better than general chunk aspectual preferences, while translators' choices were better predicted by general chunk preferences. This may have occurred because the authors of non-translated legal texts have more experience with legal language than the translators (see Section 4.1.1 for the discussion about the experience of translators working on the pre-2004 EU texts that constitute over 50% of the translated corpus used in this investigation). Based on the results of previous studies, which have shown that speakers with different usage histories are faster at processing items that are specific to their area of expertise or experience (e.g. Caldwell-Harris, Berant, & Edelman, 2012), it is reasonable to argue that translators did not have enough experience with legal language to have these genre-specific chunks entrenched in memory, so their aspectual choices were better predicted by general chunk aspectual preferences. Authors of non-translated legal texts, on the other hand, are assumed to have substantial experience with legal texts (being the legislator), so their aspectual choices were better predicted by legal chunk aspectual preferences. These results provide clear support for the usage-based assumption that language is shaped by usage and that different linguistic experience can result in different structures being formed in speakers' minds (see Dąbrowska, 2015).

8.1.2. Aspect

Aspect can be expressed in various ways by different languages - through lexical, contextual, or grammatical features (Bermel, 1997, p. 25). In Polish, the majority of verbs are morphologically marked for aspect, which means that they exist in two forms - the perfective and the imperfective. Some verbs are biaspectual (expressing both aspects at the same time), and some exist in the perfective or the imperfective

form only. Perfective verbs describe actions that have clear, definitive boundaries and are viewed as a whole, as if from the outside. The focus is on completion and totality of the action, with a change of state. Imperfective verbs, on the other hand, are said to describe actions that are unbounded in time and are in progress; they focus on the duration of the action, as if looking at it from within. Other, more cognitively-inspired models of aspect have been proposed, including Janda (2004), which is based on the idealized cognitive model (ICM) of matter. There are also studies that looked at the distributional correlations of aspectual forms with other linguistic variables, such as polarity, modality type, and state of affairs applicability (Divjak, 2009; Wiemer, 2001). This thesis argued, however, that the choice of aspectual form, at least in situations in which either form is possible, may be subject to frequency effects.

In order to test this argument, frequencies of occurrence of the two aspectual forms of modal chunks were extracted and used to establish whether the more frequent aspectual versions can contribute significantly to explaining aspectual choices. In other words, we checked whether there are any significant correlations between the more frequent aspectual versions of modal chunks and the aspectual versions of modal chunks actually chosen. The corpus predictions confirmed that this is indeed the case - in 45% of all cases the choice of aspect can be predicted with the aspectual preference of a modal chunk alone. It is important to note that the frequencies and choices did not correlate to the full extent and that frequencies cannot be relied on in isolation when predicting aspectual choice, or perhaps a different type of frequency information should be checked (as suggested in Section 7.3.4). That is, there are other factors - or perhaps other ways of measuring frequency - that play a role in choosing aspectual versions of verbs when either form is possible. The role of

frequency in the choice of aspect was also tested experimentally - participants rated the more frequent aspectual version of a modal chunk as more acceptable and read it more quickly than participants who saw the less frequent aspectual version of the same modal chunk. Moreover, of the two aspectual versions of the same modal chunk, the more frequent one was more likely to be chosen in a forced-choice task. All of this suggests that frequency - as measured in this thesis - should be seriously considered in any future attempts at modeling aspect.

8.2. Translation Studies

There are two important conclusions that can be drawn from the investigation presented in this thesis: one relates to the psychological plausibility of translation universals and the other to comparability of comparable corpora and their usefulness in investigations of the process of translation.

8.2.1. Comparability of comparable corpora

The analyses in Part I and II enabled us to answer the research question discussed in Section 1.3.1. It was argued there that if we want to make sure that the language-external explanations (cf. Halverson 2003) for atypical linguistic behaviour of translators observed in comparable corpora are psychologically and biologically plausible, we first need to ensure that the atypical features observed do not result from the limited comparability of the corpora used.

It was shown in Chapter 5 that the majority of differences between translated and non-translated texts observed in the corpus analysis in Chapter 4 occurred due to differences in the type of situations described in the translated and non-translated samples: the situations or information conveyed in the translated samples required

the use of perfective verbs more often than those in non-translated samples. In Chapter 6 and 7, the remaining differences in aspectual choices were shown to have occurred because the translated and non-translated samples contained verbs with varying aspectual preferences: the former contained verbs that more frequently occur in perfective form than the latter, resulting in differences in the distribution of the aspectual forms. Had the type of situations and frequency effects not been checked, misleading conclusions about the role of translation universals in the differences would have been drawn. The analysis has therefore shown that comparable corpora may lack the necessary comparability that would allow translation scholars to draw reliable conclusions about the nature of the translation process. It is therefore necessary to either reconsider the application of corpus methods in such investigations or include a careful assessment of the content of the two corpora before explaining any observed differences by positing translation universals.

Based on the issues associated with corpus-based methods, discussed in details in Chapter 1, as well as the results of the current investigation, we can identify a number of methodological steps that would allow us to reliably investigate the source of any differences observed in comparisons of translated and non-translated texts. Such steps include analysing the following:

1. **Source language/culture conventions**. It may be the case that certain features of translated texts result from shining through of source language/culture conventions. For example, increased use of formal language in translated texts as compared to non-translated texts may result from the fact that a given type of text in the source language requires a more formal use of language than its equivalent type of text in the target language. See Section 1.1 for more details.

2. **S-universals**[25]. It may be the case that certain features of translated texts result from strategies employed to deal with problematic items present in the source texts. For example, increased explicitness in translated texts as compared to non-translated texts may result from implicitness in the source texts. The implicitly encoded information in the source text might be encoded explicitly in the target text in order to avoid ambiguity, to reduce communicative risk, or because the target language requires more explicit encoding for a given item. See Section 1.2 for more details.

3. **Differences between translated and comparable non-translated corpora**. It may be the case that certain features of translated texts result from differences in the type of information conveyed in the translated and non-translated texts. For example, the increased use of perfective forms in the translated texts as compared to non-translated texts resulted from the fact that the translated texts describe situations that impose the perfective form of a verb more often than non-translated texts. We saw in Section 5.1 that situation types helped to explain the majority of differences in the distribution of aspectual forms in translated and non-translated texts. Moreover, the implicit linguistic knowledge that underlies the language structures included in the analysed texts should also be considered. It may be the case that certain features of translated texts result from the linguistic mechanisms that underlie the production of the investigated items, including a variety of frequency effects. For example, the increased use of perfective infinitives in the translated *musieć* [must, have to] sample as compared to the non-translated *musieć* [must, have to] sample resulted from the fact that infinitives and modal verbs form chunks and the more frequent aspectual versions of modal chunks are more deeply entrenched in speakers' memory and therefore activated more quickly. The translated sample contained more modal

[25] Here, s-universals are understood as strategies employed by translators to deal with problems in the source texts, as argued in Section 1.2.1.2.

chunks that are more frequent in the perfective form than the non-translated sample, and that is why there were more perfective infinitives in the translated sample. That is, the two corpora differ in the types of verbs contained in them, which has resulted in certain differences being observed (see Chapter 7 for more details). Processes and mechanisms related to the organisation of implicit linguistic knowledge in bilingual speakers are also likely to play a role, as discussed in Section 1.3.4, and therefore need to be considered too.

Only if the influence of the above factors on the differences between translated and non-translated texts can be ruled out, should researchers look for inherent and unique properties of the translation process that could explain the observed differences. If any of the above can be said to underlie the differences, attributing them to translation universals will be unjustified. Alternatively, the currently used corpus methods could be replaced by experimental investigations that would elicit linguistic input from translators and non-translators responding to the same stimuli, thus enabling researchers to compare utterances describing the same type of information and therefore directly compare the cognitive processing that underlies them.

8.2.2. Psychological plausibility of translation universals

The notion of translation universals was first analysed by looking at its plausibility from the point of view of implicit and explicit linguistic knowledge, and conscious and unconscious cognitive processing. It was argued in Section 1.2 that linguistic behaviour in translation can either be translation-specific, but not universal at the same time, or it can be universal in the cognitive-linguistic sense, but not translation-specific. This is because the decisions that are specific to translation are made in the mode of conscious cognitive processing, which is subject to individual differences,

ranging from general intelligence to differences in educational histories. These decisions cannot therefore be considered universal. At the same time, linguistic decisions that can be considered universal because they take place in the mode of unconscious cognitive processing and result from implicit linguistic knowledge, cannot be translation-specific in the professional sense. Implicit linguistic knowledge results from general cognitive abilities which are universally shared by all humans, whether bilingual or monolingual. Although bilingual speakers exhibit certain linguistic mechanisms not exhibited by monolingual speakers, such as code-switching, the knowledge of translation theory, procedures, and issues acquired in the course of professional training and experience is not one of them. This means that even if certain linguistic behaviour exhibited by translators differs from the behaviour exhibited by monolingual speakers, it cannot be inherent to translation in the professional sense because bilingual speakers who are not translators will also exhibit this behaviour. It was also argued that the supposed evidence supporting the notion, observed in comparisons of translated and non-translated texts, may be misleading due to limited comparability of those texts, casting further doubt on the existence of translation universals.

The abovementioned theoretical claims were tested and confirmed empirically in Part II and III, where the differences observed in the comparison of translated and non-translated texts were analysed by assessing the comparability of the samples. The investigations provide evidence that translated and non-translated texts used for comparison of translated and non-translated language differ at a very fine-grained level of linguistic analysis, not previously considered. It may therefore be the case that no support for translation universals would have been found if translation

scholars considered the use of comparable corpora more carefully, as was done in this thesis.

Although the current study tested the comparability of comparable corpora by investigating the use of modal verbs in Polish legal texts so the results cannot be treated as the ultimate proof for non-existence of translation universals, it offers a promising avenue of research for future investigations. Such investigations should employ the methodological steps listed in Section 8.2.1 and should involve a wide variety of languages, genres, and linguistic phenomena. Ideally, previous studies that claimed to support translation universals should also be re-visited from the point of view of comparability of the corpora they use to verify whether the differences observed in them truly result from the process of translation rather than differences in the type of information conveyed in them. This type of methodologically rigorous work will hopefully allow researchers to resolve the question of translation universals once and for all.

Bibliography

Aarts, H., & Dijksterhuis, A. (2000). Habits as Knowledge Structures: Automaticity in Goal-Directed Behavior. *Journal of Personality and Social Psychology, 78*(1), 53-63.

Alves, F., Pagano, A., & Silva, I. d. (2010). A new window on translators' cognitive activity: methodological issues in the combined use of eye tracking, key logging and retrospective protocols. In I. M. Mees, F. Alves, & S. Göpferich (Eds.), *Methodology, Technology and Innovation in Translation Process Research. A Tribute to Arnt Lykke Jakobsen* (pp. 297-291). Copenhagen: Samfundslitteratur.

Ambridge, B., Kidd, E., Rowland, C. F., & Theakston, A. L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language, 42*, 239-273.

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language, 62*, 67-82.

Arppe, A., & Jarvikivi, J. (2007). Every method counts: Combining corpus-based and experimental evidence in the study of synonymy. *Corpus Linguistics and Linguistic Theory, 3*(2), 131-159.

Arppe, A., Gilquin, G., Glynn, D., Hilpert, M., & Zeschel, A. (2010). Cognitive Corpus Linguistics: five points of debate on current theory and methodology. *Corpora, 5*(1), 1-27.

Baayen, R. H. (2008). *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R.* Cambridge: Cambrudge University Press.

Baayen, R. H. (2010). Demythologizing the word frequency effect: A discriminative learning perspective. *The Mental Lexicon, 5*, 436-461.

Baker, M. (1993). Corpus Linguistics and Translation Studies - Implications and applications. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and Technology. In Honour of John Sinclair* (pp. 233-252). Philadelphia/Amsterdam: John Benjamins Publishing Company.

Baker, M. (1996). Corpus-based translation studies: the challenges that lie ahead. In *Terminology, LSP and Translation: Studies in Language Engineering, in Honour of Juan C. Sager* (pp. 175-186). Amsterdam: John Benjamins.

Baker, M. (2004). A corpus-based view of similarity and difference in translation. *International Journal of Corpus Linguistics , 9*(2), 167-193.

Barlow, M., & Kemmer, S. (2000). *Usage-Based Models of Language.* Stanford, California: CSLI Publications.

Bassetti, B., & Cook, V. (2011). Relating language and cognition: The second language user. In V. Cook, & B. Bassetti (Eds.), *Language and Bilingual Cognition* (pp. 143-190). New York/Hove: Psychology Press.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Baumgarten, N., Meyer, B., & Özçetin, D. (2008). Explicitness in Translation and Interpreting: A Critical Review of Some Empirical Evidence (of an Elusive Concept). *Across Languages and Cultures, 9*(2), 177-203.

Becher, V. (2010). Abandoning the Notion of "Translation-Inherent" Explicitation: Against a Dogma of Translation Studies. *Acorss Languages and Cultures, 11*(1), 1-28.

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

Benaglia, T., Chauveau, D., Hunter, D. R., & Young, D. (2009). mixtools: An R Package for Analyzing Finite Mixture Models. *Journal of Statistical Software, 32*(6), 1-29.

Berez, A. L., & Gries, S. T. (2009). In defense of corpus-based methods: A behavioral profile analysis of polysemous get in English. In *Proceedings of the 2008 Northwest Linguistics Conference* (pp. 157-166). Seattle, WA: Department of Linguistics.

Bermel, N. (1997). *Context and the lexicon in the development of Russian aspect.* Berkeley/London: University of California Press.

Bermel, N., & Knittl, L. (2012). Corpus frequency and acceptability judgements: A study of morphosyntactic variants in Czech. *Corpus Linguistics and Linguistic Theory, 8*(2), 241-275.

Bermel, N., Knittl, L., & Russell, J. (2015). Morphological variation and sensitivity to frequency of forms among native speakers of Czech. *Russian Linguistics, 39*, 283-308.

Bernardini, S., & Ferraresi, A. (2011). Practice, Description and Theory Come Together – Normalization or Interference in Italian. *Meta: Translators' Journal, 56*(2), 226-246.

Bernardini, S., & Zanettin, F. (2004). When is a universal not a universal? Some limits of current corpus-based methodologies for the investigation of translation universals. In A. Mauranen, & P. Kujamäki, *Translation Universals. Do they exist?* (pp. 51-62). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Biber, D. (1993). Representativness in Corpus Design. *Literary and Linguistic Computing, 8*(4), 243-257.

Biber, D., & Jones, J. K. (2009). Qunatitative methods in corpus linguistics. In A. Lüdeling, & M. Kytö (Eds.), *Corpus Linguistics: An International Handbook, Volume 2* (pp. 1287-1304). Berlin/New York: Walter de Gruyter.

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at...: Lexical Bundles in University Teaching and Textbooks. *Applied Linguistics, 25*(3), 371-405.

Błaszczyk-Szabat, A. (2005). The relationship between inherent aspect and past tense in the early and late acquisition of L2 Polish. *Poznań Studies in Contemporary Linguistics, 40*, 91-110.

Blumenthal-Drame, A. (2012). *Entrenchment in Usage-Based Theories.* Berlin/Boston: De Gruyter Mouton.

Blum-Kulka, S. (1986). Shifts of Cohesion and Coherence in Translation. In J. House, & S. Blum-Kulka (Eds.), *Interlingual and Intercultural Communication. Discourse and Cognition in Translation and Second Language Acquisition Studies* (pp. 17-36). Tübingen: Gunter Narr.

Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference. A Practical Information-Theoretic Approach.* New York: Springer-Verlag.

Bybee, J. (2007). *Frequency of Use and the Organization of Language.* Oxford, New York: Oxfor University Press.

Bybee, J. (2010). *Language, Usage and Cognition.* Cambridge: Cambridge University Press.

Bybee, J. L., & Fleischman, S. (1995). *Modality in Grammar and Discourse.* Amsterdam/Philadelphia: John Benjamins.

Bybee, J. L., & Hopper, P. J. (2001). *Frequency and the Emergence of Linguistic Structure.* John Benjamins.

Bybee, J. L., Perkins, R. D., & Pagliuca, W. (1994). *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World.* Chicago: University of Chicago Press.

Caldwell-Harris, C., Berant, J., & Edelman, S. (2012). Measuring Mental Entrenchment of Phrases with Perceptual Identification, Familiarity Ratings, and Corpus Frequency Statistics. In D. Divjak, & S. T. Gries (Eds.), *Frequency Effects in Language Representation* (pp. 165-194). Berlin/Boston: De Gruyter Mouton.

Chesterman, A. (2004). Beyond the particular. In A. Mauranen, & P. Kujamäki (Eds.), *Translation Universals. Do they exist?* (pp. 33-50). Amsterdam/Philadelphia: John Benjamins.

Chesterman, A. (2010). Why study translation universals? *Acta Translatologica Helsingiensia, 1*, 38-48.

Coates, J. (1995). The expression of root and epistemic possibility in English. In J. Bybee, & S. Fleischman (Eds.), *Modality in grammar and discourse* (pp. 55-66). Amsterdam/Philadelphia: John Benjamins.

Comrie, B. (1976). *Aspect. An Introduction to the Study of Verbal Aspect and Related Problems.* Cambridge: Cambridge University Press.

Cook, V. (2003). Introduction: The Changing L1 in the L2 User's Mind. In V. Cook (Ed.), *Effects of the Second Language on the First* (pp. 1-18). Clevedon: Multilingual Matters.

Crawley, M. J. (2015). *Statistics. An introduction using R.* Chichester: John Wiley & Sons.

Dąbrowska, E. (2010). Naive v. expert intuitions: An empirical study of acceptability judgments. *The Linguistic Review, 27*, 1-23.

Dąbrowska, E. (2015). Individual differences in grammatical knowledge. In E. Dąbrowska, & D. Divjak (Eds.), *Handbook of Cognitive Linguistics* (pp. 650-667). Berlin/Boston: Walter de Gruyter.

Dąbrowska, E., & Divjak, D. (2015). *Handbook of Cognitive Linguistics.* Berlin/Boston: Walter de Gruyter.

Dayrell, C. (2008). Investigating the Preference of Translators for Recurrent Lexical Patterns: A Corpus-based Study. *Trans-kom, 1*(1), 36-57.

De Sutter, G., & Van de Velde, M. (2008). Do the mechanisms that govern syntactic choices differ between original and translated language? A corpus-based translation study of PP placement in Dutch and German. In R. Xiao, L. He, & M. Yue (Ed.), *Proceedings of the International Symposium on Using Corpora in Contrastive and Translation Studies (UCCTS)*, (pp. 1-38). Hangzhou.

Delaere, I., De Sutter, G., & Plevoets, K. (2012). Is translated language more standardized than non-translated language? Using profile-based correspondence analysis for measuring linguistic distances between language varieties. *Target, 24*(2), 203-224.

Deshors, S. C., & Gries, S. T. (2014). A case for the multifactorial assessment of learner language: the uses of may and can in French-English interlanguage. In D. Glynn, & J. Robinson (Eds.), *Corpus methods for semantics: quantitative studies in polysemy and synonymy* (pp. 179-204). Amsterdam & Philadelphia: John Benjamins.

Dienes, Z. (2012). Conscious versus unconscious learning of structure. In P. Rebuschat, & J. N. Williams (Eds.), *Statistical Learning and Language Acquisition* (pp. 337-364). Boston/Berlin: Walter de Gruyter.

Dilin, L. (2012). Is it a chief, main, major, primary, or principal concern? A corpus-based behavioral profile study of the near-synonyms and its implications. *International Journal of Corpus Linguistics, 15*, 56-87.

Divjak, D. (2004). *Degrees of verb integration. Conceptualizing and categorizing events in Russian.* Katkolieke Universiteit Leuven, Belgium: Unpublished PhD dissertation.

Divjak, D. (2009). Mapping between domains. The aspect–modality interaction in Russian. *Russian Linguistics, 33*(3), 249-269.

Divjak, D. (2011). Predicting Aspectual Choice in Modal Constructions: A Quest for the Holy Grail? In *Slavic Linguistics in a Cognitive Framework* (pp. 67-86). Frankfurt a. M.: Lang.

Divjak, D., & Arppe, A. (2013). Extracting prototypes from exemplars. What can corpus data tell us about concept representation? *Cognitive Linguistics, 24*(2), 221-274.

Divjak, D., & Caldwell-Harris, C. (2015). Frequency and entrenchment. In E. Dąbrowska, & D. Divjak (Eds.), *Handbook of Cognitive Linguistics* (pp. 53-74). Berlin/Boston: Walter de Gruyter GmbH.

Divjak, D., & Gries, S. T. (2006). Ways of Trying in Russian. Clustering Behavioral Profiles. *Journal of Corpus Linguistics and Linguistic Theory*, *2*(1), 23-60.

Divjak, D., & Gries, S. T. (2006). Ways of Trying in Russian. Clustering Behavioral Profiles. *Journal of Corpus Linguistics and Linguistic Theory*, *2*(1), 23-60.

Divjak, D., & Gries, S. T. (2008). Clusters in the mind? Converging evidence from near-synonymy in Russian. *The Mental Lexicon, 3*(2), 59-81.

Divjak, D., & Gries, S. T. (2009). Corpus-based Cognitive Semantics: A Contrastive Study of Phrasal Verbs in English and Russian. In B. Lewandowska-Tomaszczyk, & K. Dziwirek (Eds.), *Studies in Cognitive Corpus Linguistics* (pp. 273-296). Frankfurt a. M. : Lang.

Divjak, D., Dąbrowska, E., & Arppe, A. (2016). Machine Meets Man: Evaluating the Psychological Reality of Corpus-based Probabilistic Models. *Cognitive Linguistics, 27*(1), 1-33.

Divjak, D., Szymor, N., & Socha-Michalik, A. (2015). Less is more: possibility and necessity as centres of gravity in a usage-based classification of core modals in Polish. *Russian Linguistics, 39*(1), 327-349.

Dragsted, B. (2010). Coordination of reading and writing processes in translation. In G. M. Shreve, & E. Angelone (Eds.), *Translation and Cognition* (pp. 41-61). Amsterdam/Philadelphia: John Benjamins.

European Union. (2010). *Lawmaking in the EU multilingual environment.* Luxembourg: Publications Office of the European Union.

Evans, J. S. (2008). Dual-Processing Accounts of Reasoning, Judgement, and Social Cognition. *Annual Review of Psychology, 59*, 255-278.

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences, 32*, 429-492.

Evans, V., & Green, M. (2006). *Cognitive Linguistics. An Introduction.* Edinburgh: Edinburgh University Press.

Faber, D., & Hjort-Pedersen, M. (2009). Manifestations of inference processes in legal translation. In S. Göpferich, A. L. Jakobsen, & I. M. Mees (Eds.), *Behind the Mind: Methods. Models and Results in Translation Process Research* (pp. 107-124). Copenhagen Studies in Language 37.

Featherston, S. (2008). Thermometer judgements as linguistic evidence. In C. M. Riehl, & A. Rothe (Eds.), *Was ist linguistische Evidenz?* (pp. 69-89). Aachen: Shaker Verlag.

Fox, J. (2003). Effect Displays in R for Generalised Linear Models. *Journal of Statistical Software, 8*(15), 1-27.

Fox, J., & Weisberg, S. (2011). *An {R} Companion to Applied Regression, Second Edition.* Thousand Oaks CA: Sage.

Gilquin, G., & Gries, S. T. (2009). Corpora and experimental methods: A state-of-the-art review. *Corpus Linguistics and Linguistic Theory, 5*(1), 1-26.

Gries, S. T. (2014). Frequency tables. Tests, effect sizes, and explorations. In D. Glynn, & J. A. Robinson (Eds.), *Corpus Methods for Semantics. Quantitative studies in polysemy and synonymy* (pp. 365-390). Amsterdam/Philadelphia: John Benjamins.

Gries, S. T., & Divjak, D. (2010). Quantitative approaches in usage-based Cognitive Semantics: Myths, erroneous assumptions, and a proposal. In D. Glynn, & K. Fischer (Eds.), *Quantitative Methods in Cognitive Semantics. Corpus-driven approaches* (pp. 333-354). Berlin/New York: Mouton de Gruyter.

Groot, G.-R. d. (1987). The point of view of a comparative lawyer. *Les Cashiers de droit, 28*(4), 793-812.

Grzegorczykowa, R. (2001). *Wprowadzenie do semantyki jezykoznawczej.* Warszawa: Wydawnictwo Naukowe PWN.

Halverson, S. L. (2003). The cognitive basis of translation universals. *Target, 15*(2), 197-241.

Halverson, S. L. (2013). Implications of Cognitive Linguistics for Translation Studies. In A. Rojo, & I. Ibarretxe-Antuñano (Eds.), *Cognitive Linguistics and Translation. Advances in some theoretical models and applications.* (pp. 33-74). Berlin/Boston: Walter de Gruyter GmbH.

Halverson, S. L. (2017). Gravitational pull in translation: testing a revised model. In G. De Sutter, M.-A. Lefer, & I. Delaere, *Empirical Translation Studies. New*

Methodological and Theoretical Traditions (pp. 9-46). Berlin, Boston: De Gruyter Mouton.

Hansen, B. (2004). Modals and the boundaries of grammaticalization: the case of Russian, Polish and Serbo-Croatian. In W. Bisang, N. P. Himmelmenn, & B. Wiemer (Eds.), *What makes grammaticalization: A look from its Fringes and Components* (pp. 245-270). Berlin: Walter de Gruyter.

Hansen, S., & Teich, E. (2001). Multi-layer analysis of translation corpora: methodological issues and practical implications. In D. Cristea, N. Ide, D. Marcu, & M. Poesio (Eds.), *Proceedings of EUROLAN 2001 Workshop on Multi-layer Corpus-based Analysis.* (pp. 44-55). Iasi.

Harrell, F. (2001). *Regression Modeling Strategies. With Applications to Linear Models, Logistic Regression, and Survival Analysis.* New York: Springer-Verla.

Hartigan, J. A., & Hartigan, P. M. (1985). The Dip Test of Unimodality. *The Annals of Statistics, 13*(1), 70-84.

Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulleting & Review, 9*(4), 780-789.

Hengeveld, K. (2004). Illocution, mood and modality. In G. Booij, C. Lehmann, J. Mudgan, & S. Skopeteas (Eds.), *An International Handbook on Inflection and Word-Formation Vol. 2* (pp. 1190-1201). Berlin/New York: De Gruyter.

Hofmann, T. R. (1976). Past tense replacement and the modal system. In J. McCawley (Ed.), *Syntax and semantics. Volume 7: Notes from the linguistic underground* (pp. 85-100). New York: Academic Press.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression* (2nd ed.). New York: John Wiley & Sons.

House, J. (2008). Beyond Intervention: Universals in Translation? *Trans-kom, 1*(1), 6-19.

Janda, L. (2004). A metaphor in search of a source domain: The categories of Slavic aspect. *Cognitive Linguistics, 15*(4), 471-527.

Janda, L. A., Eckhoff, H. M., Lyashevskaya, O., & Reynolds, R. J. (2016). How learnable is Russian aspect? *Time and Language Conference.*

Jansegers, M., Vanderschueren, C., & Enghels, R. (2015). The polysemy of the Spanish verb sentir: a behavioral profile analysis. *Cognitive Linguistics, 26*(3), 381-421.

Juliano, C., & Tanenhaus, M. K. (1993). Contingent frequency effects in syntactic ambiguity resoluti. In *Proceedings of the 15th annual conference of the cognitive science society, COGSCI-94* (pp. 593-598). Boulder: Psychology Press.

Jurafsky, D. (1996). A Probabilistic Model of Lexical and Syntactic Access and Disambiguation. *Cognitive Science, 20*, 137-194.

Kaiser, E. (2013). Experimental paradigms in psycholinguistics. In R. J. Podesva, & D. Sharma (Eds.), *Research Methods in Linguistics* (pp. 135-168). Cambridge: Cambridge University Press.

Kaleta, Z. (1995). *Gramatyka języka polskiego dla cudzoziemców.* Kraków: Uniwersytet Jagielloński.

Kancelaria Prezesa Rady, M. (2001). *Raport o stanie państwa i działaniach rządu w latach 1997-2001*. Retrieved from http://kuczyn.com/papers/sluzba/raportbuzka_20050923.pdf

Karlsson, F. (2009). Early generative linguistics and empirical methodology. In A. Lüdeling, & M. Kytö (Eds.), *Corpus Linguistics. An International Handbook. Volume 1* (pp. 14-32). Berlin: Walter de Gruyter.

Kenny, D. (2001). *Lexis and Creativity in Translation. A Corpus-based Study.* Manchester: St. Jerome Publishing.

Kilgarriff, A. (2005). Language is never, ever, ever, random. *Corpus Linguistics and Linguistic Theory, 1*(2), 263-276.

Kjaer, A. L. (2004). A common legal language in Europe? In M. v. Hoecke (Ed.), *Epistemology and Methodology for Comparative Law* (pp. 377-408). Oxford: Hart Publishing.

Klaudy, K. (2008). Explicitation. In M. Baker, & G. Saldanha, *Routledge Encyclopedia of Translation Studies* (pp. 80-85). London: Routledge.

Kruger, H. (2012). A corpus-based study of the mediation effect in translated and edited language. *Target, 24*(2), 355-388.

Krüger, H., & Van Rooy, B. (2012). Register and the features of translated language. *Across Languages and Cultures, 13*(1), 36-65.

Krüger, R. (2013). A Cognitive Linguistic Perspective on Explicitation and Implicitation in Scientific and Technical Translation. *Trans-kom, 6*(2), 285-314.

Krüger, R. (2015). *The Interface between Scientific and Technical Translation Studies and Cognitive Linguistics: With Particular Emphasis on Explicitation*

and Implicitation as Indicators of Translational Text-Context Interaction. Frank & Timme GmbH.

Langacker, R. W. (2000). A Dynamic Usage-Based Model. In M. Barlow, & S. Kemmer (Eds.), *Usage -based models of language* (pp. 1-64). Stanford, CA: CSLI.

Langacker, R. W. (2008). *Cognitive Grammar. A Basic Introduction.* Oxford/New York: Oxford University Press.

Lanstyák, I., & Heltai, P. (2012). Universals in language contact and translation. *Across Languages and Cultures, 13*(1), 99-121.

Laviosa, S. (2002). *Corpus-based Translation Studies. Theory, Findings, Applications.* Amsterdam/New York: Rodopi.

Lyons, J. (1977). *Semantics.* Cambridge: Cabridge University Press.

Malmkjaer, K. (1997). Punctuation in Hans Christian Andersen's stories and in their translations into English. In F. Poyatos (Ed.), *Nonverbal Communication and Translation : New Perspectives and Challenges in Literature, Interpretation and the Media* (pp. 151-162). John Benjamins Publishing Company.

Malmkjaer, K. (1998). Love thy Neighbour : Will Parallel Corpora Endear Linguists toTranslators? *Meta: Translators' Journal, 43*(4), 534-541.

Marinis, T. (2010). Using on-line processing methods in language acquisition research. In E. Blom, & S. Unsworth (Eds.), *Experimental Methods in Language Acquisition Research* (pp. 139-162). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Maslen, R. J., Theakston, A. L., Lieven, E. V., & Tomasello, M. (2004). A Dense Corpus Study of Past Tense and Plural Overregularization in English. *Journal of Speech, Language, and Hearing Research, 47,* 1319-1333.

Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2005). The role of frequency in the acquisition of English word order. *Cognitive Development, 20*(1), 121-136.

Matulewska, A., & Gortych, K. (2009). Translacyjne problemy wyrażania modalności deontycznej w tekstach aktów normatywnych w języku polskim, angielskim i greckim. In A. Mróz, A. Niewiadomski, & M. Pawelec (Eds.), *Prawo i język* (pp. 65-78). Warszawa: Lingua Iuris.

Mauranen, A. (2000). Strange Strings in Translated Language: A Study on Corpora. In M. Olohan (Ed.), *Intercultural Faultlines. Research Methods in Translation Studies I: Textual and Cognitive Aspects* (pp. 105-118). Manchester: St. Jerome.

Mauranen, A. (2008). Universal Tendencies in Translation. In G. M. Anderman, & M. Rogers (Eds.), *Incorporating Corpora: The Linguist and the Translator* (pp. 32-48). Multilingual Matters.

May, R. (1997). Sensible Elocution. *The Translator, 3*(1), 1-20.

McEnery, T., & Wilson, A. (2001). *Corpus Linguistics. An Introduction* (2nd ed.). Edinburgh: Edinburgh University Press.

Merryman, J. H. (1981). On the Convergence (and Divergence) of the Civil Law and the Common Law. *Stanford Journal of International Law, 17,* 357-388.

Munday, J. (2001). *Introducing Translation Studies.* London/New York: Routledge.

Newmark, P. (1988). *A textbook of translation.* Prentice Hall International.

Nord, C. (1997). *Translating as a purposeful activity: functionalist approaches explained.* Manchester: St. Jerome.

NSA. (2010). *Informacja o działalności sądów administracyjnych w 2009 roku.* Retrieved from http://www.nsa.gov.pl/index.php/pol/Media/Files/20092

Nuyts, J. (2006). Modality: overview and linguistic issues. In W. Frawley (Ed.), *The expression of modality* (pp. 1-26). Berlin/New York: De Gruyter Mouton.

Olohan, M. (2001). Spelling out the Optionals in Translation: A Corpus Study. *UCREL Technical Papers, 13*, 423-432.

Olohan, M., & Baker, M. (2000). Reporting that in translated English. Evidence for subconscious processes of explicitation? *Across Languages and Cultures, 1*(2), 141-158.

Øverås, L. (1998). In Search of the Third Code: An Investigation of Norms in Literary Translation. *Meta, 43*(4), 571-588.

Palmer, F. R. (1986). *Mood and modality.* Cambridge: Cambridge University Press.

Palmer, F. R. (2001). *Mood and Modality (2nd ed.).* Cambridge: Cambridge University Press.

Pápai, V. (2004). Explicitation: A Universal of Translated Texts? In A. Mauranen, & P. Kujamäki (Eds.), *Translation Universals. Do They Exist?* (pp. 143-165). Amsterdam/Philadelphia: John Benjamins.

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1-2), 8-13.

Perkins, M. R. (1983). *Modal Expressions in English.* London: Frances Pinter & Norwood.

Pesaran, M. H., & Weeks, M. (1999). *Non-nested Hypothesis Testing: An Overview.* Retrieved from https://www.repository.cam.ac.uk/bitstream/handle/1810/429/nnest.pdf?sequence=1

Portner, P. (2009). *Modality.* Oxford University Press.

Poscher, R. (2015). Ambiguity and Vagueness in Legal Interpretation. In *The Oxford Handbook of Language and Law* (pp. 128-144). Oxford/New Yok: Oxford University Press.

Pullum, G. K. (2007). Ungrammaticality, rarity, and corpus use. *Corpus Linguistics and Linguistic Theory, 3*(1), 33-47.

Pym, A. (2008). On Toury's laws of how translators translate. In A. Pym, M. Shlesinger, & D. Simeoni (Eds.), *Beyond Descriptive Translation Studies* (pp. 311-238). Amsterdam: John Benjamins.

Reber, A. S., Allen, R., & Reber, P. J. (1999). Implicit versus Explicit Learning. In R. J. Sternberg (Ed.), *The Nature of Cognition* (pp. 475-513). Cambridge, Massachusetts: MIT Press.

Roehr-Brackin, K. (2015). Explicit knowledge about language in L2 learning. A usgae-based perspective. In P. Rebuschat (Ed.), *Implicit and Explicit Learning of Languages* (pp. 117-138). Amsetrdam/Philadelphia: John Benjamins.

Rossini, C. (1998). *English as a Legal Language.* London: Kluwer Law International.

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic Entries in the Internal Lexicon. *Journal of Verbal Learning and Verbal Behavior, 9*(5), 487-494.

Saldanha, G. (2008). Explicitation Revisited: Bringing the Reader into the Picture. *Trans-kom, 1*(1), 20-35.

Šarčević, S. (1997). *New Approach to Legal Translation.* Hague/London: Kluwer Law International.

Sarkar, D. (2008). *Lattice: Multivariate Data Visualization with R.* New York: Springer.

Schmid, H.-J. (2000). *English Abstract Nouns as Conceptual Shells. From Corpus to Cognition.* Berlin/New York: Mouton de Gruyter.

Schmid, H.-J. (2010). Does frequency in text instantiate entrenchment in the cognitive system? In D. Glynn, & K. Fischer (Eds.), *Quantitative methods in*

cognitive semantics: corpus-driven approaches (pp. 101-133). Berlin/New York: Walter de Gruyter.

Schmied, J., & Schäffler, H. (1996). Explicitness as a universal feature of translation. In M. Ljung (Ed.), *Corpus-based studies in English: Papers from the seventeenth International Conference on English Language Research on Computerized Corpora (ICAME 17)* (pp. 31-36). Amsterdam/Atlanta: Rodopi.

Schütze, C. T. (2005). Thinking About What We Are Asking Speakers to Do. In S. Kepser, & M. Reis (Eds.), *Linguistic Evidence. Empirical, Theoretical and Computational Perspectives* (pp. 547-484). Berlin/New York: Mouton de Gruyter.

Schütze, C. T., & Sprouse, J. (2013). Judgement data. In R. J. Podesva, & D. Sharma (Eds.), *Research Methods in Linguistics* (pp. 27-50). Cambridge: Cambridge University Press.

Schütze, C. T., & Sprouse, J. (2013). Judgement data. In R. J. Podesva, & D. Sharma (Eds.), *Research Methods in Linguistics* (pp. 27-50). Cambridge: Cambridge University Press.

Sedlmeier, P., Betsch, T., & Renkewitz, F. (2002). Frequency processing and cognition: introduction and overview. In P. Sedlmeier, & T. Betsch (Eds.), *Etc. Frequency Processing and Cognition* (pp. 1-17). Oxford, New York: Oxford University Press.

Shlesinger, M. (1991). Interpreter Latitude vs. Due Process. Simultaneous and Consecutive Interpretation in Multilingual Trials. In S. Tirkkonen-Condit (Ed.), *Empirical Research in Translation and Intercultural Studies* (pp. 147-155). Tubingen: Gunter Narr.

Snider, N., & Arnon, I. (2012). A unified lexicon and grammar? Compositional and non-compositional phrases in the lexicon. In D. Divjak, & S. T. Gries (Eds.), *Frequency effects in language representation* (pp. 127-164). Berlin/Boston: De Gruyter Mouton.

Speelman, D. (2014). Logistic regression. A confirmatory technique for comparisons in corpus ilnguistics. In D. Glynn, & J. A. Robinson (Eds.), *Corpus Methods for Semantics. Quantitative studies in polysemy and synonymy* (pp. 487-533). Amsterdam/Philadelphia: John Benjamins Publishing.

Stefanowitsch, A., & Gries, S. T. (2003). Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics, 8*(2), 209-243.

Sunderman, G., & Kroll, J. F. (2006). First Language Activation during Second Language Processing. An Investigation of Lexical Form, Meaning, and Grammatical Class. *Studies in Second Language Acquisition, 28*(3), 387-422.

Sweetser, E. E. (1982). Root and Epistemic Modals: Causality in Two Worlds. In *Proceedings of the Eighth Annual Meeting of the Berkeley Linguistics* (pp. 484-507). Berkley Linguistic Society.

Szymor, N. (2011). *Quality Control in Legal Translation. Translation of EU Legislation into Polish.* University of Sheffield: Unpublished MA dissertation.

Szymor, N. (2015). Behavioral Profiling in Translation Studies. *trans-kom, 8*(2), 483-498.

Teddiman, L. (2012). Conversion and the lexicon: Comparing evidence from corpora and experimentation. In D. Divjak, & S. T. Gries (Eds.), *Frequency Effects in Language Representation* (pp. 235-254). Berlin/Boston: De Gruyter Mouton.

Teich, E. (2003). *Cross-linguistic Variation in System and Text: A Methodology for the Investigation of Translations and Comparable Texts.* Berlin/New York: Walter de Gruyter.

Tetley, W. (2003). Nationalism in a Mixed Jurisdiction and the Importance of Language. *Tulane Law Review, 78*, 175-218.

Theakston, A. L., Lieven, E. V., Pine, J. M., & Rowland, C. F. (2004). Semantic generality, input frequency and the acquisition of syntax. *Journal of Child Language, 31*(1), 61-99.

Thráinsson, H., Angantýsson, Á., Svavarsdóttir, Á., Eythórsson, T., & Jónsson, J. G. (2007). The Icelandic (Pilot) Project in ScanDiaSyn. *Nordlyd : Tromsø University working papers on language & linguistics, 34*(1), 87-124.

Toury, G. (1985). A Rationale for Descriptive Translation Studies. In T. Hermans (Ed.), *The Manipulation of Literature: Studies in Literary Translation* (pp. 16-41). London/Sydney: Croom Helm.

Toury, G. (1991). Experimentation in Translation Studies: Achievements, Prospects and Some Pitfals. In S. Tirkkonen-Condit, *Empirical Research in Translation and Intercutlural Studies* (pp. 45-66). Tubingen: Gunter Narr Verlag.

Toury, G. (1995). *Descriptive Translation Studies and beyond.* Amsterdam/Philadelphia: John Benjamins.

Ulrych, M., & Murphy, A. (2008). Descriptive Translation Studies and the Use of Corpora: Investigating Mediation Universals. In C. T. Torsello, K. Ackerley,

& E. Castello, *Corpora for University Language Teachers* (pp. 141-166). Bern: Peter Lang.

van der Auwera, J., & Plungian, V. A. (1998). Modality's Semantic Map. *Linguistic Typology, 2*(1), 79-124.

Vanderauwera, R. (1985). *Dutch Novels Translated into English: The Transformation of a "Minority" Literature.* Amsterdam: Rodopi.

Vandevoorde, L. (2016). *On semantic differences. A multivariate corpus-based study of the semantic field of inchoativity in translated and non-translated Dutch.* Ghent: Unpublished.

Venables, W., & Ripley, B. (2002). *Modern Applied Statistics with S* (4th ed.). New York: Springer.

Von Fintel, K. (2006). Modality and language. In D. M. Borchert (Ed.), *Encyclopedia of philosophy, 2nd Ed.* (pp. 20-27). Detroit: MacMillan Reference US.

Wasow, T., & Arnold, J. (2005). Intuitions in linguistic argumentation. *Lingua, 115,* 1481-1496.

Wiemer, B. (2001). Aspect choice in non-declarative and modalized utterances as extensions from assertive domains (Lexical semantics, scopes, and categorial distinctions in Russian and Polish). In H. Bartels, N. Störmer, & E. Walusiak (Eds.), *Untersuchungen zur Morphologie und Syntax im Slavischen* (pp. 195-221). Oldenburg: BIS-Verlag.

Williams, D. A. (2005). *Recurrent Features of Translation in Canada: A Corpus-based Study.*

Winter, B. (2013). *Linear models and linear mixed effects models in R with linguistic applications.* Retrieved from https://arxiv.org/ftp/arxiv/papers/1308/1308.5499.pdf

Appendix 1. Consent forms[26]

1. Consent form for professional translators (in Polish and English)

Zgoda na udział w badaniu

Proszę o uważne przeczytanie poniższych informacji. Mogą Państwo poprosić o kopię formularza do własnego użytku.

Badanie:	Przetwarzanie informacji w procesie tłumaczenia z języka angielskiego na język polski
Organizator badania:	Nina Szymor
Instytucja prowadząca:	Uniwersytet w Sheffield (The University of Sheffield)

OPIS BADANIA: Czy kiedykolwiek zastanawiali się Państwo nad tym, jakie procesy zachodzą w umyśle i w jaki sposób przetwarzamy informacje w procesie tłumaczenia z języka obcego na język ojczysty? W jaki sposób wybieramy słowa, których chcemy użyć do przetłumaczenia na własny język znaczenia wyrażonego w języku obcym? Czy procesy i wybory zachodzące w trakcie tłumaczenia są podobne

[26] Initially, the three tasks were: judgement task, self-paced reading, and proofreading. Due to a glitch in the keystroke logging software, the proofreading task had to be replaced by the forced-choice task. The consent to the forced-choice task was given online.

do tych, które mają miejsce gdy piszemy w języku ojczystym? Udzielone przez Państwa odpowiedzi pomogą nam w lepszym zrozumieniu tych zjawisk.

KONSTRUKCJA BADANIA: Badanie składa się z trzech części i potrwa około półtorej godziny. W pierwszej części badania poproszę Państwa o ocenienie poprawności wypowiedzi rodzimych użytkowników języka polskiego (na skali od 1 do 100). Ta część odbędzie się online i mogą ją Państwo wykonać we własnym czasie. W drugiej części, poproszę Państwa o przeczytanie wypowiedzi podzielonych na fragmenty. Będą Państwo mogli samodzielnie kontrolować tempto pojawiania się kolejnego fragmentu zdania (ang. *self-paced reading*). W ostatniej części zostaną Państwo poproszeni o koretkę zdań przetłumaczonych z języka angielskiego na język polski. W tej części, wszystkie czynności wykonane przez Państwa klawiaturą i myszką będą monitorowane i zapisywane przy pomocy specjalnego oprogramowania (ang. *key stroke logger*). Aby wykonać część drugą i trzecią, konieczne będzie spotkanie.

CZAS TRWANIA: Około półtorej godziny.

PRAWA UCZESTNIKA: Jeżeli zgadzają się Państwo na udział w badaniu, proszę pamiętać, iż mogą Państwo je przerwać zarówno na początku, jak również w dowolnym momencie podczas jego trwania, bez podawania przyczyny. Jeśli nie chcą Państwo odpowiadać na którekolwiek z pytań, mogą Państwo je pominąć bez podawania przyczyny. Wszelkie dane podane przez Państwa (np. wykształcenie) użyte zostaną wyłącznie do celów statystycznych, a wyniki ankiety nie zostaną powiązane z Państwa imieniem i nazwiskiem. Dane uzyskane w trakcie trwania eksperymentu przechowywane będą na chronionych hasłem komputerach Uniwersytetu w Sheffield.

Uczestnicy badania otrzymają symboliczne wynagrodzenie pieniężne w postaci karty podarunkowej o wartości 20GBP (lub 120PLN), pod warunkiem wykonania wszystkich trzech części badania i przekazania swoich prawidłowych danych osobowych, przy czym przekazanie danych osobowych jest dobrowolne.

Jeżeli akceptują Państwo powyższe warunki oraz zgadzają się na udział w badaniu, proszę złożyć
swój podpis poniżej:

Imię i nazwisko:

_____

Data: _____ Podpis:_____

Informed consent form for Experimental Participants

Please read the following information carefully. You can also request a copy for future reference.

Experiment: Processing of information in the process of translation from English into Polish

Experimenter: Nina Szymor

Institution: The University of Sheffield

DESCRIPTION: Have you ever wondered what process take place in our minds and how we process information when we translate from a foreign language into our mother tongue? How do we choose words that we want to use to translate the meaning expressed in a foreign language into our mother tongue? Are the choices and processed taking place while we translate similar to those that take place when we write originally in our mother tongue? Your answers will help us to better understand these processes and phenomena.

EXPERIMENT: The test consists of three tasks and will take approximately 1.5 hours. In task one, you will be asked to judge the acceptability of sentences on a scale from 1 to 100. This part is done online so you can complete it in your own time. In task two, you will be asked to read fragments of utterances. You will be able to control when the next part of the utterance appears on the screen (*self-paced reading*). In task 3, you will be asked to proofread sentences translated from English

into Polish. All of your keyboard and mouse movements will be monitored and recorded by special software (*key stroke logger*). In order to complete tasks 2 and 3, a meeting in person will be necessary.

TIME INVOLVEMENT: Approx. 1.5 hours.

SUBJECT'S RIGHTS: If you have read this form and have decided to participate in this experiment, please understand your participation is voluntary and you have the right to withdraw your consent or discontinue participation at any time without stating the reason. You have the right to refuse to answer particular questions. All data obtained from this questionnaire (e.g. your educational attainments) will be used for statistical purposes only and the results of the study will not be linked to your name. The data obtained as a result of this questionnaire will be stored on password-protected computers at the University of Sheffield.

Participants will be given a small compensation in the form of a gift card for the amount of 20GBP (or 120PLN), provided they completed all three tasks and provided a valid email address.

If you agree with the above-stated conditions and are willing to participate in the experiment, please sign below.

Name:

_____

Date: _____ Signature:_____

2. Consent form for trainee translators (in Polish and English)

Zgoda na udział w badaniu

Proszę o uważne przeczytanie poniższych informacji. Mogą Państwo poprosić o kopię formularza do własnego użytku.

Badanie:	Przetwarzanie informacji w procesie tłumaczenia z języka angielskiego na język polski
Organizator badania:	Nina Szymor
Instytucja prowadząca:	Uniwersytet w Sheffield (The University of Sheffield)

OPIS BADANIA: Czy kiedykolwiek zastanawiali się Państwo nad tym, jakie procesy zachodzą w umyśle i w jaki sposób przetwarzamy informacje w procesie tłumaczenia z języka obcego na język ojczysty? W jaki sposób wybieramy słowa, których chcemy użyć do przetłumaczenia na własny język znaczenia wyrażonego w języku obcym? Czy procesy i wybory zachodzące w trakcie tłumaczenia są podobne do tych, które mają miejsce gdy piszemy w języku ojczystym? Udzielone przez Państwa odpowiedzi pomogą nam w lepszym zrozumieniu tych zjawisk.

KONSTRUKCJA BADANIA: Badanie składa się z trzech części i potrwa około półtorej godziny. W pierwszej części badania poproszę Państwa o ocenienie poprawności wypowiedzi rodzimych użytkowników języka polskiego (na skali od 1 do 100). Ta część odbędzie się online i mogą ją Państwo wykonać we własnym czasie. W drugiej części, poproszę Państwa o przeczytanie wypowiedzi podzielonych na fragmenty. Będą Państwo mogli samodzielnie kontrolować tempto pojawiania się kolejnego fragmentu zdania (ang. *self-paced reading*). W ostatniej części zostaną Państwo poproszeni o koretkę zdań przetłumaczonych z języka angielskiego na język polski. W tej części, wszystkie czynności wykonane przez Państwa klawiaturą i myszką będą monitorowane i zapisywane przy pomocy specjalnego oprogramowania (ang. *key stroke logger*). Aby wykonać część drugą i trzecią, konieczne będzie spotkanie.

CZAS TRWANIA: Około półtorej godziny.

PRAWA UCZESTNIKA: Jeżeli zgadzają się Państwo na udział w badaniu, proszę pamiętać, iż mogą Państwo je przerwać zarówno na początku, jak również w dowolnym momencie podczas jego trwania, bez podawania przyczyny. Jeśli nie chcą Państwo odpowiadać na którekolwiek z pytań, mogą Państwo je pominąć bez podawania przyczyny. Wszelkie dane podane przez Państwa (np. wykształcenie) użyte zostaną wyłącznie do celów statystycznych, a wyniki ankiety nie zostaną

powiązane z Państwa imieniem i nazwiskiem. Dane uzyskane w trakcie trwania eksperymentu przechowywane będą na chronionych hasłem komputerach Uniwersytetu w Sheffield.

Uczestnicy badania otrzymają symboliczne wynagrodzenie pieniężne w postaci karty podarunkowej o wartości 7.50GBP (lub 45PLN), pod warunkiem wykonania wszystkich trzech części badania i przekazania swoich prawidłowych danych osobowych, przy czym przekazanie danych osobowych jest dobrowolne.

Jeżeli akceptują Państwo powyższe warunki oraz zgadzają się na udział w badaniu, proszę złożyć
swój podpis poniżej:

Imię i nazwisko:
_____

Data: _____ Podpis:_____

3. Consent form for naive native speakers (in Polish and English)

Zgoda na udział w badaniu

Proszę o uważne przeczytanie poniższych informacji. Mogą Państwo poprosić o kopię formularza do własnego użytku.

Badanie:	Przetwarzanie informacji w procesie tłumaczenia z języka angielskiego na język polski
Organizator badania:	Nina Szymor
Instytucja prowadząca:	Uniwersytet w Sheffield (The University of Sheffield)

OPIS BADANIA: Czy kiedykolwiek zastanawiali się Państwo nad tym, jakie procesy zachodzą w umyśle i w jaki sposób przetwarzamy informacje w procesie tłumaczenia z języka obcego na język ojczysty? W jaki sposób wybieramy słowa,

których chcemy użyć do przetłumaczenia na własny język znaczenia wyrażonego w języku obcym? Czy procesy i wybory zachodzące w trakcie tłumaczenia są podobne do tych, które mają miejsce gdy piszemy w języku ojczystym? Udzielone przez Państwa odpowiedzi pomogą nam w lepszym zrozumieniu tych zjawisk.

KONSTRUKCJA BADANIA: Badanie składa się z dwóch części i potrwa około pół godziny. W pierwszej części badania poproszę Państwa o ocenienie poprawności wypowiedzi rodzimych użytkowników języka polskiego (na skali od 1 do 100). Ta część odbędzie się online i mogą ją Państwo wykonać we własnym czasie. W drugiej części, poproszę Państwa o przeczytanie wypowiedzi podzielonych na fragmenty. Będą Państwo mogli samodzielnie kontrolować tempto pojawiania się kolejnego fragmentu zdania (ang. *self-paced reading*). Aby wykonać część drugą część, konieczne będzie spotkanie.

CZAS TRWANIA: Około pół godziny.

PRAWA UCZESTNIKA: Jeżeli zgadzają się Państwo na udział w badaniu, proszę pamiętać, iż mogą Państwo je przerwać zarówno na początku, jak również w dowolnym momencie podczas jego trwania, bez podawania przyczyny. Jeśli nie chcą Państwo odpowiadać na którekolwiek z pytań, mogą Państwo je pominąć bez podawania przyczyny. Wszelkie dane podane przez Państwa (np. wykształcenie) użyte zostaną wyłącznie do celów statystycznych, a wyniki ankiety nie zostaną powiązane z Państwa imieniem i nazwiskiem. Dane uzyskane w trakcie trwania eksperymentu przechowywane będą na chronionych hasłem komputerach Uniwersytetu w Sheffield.

Uczestnicy badania otrzymają symboliczne wynagrodzenie pieniężne w postaci karty podarunkowej o wartości 5GBP (lub 30PLN), pod warunkiem wykonania wszystkich trzech części badania i przekazania swoich prawidłowych danych osobowych, przy czym przekazanie danych osobowych jest dobrowolne.

Jeżeli akceptują Państwo powyższe warunki oraz zgadzają się na udział w badaniu, proszę złożyć
swój podpis poniżej:

Imię i nazwisko: _____

Data: _____ Podpis:_____

Informed consent form for Experimental Participants

Please read the following information carefully. You can also request a copy for future reference.

Experiment: Processing of information in the process of translation from English into Polish

Experimenter: Nina Szymor

Institution: The University of Sheffield

DESCRIPTION: Have you ever wondered what process take place in our minds and how we process information when we translate from a foreign language into our mother tongue? How do we choose words that we want to use to translate the meaning expressed in a foreign language into our mother tongue? Are the choices and processed taking place while we translate similar to those that take place when we write originally in our mother tongue? Your answers will help us to better understand these processes and phenomena.

EXPERIMENT: The test consists of two tasks and will take approximately 30 minutes. In task one, you will be asked to judge the acceptability of sentences on a scale from 1 to 100. This part is done online so you can complete it in your own time. In task two, you will be asked to read fragments of utterances. You will be able to control when the next part of the utterance appears on the screen (*self-paced reading*). In order to complete task 2 , a meeting in person will be necessary.

TIME INVOLVEMENT: Approx. 30 minutes.

SUBJECT'S RIGHTS: If you have read this form and have decided to participate in this experiment, please understand your participation is voluntary and you have the right to withdraw your consent or discontinue participation at any time without stating the reason. You have the right to refuse to answer particular questions. All data obtained from this questionnaire (e.g. your educational attainments) will be used for statistical purposes only and the results of the study will not be linked to your name. The data obtained as a result of this questionnaire will be stored on password-protected computers at the University of Sheffield.

Participants will be given a small compensation in the form of a gift card for the amount of 5GBP (or 30PLN), provided they completed all three tasks and provided a valid email address.

If you agree with the above-stated conditions and are willing to participate in the experiment, please sign below.

Name:

_____

Date: _____ Signature:_____

Informed consent form for Experimental Participants

Please read the following information carefully. You can also request a copy for future reference.

Experiment: Processing of information in the process of translation from English into Polish

Experimenter: Nina Szymor

Institution: The University of Sheffield

DESCRIPTION: Have you ever wondered what process take place in our minds and how we process information when we translate from a foreign language into our

mother tongue? How do we choose words that we want to use to translate the meaning expressed in a foreign language into our mother tongue? Are the choices and processed taking place while we translate similar to those that take place when we write originally in our mother tongue? Your answers will help us to better understand these processes and phenomena.

EXPERIMENT: The test consists of three tasks and will take approximately 1.5 hours. In task one, you will be asked to judge the acceptability of sentences on a scale from 1 to 100. This part is done online so you can complete it in your own time. In task two, you will be asked to read fragments of utterances. You will be able to control when the next part of the utterance appears on the screen (*self-paced reading*). In task 3, you will be asked to proofread sentences translated from English into Polish. All of your keyboard and mouse movements will be monitored and recorded by special software (*key stroke logger*). In order to complete tasks 2 and 3, a meeting in person will be necessary.

TIME INVOLVEMENT: Approx. 1.5 hours.

SUBJECT'S RIGHTS: If you have read this form and have decided to participate in this experiment, please understand your participation is voluntary and you have the right to withdraw your consent or discontinue participation at any time without stating the reason. You have the right to refuse to answer particular questions. All data obtained from this questionnaire (e.g. your educational attainments) will be used for statistical purposes only and the results of the study will not be linked to your name. The data obtained as a result of this questionnaire will be stored on password-protected computers at the University of Sheffield.

Participants will be given a small compensation in the form of a gift card for the amount of 7.50GBP (or 45PLN), provided they completed all three tasks and provided a valid email address.

If you agree with the above-stated conditions and are willing to participate in the experiment, please sign below.

Name:

_____

Date: _____ Signature:_____

Appendix 2. Stimuli and experimental sets

1. Stimuli

1.1. Judgement task and self-paced reading task

TABLE 1: EXPERIMENTAL STIMULI					
ID	**general chunk pref**	**legal chunk pref**	**general unigram pref**	**legal unigram pref**	**sentence_PL**
E1	impf	impf	pf	pf	Członkostwa w Radzie nie można łączyć z członkostwem w Państwowej Komisji Akredytacyjnej.
E2	impf	impf	pf	impf	Strefy pożarowe w podziemnej części budynków nie powinny przekraczać 50% powierzchni określonych w tabeli.
E3	impf	impf	pf	pf	Młodocianego nie wolno zatrudniać w godzinach nadliczbowych ani w porze nocnej.
E4	impf	impf	pf	impf	Należy określić zakres uprawnień zainteresowanych osób oraz to, czy mogą wykonywać je samodzielnie.
E5	impf	impf	pf	pf	Programu pomocy nie można łączyć z innymi programami, jest on skierowany do przedsiębiorstw pozostałych na liście.
E6	impf	impf	pf	impf	Przegląd będzie dotyczył roli, jaką Stały Komitet ds. Leśnictw powinien spełniać we wdrażaniu planu działania.
E7	impf	impf	pf	pf	Wspólnota Europejska powinna określić wspólne zasady oraz ogólne ramy dla działań.
E8	pf	pf	impf	impf	Potwierdzenie dostawy może nastąpić poprzez przedłożenie innych dokumentów wskazujących, że została ona zrealizowana.
E9	pf	pf	impf	pf	Dokument potwierdzający kwalifikacje można wydać tylko funkcjonariuszowi uprawnionemu do jego uzyskania.

E10	pf	pf	impf	pf	W opisie statku morskiego należy wskazać rok i miejsce budowy z oznaczeniem stoczni.
E11	pf	pf	impf	impf	Użycie broni palnej powinno następować w sposób wyrządzający możliwie najmniejszą szkodę ofierze.
E12	pf	pf	impf	pf	Każdy wierzyciel może wnieść swoje roszczenie w języku urzędowym Państwa Członkowskiego.
ID	**general chunk pref**	**legal chunk pref**	**general unigram pref**	**legal unigram pref**	**sentence_PL**
E13	pf	pf	impf	pf	Opinia zawiera wymogi służące do określenia, czy danemu łowisku można nadać oznaczenie ekologiczne.
E14	pf	pf	impf	pf	Zasady te określają informacje, które należy wskazać we wnioskach o pozwolenie i w pozwoleniach.
E15	pf	pf	impf	impf	Utworzenie Europejskiego Instytutu ds. Równości Płci powinno następować neutralnie dla budżetu.
E16	pf	pf	impf	pf	Dowód pochodzenia musi być wystawiony zgodnie z przepisami i należy wskazać w nim nazwę i adres danej wystawy.
E17	pf	impf	impf	impf	Wielkość ogłoszenia nie może wynosić więcej niż 650 wyrazów.
E18	pf	impf	impf	impf	Całkowita kwota pomocy wspólnotowej może wynieść 20 % całkowitego kosztu inwestycji.
E19	impf	pf	impf	pf	Naczelnik urzędu celnego może wydać, na uzasadniony wniosek, większą liczbę egzemplarzy zwolnienia.
E20	impf	pf	impf	pf	Państwa Członkowskie nie mogą wydawać dokumentu określonego w dyrektywie 74/150/EWG art. 10 w odniesieniu do typu ciągnika.
E21	pf	impf	pf	pf	Rozporządzenie dotyczy jednostek badawczych, które mogą nabyć mienie niezbędne do prowadzenia badań naukowych.
E22	pf	impf	pf	impf	Odszkodowania takiego można dochodzić od osoby fizycznej lub prawnej, lub od każdego innego podmiotu wyrządzającego szkody.
E23	pf	impf	pf	impf	W skład oddziału powinny wchodzić magazyny depozytów przedmiotów osobistych pacjenta.
E24	pf	impf	pf	impf	Kwestia odrzucenia wniosku Słowenii

					może wchodzić w rachubę w myśl zasady "pacta sunt servanda".

TABLE 2: FILLER STIMULI	
ID	**sentence_PL**
F1	Rada Ministrów może rozstrzygać poszczególne sprawy w drodze korespondencyjnego uzgodnienia stanowisk.
F2	Funkcję członka Rady można pełnić tylko jedną kadencję.
F3	Osoba ubiegająca się o licencję musi spełniać następujące wymagania w zakresie wieku i wykształcenia.
F4	W księdze należy ewidencjować wyłącznie przychody i koszty z działalności gospodarczej.
F5	W pomieszczeniu, w którym przebywają krowy, nie wolno utrzymywać drobiu i trzody chlewnej.
F6	Wniosek o określenie warunków przyłączenia może zawierać wymagania dotyczące odmiennych parametrów technicncyh.
F7	Można stosować jednocześnie różne środki przymusu bezpośredniego, jeżeli jest to konieczne do osiągnięcia podporządkowania się wydanym poleceniom.
F8	Wykaz podpisów musi zawierać na każdej stronie nazwę komitetu wyborczego zgłaszającego listę.
F9	Sprawozdania należy sporządzać w sposób czytelny i trwały.
F10	Powiadomienie, o którym mowa w ust. 2 i 3, powinno zawierać nazwę i siedzibę kościelnej osoby prawnej.
F11	Wolno zamieszczać rozpowszechnione drobne utwory lub fragmenty większych utworów w podręcznikach.
F12	Jeżeli nie można stwierdzić, czy osoba przewozi towary niezgłoszone, organ celny może przeszukać tę osobę.
F13	Rada Ministrów określi warunki, jakie musi spełnić towarzystwo w celu zapewnienia bezpieczeństwa działalności.
F14	W rozporządzeniu należy uwzględnić potrzeby szczegółowej ewidencji spraw prowadzoncyh w kancelarii.
F15	Instruktaż powinien zapoznać uczestników szkolenia z podstawowymi przepisami bezpieczeństwa i higieny pracy.
F16	Nie wolno zatrudniać młodocianych przy pracach wzbronionych, których wykaz ustala Rada Ministrów.
F17	Student może otrzymać jedną pożyczkę studencką albo jeden kredyt studencki.
F18	Policjantowi można przyznać nagrodę za wzorowe wykonywanie zadań służbowych i przejawianą inicjatywę w służbie.
F19	Kandydat musi wykazać się wiedzą z zakresu zadań Państwowej Inspekcji Sanitarnej.
F20	W treści środka zaskarżenia należy podać, czego domaga się skarżący.
F21	Przed przystąpieniem do robót użytkownik powinien w odpowiednim terminie zawiadomić właściciela o swym zamiarze.

F22	Nie wolno dopuścić pracownika do pracy, do której wykonywania nie posiada on wymaganych kwalifikacji.
F23	Rada Ministrów może określać rezerwy gospodarcze, zobowiązując do ich tworzenia innych ministrów.
F24	Utrwalenia obrazu i dźwięku można dokonywać na środkach technicznych przeznaczonych do utrwalania obrazu i dźwięku.
F25	Na statku musi znajdować się ekspert do spraw bezpieczeństwa przewozu materiałów niebezpiecznych statkami żeglugi śródlądowej.
ID	**sentence_PL**
F26	Żądanie odszkodowania należy zgłosić w terminie przewidzianym w art. 555.
F27	Rozporządzenie powinno ustalać sposób przydziału lokali mieszkalnych i kwater tymczasowych.
F28	Organizacje mogą głosować na dwóch kandydatów na przedstawicieli partnerów społecznych i gospodarczych.
F29	Uchwała senatu uczelni określa plan studiów i programy nauczania, które muszą wynikać ze standardów kształcenia.
F30	Jeżeli zbycie uzależnione jest od zezwolenia spółki, należy stosować poniższe przepisy.
F31	Kapitał zakładowy spółki powinien wynosić co najmniej 50.000 złotych.
F32	Do czasu wydania uchwały zezwalającej na pociągnięcie sędziego do odpowiedzialności karnej wolno podejmować tylko czynności niecierpiące zwłoki.
F33	Komendant Główny Straży Granicznej może przedłużyć funkcjonariuszom uprawnienia do normy LOT na czas nie dłuższy niż 6 miesięcy.
F34	Osobie, o której mowa w ust. 1, można przyznać prawo wykonywania zawodu jeżeli ukończyła studia magisterskie.
F35	Koordynacja działań pomiędzy strefami musi zapewnić właściwą ocenę poziomu ozonu.
F36	W dokumencie gwarancyjnym należy zamieścić podstawowe dane potrzebne do dochodzenia roszczeń z gwarancji.
F37	Do wniosku operator powinien dołączyć plan inwestycyjny oraz oświadczenie zawierające zobowiązanie.
F38	Wydział pracy i ubezpieczeń społecznych można utworzyć w sądzie rejonowym mającym siedzibę w mieście na prawach powiatu.
F39	Państwa Członkowskie mogą wprowadzić uproszczenie przepisów dotyczących urządzeń zamykających.
F40	W wyborach do Sejmu i do Senatu głosować można tylko osobiście i tylko jeden raz.
F41	Pomieszczenie do przechowywania mleka musi być zabezpieczone przed dostępem szkodników, zwłaszcza gryzoni i owadów.
F42	Przewody powinny być wyposażone w otwory rewizyjne umożliwiające oczyszczenie wnętrza tych przewodów.
F43	Sędziemu nie wolno podejmować innego zajęcia, które przeszkadzałoby mu w pełnieniu obowiązków sędziego.
F44	Zwolnienia wolno odmówić tylko wtedy, gdyby złożenie zeznania wyrządzić mogło poważną szkodę państwu.
F45	Cudzoziemcowi nie wolno opuścić wyznaczonego miejsca zamieszkania bez zgody organu, który wydał decyzję.
F46	Zainteresowane strony mogą przesyłać swoje uwagi w terminie do jednego

	miesiąca od daty niniejszego streszczenia.
F47	Oznakowanie ekologiczne można przyznawać produktom, które mogą przyczyniać się do poprawy aspektów środowiskowych.
F48	Zezwolenia muszą być udzielane na podstawie oceny wywoływanych przez nie skutków w zakresie zdrowia ludzi i zwierząt.
F49	Komitet Doradczy doszedł do wniosku, że należy zaliczać czynnik wywołujący BSE do czynników chorobotwórczych.
ID	**sentence_PL**
F50	Ekosystemy powinny być chronione przed szkodliwymi skutkami oddziaływania dwutlenku siarki.
F51	Statkom rybackim nie wolno łowić, przetrzymywać, przeładowywać ani wyładowywać ryb, jeśli nie przyznano im licencji połowowej.
F52	Każde Państwo Członkowskie może ograniczyć ilości paliwa, które mogą być bezcłowo dopuszczone do obrotu.
F53	Przepisy ust. 1 stosuje się, gdy można dowieść naukowo, że gatunki podatne w określonych stadiach rozwoju nie przenoszą danej choroby.
F54	Każdy sprzeciw musi być dostarczony na piśmie do zgłaszającego i innych zainteresowanych organów właściwych w ciągu 30 dni.
F55	Środki niezbędne w celu wykonania niniejszej Umowy należy ustanowić zgodnie z procedurą ustanowioną w art. 30 dyrektywy 72/462/EWG.
F56	Maksymalne oprocentowanie powinno być wybrane rozsądnie zgodnie z wszelkimi odpowiednimi przepisami krajowymi.
F57	Gdy pracownikowi nie wolno ujawnić żądanej informacji ze względu na jej poufny charakter, może on wskazać powody, dla których nie może przekazać informacji.
F58	Wnioski o wydanie pozwoleń mogą być składane w ciągu pierwszych pięciu dni września 2006 r. na 2514,507 tony.
F59	Dyrektywa 70/524/EWG przewiduje, że można zezwolić na nowe dodatki uwzględniając postęp w dziedzinie wiedzy naukowo-technicznej.

TABLE 3: PRACTICE STIMULI	
ID	**sentence_PL**
T1	W przetargu mogą brać udział osoby, które wniosą wadium w terminie wyznaczonym w ogłoszeniu o przetargu.
T2	Zwrotu od Zakładu nienależnie opłaconych składek nie można dochodzić, jeżeli od daty ich opłacenia upłynęło 5 lat.
T3	Wykonujący serwis musi niezwłocznie i bez opłaty interweniować na żądanie właściwego organu skarbowego.
T4	Przez 'temat' należy rozumieć działanie, w rozumieniu ustawy z dnia 20 kwietnia 2004.
T5	Pracodawca powinien zapewnić pracownikom instrukcje dotyczące stosowanych w zakładzie pracy znaków i sygnałów.
T6	W razie przekazania sprawy do ponownego rozpoznania wolno w dalszym postępowaniu wydać orzeczenie surowsze niż uchylone.
T7	Główny Geodeta Kraju może przeprowadzać kontrole w zakresie określonym dla organów, o których mowa w ust. 1 pkt 2-4.
T8	Osoba wykwalifikowana musi zaświadczyć, że każda wytworzona seria

	spełnia wymagania określone w ust. 1.
T9	Państwa Członkowskie zapewniają, że produkty mogą być wywożone ze Wspólnoty tylko gdy są zgodne z niniejszą dyrektywą.
T10	Wnioski można składać w ciągu 13 tygodni od daty opublikowania niniejszego zawiadomienia o zaproszeniu.
T11	Dokument zawierający bilans musi podawać dane osób, które z mocy prawa są wyznaczone do poświadczania takiego dokumentu.
ID	**sentence_PL**
T12	W sytuacjach kiedy skrobia występuje w ilościach śladowych należy stosować mikroskopową metodę jakościową.
T13	Urządzenie do rozruchu zimnego powinno być zaprojektowane w taki sposób, by podczas normalnej eksploatacji silnika nie mogło się uruchomić.
T14	Producentom nie wolno pozwolić na wykorzystywanie obaw przed chorobami jako argumentu przemawiającego za dokonaniem zakupu ich produktu.
T15	Komisja może ustanowić szczególne warunki w celu zastosowania odstępstw przewidzianych w ust. 1.
T16	Komisja będzie musiała rozstrzygnąć, czy konieczne jest wprowadzenie zmian do rozporządzenia finansowego.
T17	Żołnierze zawodowi mogą być zwolnieni z zawodowej służby wojskowej bez zachowania okresu wypowiedzenia.
T18	Konieczne jest ustanowienie warunków, na podstawie których można nakładać środki tymczasowe
T19	Decyzje określone w art. 6 ust. 1 muszą być podjęte najdalej w ciągu jednego miesiąca.
T20	Należy wziąć pod uwagę rozmiary drogowego transportu pasażerskiego, który w ostatnich latach wzrastał w postępie geometrycznym.
T21	Przebudowa obiektu budowlanego powinna uwzględniać poprawę warunków bezpieczeństwa i higieny pracy.
T22	Pracownicom w ciąży i pracownicom karmiącym piersią nie wolno wykonywać czynności, które zostały ocenione jako niosące ryzyko.
T23	Zawód rzecznika patentowego może wykonywać osoba, która spełnia wymagania określone niniejszą ustawą.
T24	W wyżej wymienionych dokumentach zawarta musi być wzmianka o istnieniu prospektu oraz miejscu jego publikacji.

1.2. Forced-choice task

TABLE 4: EXPERIMENTAL STIMULI					
ID	**general chunk pref**	**legal chunk pref**	**general unigram pref**	**legal unigram pref**	**sentence_PL**
E1	impf	impf	pf	impf	Ośrodek może wykonywać inne zadania z zakresu bezpieczeństwa ruchu drogowego.
E2	impf	impf	pf	impf	Lekarz może wykonać czynności, o których mowa w ust. 1, bez zgody przedstawiciela pacjenta bądź zgody właściwego sądu opiekuńczego, gdy zwłoka groziłaby pacjentowi niebezpieczeństwem utraty

					życia.
E3	impf	impf	pf	pf	W uzasadnionych przypadkach można łączyć karę wyznaczenia na niższe stanowisko służbowe oraz karę wydalenia ze służby z karą obniżenia stopnia.
ID	**general chunk pref**	**legal chunk pref**	**general unigram pref**	**legal unigram pref**	**sentence_PL**
E4	impf	impf	pf	pf	Ogłoszenie oraz zawiadomienie o zamiarze sprzedaży można połączyć z ogłoszeniem i zawiadomieniem, o którym mowa w art. 27 ust. 1 i 2.
E5	impf	impf	pf	pf	Rozporządzenie powinno określać sposób udokumentowania przez funkcjonariusza celnego poniesionych kosztów przeniesienia, w tym podróży.
E6	impf	impf	pf	pf	Rozporządzenie powinno określić nazwy komisji, ich siedziby oraz strukturę organizacyjną.
E7	pf	pf	impf	impf	Oświadczenie pracownika o rozwiązaniu umowy o pracę bez wypowiedzenia powinno nastąpić na piśmie, z podaniem przyczyny uzasadniającej rozwiązanie umowy.
E8	pf	pf	impf	impf	Korzystanie przez skazanego z przysługujących mu praw powinno następować w sposób nie naruszający praw innych osób oraz nie zakłócający ustalonego w zakładzie karnym porządku.
E9	pf	pf	impf	pf	Wniosek o wznowienie postępowania dyscyplinarnego mogą wnieść osoba ukarana lub rzecznik dyscyplinarny do spraw studentów w terminie trzydziestu dni od dnia powzięcia wiadomości o przyczynie uzasadniającej wznowienie.
E10	pf	pf	impf	pf	Skarb Państwa może wnosić do funduszu wkłady niepieniężne w postaci akcji spółek akcyjnych, zwanych dalej spółkami, na zasadach określonych w art. 10.
E11	pf	pf	impf	pf	Upoważnieniu do wydania rozporządzenia można nadać charakter obligatoryjny albo fakultatywny.
E13	pf	pf	impf	pf	W protokole czynności przeprowadzonej z udziałem specjalistów należy wskazać ich imiona i nazwiska, specjalność, miejsce zamieszkania, miejsce pracy i stanowisko.
E15	pf	impf	impf	impf	Wysokość dofinansowania nie może wynieść więcej niż 7.500 złotych.

E16	pf	impf	impf	impf	W zależności od stopnia znajomości języka polskiego, liczba godzin dydaktycznych kursu języka polskiego może wynosić od 30 do 100 godzin.
E17	impf	impf	pf	impf	Liczba dozorów lub nadzorów sprawowanych przez kuratora społecznego nie powinna przekraczać 10.
ID	**general chunk pref**	**legal chunk pref**	**general unigram pref**	**legal unigram pref**	**sentence_PL**
E18	impf	impf	pf	impf	Wydatki na działalność przewidzianą w art. 66 nie powinny przekroczyć 1% planowanych wydatków funduszu w danym roku.
E19	impf	impf	pf	pf	Wolno zatrudniać tylko tych młodocianych, którzy: 1) ukończyli co najmniej szkołę podstawową, 2) przedstawią świadectwo lekarskie stwierdzające, że praca danego rodzaju nie zagraża ich zdrowiu.
E21	impf	impf	pf	impf	Materiały stosowane do budowy zbiorników powinny spełniać wymagania określone w Polskich Normach lub specyfikacjach technicznych uzgodnionych z organem właściwej jednostki dozoru.
E22	impf	impf	pf	impf	Programy studiów w uczelniach technicznych i rolniczych powinny spełnić zalecenia FEANI: 10% godzin - przedmioty kształcenia ogólnego, 35% - przedmioty podstawowe, 55% - przedmioty kierunkowe.
E23	pf	pf	impf	impf	Połączenie zakładów ubezpieczeń może nastąpić, jeżeli wykonują one działalność w takiej samej formie organizacyjnej.
E24	pf	pf	impf	impf	Zmiany kwot dotacji na zadania zlecone jednostkom samorządu terytorialnego mogą następować w terminie do dnia 15 listopada roku budżetowego.
E25	pf	pf	impf	pf	Pismo sądowe wraz z tekstem ogłoszenia można wydać stronie z zaznaczeniem, że koszty ogłoszenia wyłoży ta strona.
E27	impf	pf	impf	pf	W szczególnie uzasadnionych przypadkach Minister Spraw Wewnętrznych i Administracji może wydać zezwolenie na większą liczbę pojazdów straży gminnej niż określona w ust. 1.
E28	impf	pf	impf	pf	Organ nadzoru może wydawać zalecenia mające na celu usunięcie stwierdzonych nieprawidłowości i dostosowanie działalności brokerskiej do przepisów prawa.

E29	pf	impf	pf	pf	Za zgodą Komisji akcje spółki mogą nabywać inne krajowe i zagraniczne osoby prawne.
E30	pf	impf	pf	pf	Fundusz inwestycyjny otwarty nie może nabyć papierów wartościowych dających więcej niż 10% głosów w którymkolwiek organie emitenta tych papierów.
ID	**general chunk pref**	**legal chunk pref**	**general unigram pref**	**legal unigram pref**	**sentence_PL**
E31	pf	impf	pf	impf	W skład grup roboczych mogą wchodzić również specjaliści, eksperci, i osoby zaufania społecznego.
E32	pf	impf	pf	impf	Właściciel gruntu może wejść na grunt sąsiedni w celu usunięcia zwieszających się z jego drzew gałęzi lub owoców.
E33	pf	impf	pf	impf	W skład Komisji powinny wchodzić co najmniej dwie osoby posiadające wiedzę z zakresu matematyki ubezpieczeniowej, finansowej i statystyki.
E35	pf	impf	pf	impf	Ustalenia spełnienia lub nieistnienia obowiązku ubezpieczenia można dochodzić przed sądem powszechnym.

TABLE 5: FILLER STIMULI			
ID	**sentence_PL**	**option_1**	**option_2**
F1	Jeżeli dostawa towaru lub wykonanie usługi powinny być potwierdzone fakturą, obowiązek podatkowy powstaje z chwilą wystawienia faktury.	powinny	muszą
F2	Notatki, o których mowa w ust. 6, powinny zawierać omówienie wyniku ponownego rozpatrzenia sprawy, wyniku skargi lub rewizji.	o których mowa	wspomniane
F3	Nauczyciele gimnazjum muszą posiadać kwalifikacje określone w odrębnych przepisach.	odrębnych	osobnych
F4	Przechowawcy nie wolno używać rzeczy bez zgody składającego, chyba że jest to konieczne do jej zachowania w stanie nie pogorszonym.	konieczne do	niezbędne dla
F5	Rozporządzenie, o którym mowa w ust. 1, może być wydane także na wniosek zainteresowanej rady powiatu, rady miasta na prawach powiatu lub rady gminy.	także	również
F6	Postanowienie to należy uchylić , jeżeli ustała potrzeba dalszej ochrony osoby, która ma być ubezwłasnowolniona, lub jej mienia.	potrzeba	konieczność
F7	W przypadku zmiany rodzaju dostarczanego paliwa pobranie próbki powinno być wykonane	wystąpienia	nastąpienia

	w ciągu godziny od wystąpienia tej zmiany.		
F8	Przed salami ćwiczeń, basenów oraz salami hydroterapii powinny być urządzone szatnie oraz węzły sanitarne dla pacjentów.	powinny być urządzone	należy urządzić
F9	Umowa o oferowanie instrumentów finansowych powinna być zawarta w formie pisemnej pod rygorem nieważności.	zawarta	sporządzona
ID	**sentence_PL**	**option_1**	**option_2**
F10	Zapisy na koncie 130 są dokonywane na podstawie wyciągów bankowych, w związku z czym musi zachodzić pełna zgodność zapisów między jednostką a bankiem.	w związku z czym	dlatego też
F11	W jednym ogłoszeniu można zamieścić informacje o przetargach na jedną lub więcej nieruchomości.	można zamieścić	można zamieszczać
F12	Po upływie 3 lat od chwili czynu nie można wszcząć postępowania dyscyplinarnego, a w razie wszczęcia - ulega ono umorzeniu.	postępowania dyscyplinarnego	procedury dyscyplinarnej
F13	Nie jest dopuszczalne zastrzeżenie, że konsumentowi wolno odstąpić od umowy za zapłatą oznaczonej sumy (odstępne).	Nie jest dopuszczalne	Nie zezwala się na
F14	Wniosek o przyznanie prawa pomocy można złożyć za pośrednictwem Ministerstwa Sprawiedliwości Rzeczypospolitej Polskiej lub wyznaczonego w tym celu organu państwa członkowskiego.	w tym celu	ku temu
F15	Maszyn będących w ruchu nie wolno pozostawiać bez obsługi lub nadzoru, chyba że dokumentacja techniczna stanowi inaczej.	dokumentacja techniczna stanowi inaczej	w dokumentacji technicznej zaznaczono inaczej
F16	Dokument, o którym mowa w ust. 4, musi być przechowywany przez okres dłuższy o rok od terminu ważności produktu leczniczego.	terminu ważności	daty ważności
F17	Organy administracji rządowej i samorządowej mogą zlecać realizację zadań w trybie przewidzianym w ustawie z dnia 12 marca 2004 r. o pomocy społecznej.	z dnia 12 marca 2004 r. o pomocy społecznej.	o pomocy społecznej z dnia 12 marca 2004 r.
F18	Jeżeli istnieje możliwość powstania stężeń wybuchowych w zbiorniku, należy zastosować środki zapobiegające wybuchowi.	Jeżeli	Gdy
F19	Powództwo o roszczenia wynikające z umów ubezpieczenia można wytoczyć albo według przepisów o właściwości ogólnej albo przed sąd właściwy dla miejsca zamieszkania lub siedziby stron.	Powództwo o roszczenia	Wniosek o roszczenia
F20	Jeżeli umowę ubezpieczenia OC zawarto na okres dłuższy niż 12 miesięcy, każda ze stron umowy może ją wypowiedzieć najpóźniej na 30 dni przed upływem każdego 12-	najpóźniej na	nie później niż

	miesięcznego okresu trwania umowy.		
F21	Uwagi do projektu planu należy wnieść na piśmie w terminie wyznaczonym w ogłoszeniu.	Uwagi do	Uwagi na temat
F22	W przypadku zamierzonej zmiany umowy spółki należy wskazać treść zamierzonych zmian.	należy wskazać treść	konieczne jest wskazanie treści
ID	**sentence_PL**	**option_1**	**option_2**
F23	Jeżeli granice gruntów stały się sporne, a stanu prawnego nie można stwierdzić , ustala się granice według ostatniego spokojnego stanu posiadania.	stwierdzić	ustalić
F24	Uprzywilejowanie, o którym mowa w § 1, może dotyczyć w szczególności prawa głosu, prawa do dywidendy lub podziału majątku w przypadku likwidacji spółki.	w szczególności	zwłaszcza

TABLE 6: PRACTICE STIMULI			
ID	**sentence_PL**	**option_1**	**option_2**
T1	Funkcjonariuszowi można powierzyć pełnienie obowiązków służbowych na innym stanowisku na czas nieprzekraczający 12 miesięcy.	można powierzyć	można powierzać
T2	Aby zostać przyjętym na szkolenie, kandydat musi posiadać dyplom lub zaświadczenie, uprawniające go do wstępu na dany kierunek studiów na uniwersytecie w Państwie Członkowskim.	Aby zostać przyjętym	W celu przyjęcia
T3	Na opakowaniach substancji niebezpiecznych i preparatów niebezpiecznych nie wolno umieszczać oznaczeń wskazujących, że taka substancja lub taki preparat nie są niebezpieczne.	wolno umieszczać	można zamieszczać
T4	Jeżeli przychodnia (ośrodek zdrowia) mieści się w budynku o więcej niż 1 kondygnacji, budynek musi być wyposażony w dźwig.	mieści się	znajduje się
T5	Zbieranie pojazdów wycofanych z eksploatacji mogą prowadzić wyłącznie przedsiębiorcy prowadzący punkty zbierania pojazdów i przedsiębiorcy prowadzący stacje demontażu.	prowadzący	, którzy prowadzą
T6	Przy ustawianiu własnych urządzeń ogłoszeniowych w celu prowadzenia kampanii referendalnej należy stosować właściwe przepisy.	w celu	z zamiarem
T7	Nagrodę wypłaca się w formie bezpośredniej lub na rachunek w banku wskazanym w formie pisemnej, a w szczególnie uzasadnionych przypadkach może być ona przesłana na wskazany adres.	Nagrodę wypłaca się	Nagroda wypłacana jest

T8	Wspólnikowi nie wolno pobierać odsetek od wniesionych wkładów, jak również od przysługujących mu udziałów.	odsetek	oprocentowania

2. Experimental sets

2.1. Judgement task and self-paced reading task

2.1.1. Naive native speakers

TABLE 7: SUBSET 1.1			
sentence_type	**ID**	**aspectual_version**	**hypothesis**
practice	T8	na	na
practice	T1	na	na
practice	T2	na	na
practice	T23	na	na
filler	F9	na	na
filler	F6	na	na
experimental	E9	dispreferred	1
filler	F1	na	na
filler	F3	na	na
experimental	E16	dispreferred	1
filler	F43	na	na
experimental	E15	dispreferred	1
filler	F20	na	na
filler	F18	na	na
experimental	E1	preferred	1
filler	F40	na	na
filler	F12	na	na
filler	F31	na	na
experimental	E3	preferred	1
filler	F14	na	na
experimental	E8	preferred	1
filler	F39	na	na
experimental	E2	dispreferred	1
filler	F21	na	na
experimental	E18	preferred	2
filler	F13	na	na
filler	F27	na	na
experimental	E6	dispreferred	1
filler	F45	na	na
filler	F44	na	na
filler	F8	na	na

practice	T6	na	na
practice	T5	na	na
practice	T12	na	na
practice	T11	na	na

TABLE 8: SUBSET 1.2 (MIRROR)			
sentence_type	**ID**	**aspectual_version**	**hypothesis**
practice	T1	na	na
practice	T2	na	na
practice	T23	na	na
practice	T8	na	na
filler	F14	na	na
experimental	E18	dispreferred	2
filler	F6	na	na
experimental	E16	preferred	1
filler	F1	na	na
filler	F13	na	na
experimental	E1	dispreferred	1
filler	F8	na	na
filler	F31	na	na
experimental	E8	dispreferred	1
filler	F44	na	na
experimental	E15	preferred	1
filler	F9	na	na
filler	F20	na	na
experimental	E3	dispreferred	1
filler	F3	na	na
filler	F43	na	na
experimental	E9	preferred	1
filler	F39	na	na
filler	F12	na	na
filler	F27	na	na
experimental	E6	preferred	1
filler	F21	na	na
experimental	E2	preferred	1
filler	F40	na	na
filler	F45	na	na
filler	F18	na	na
practice	T12	na	na
practice	T11	na	na
practice	T5	na	na
practice	T6	na	na

TABLE 9: SUBSET 2.1			
sentence_type	**ID**	**aspectual_version**	**hypothesis**
practice	T4	na	na
practice	T10	na	na
practice	T3	na	na
practice	T17	na	na
experimental	E5	dispreferred	1
filler	F37	na	na
experimental	E17	dispreferred	2
filler	F36	na	na
filler	F34	na	na
experimental	E4	preferred	1
filler	F15	na	na
filler	F10	na	na
experimental	E11	dispreferred	1
filler	F35	na	na
experimental	E13	preferred	1
filler	F30	na	na
experimental	E12	dispreferred	1
filler	F26	na	na
filler	F25	na	na
filler	F17	na	na
experimental	E14	dispreferred	1
filler	F38	na	na
filler	F11	na	na
filler	F22	na	na
experimental	E10	dispreferred	1
filler	F23	na	na
filler	F16	na	na
filler	F29	na	na
experimental	E7	dispreferred	1
filler	F28	na	na
filler	F24	na	na
practice	T9	na	na
practice	T7	na	na
practice	T21	na	na
practice	T19	na	na

TABLE 10: SUBSET 2.2 (MIRROR)			
sentence_type	**ID**	**aspectual_version**	**hypothesis**
practice	T17	na	na
practice	T10	na	na
practice	T3	na	na
practice	T4	na	na
filler	F10	na	na
filler	F37	na	na
experimental	E17	preferred	2
filler	F26	na	na
filler	F16	na	na
experimental	E13	dispreferred	1
filler	F11	na	na
filler	F15	na	na
filler	F29	na	na
experimental	E12	preferred	1
filler	F38	na	na
filler	F35	na	na
experimental	E11	preferred	1
filler	F24	na	na
filler	F23	na	na
experimental	E10	preferred	1
filler	F30	na	na
filler	F25	na	na
experimental	E7	preferred	1
filler	F28	na	na
experimental	E5	preferred	1
filler	F17	na	na
experimental	E4	dispreferred	1
filler	F36	na	na
filler	F22	na	na
filler	F34	na	na
experimental	E14	preferred	1
practice	T7	na	na
practice	T19	na	na
practice	T9	na	na
practice	T21	na	na

2.1.2. Trainee translators

TABLE 11: SUBSET 1.1			
sentence_type	**ID**	**aspectual_version**	**hypothesis**
practice	T2	na	na
practice	T23	na	na
practice	T1	na	na
practice	T8	na	na
filler	F12	na	na
filler	F3	na	na
filler	F43	na	na
experimental	E11	preferred	1
filler	F8	na	na
experimental	E12	preferred	1
filler	F44	na	na
filler	F14	na	na
filler	F1	na	na
experimental	E2	dispreferred	1
filler	F18	na	na
experimental	E9	dispreferred	1
filler	F21	na	na
filler	F9	na	na
experimental	E10	preferred	1
filler	F27	na	na
filler	F6	na	na
experimental	E1	preferred	1
practice	T5	na	na
practice	T11	na	na
practice	T12	na	na
practice	T6	na	na

TABLE 12: SUBSET 1.2 (MIRROR)			
sentence_type	**ID**	**aspectual_version**	**hypothesis**
practice	T1	na	na
practice	T23	na	na
practice	T8	na	na
practice	T2	na	na

filler	F9	na	na
experimental	E12	dispreferred	1
filler	F44	na	na
filler	F3	na	na
experimental	E1	dispreferred	1
filler	F12	na	na
filler	F14	na	na
experimental	E11	dispreferred	1
filler	F21	na	na
experimental	E2	preferred	1
filler	F8	na	na
filler	F18	na	na
experimental	E10	dispreferred	1
filler	F1	na	na
experimental	E9	preferred	1
filler	F6	na	na
filler	F43	na	na
filler	F27	na	na
practice	T11	na	na
practice	T6	na	na
practice	T5	na	na
practice	T12	na	na

TABLE 13: SUBSET 2.1			
sentence_type	**ID**	**aspectual_version**	**hypothesis**
practice	T10	na	na
practice	T4	na	na
practice	T3	na	na
practice	T17	na	na
experimental	E17	preferred	2
filler	F16	na	na
filler	F23	na	na
experimental	E15	preferred	1

filler	F17	na	na
filler	F22	na	na
experimental	E3	preferred	1
filler	F10	na	na
filler	F29	na	na
experimental	E5	dispreferred	1
filler	F15	na	na
experimental	E8	preferred	1
filler	F30	na	na
filler	F36	na	na
experimental	E16	preferred	1
filler	F24	na	na
filler	F34	na	na
filler	F35	na	na
practice	T7	na	na
practice	T9	na	na
practice	T21	na	na
practice	T19	na	na

TABLE 14: SUBSET 2.2 (MIRROR)			
sentence_type	**ID**	**aspectual_version**	**hypothesis**
practice	T10	na	na
practice	T3	na	na
practice	T4	na	na
practice	T17	na	na
filler	F30	na	na
experimental	E17	dispreferred	2
filler	F34	na	na
filler	F23	na	na
experimental	E5	preferred	1
filler	F17	na	na
experimental	E3	dispreferred	1
filler	F15	na	na

filler	F36	na	na
experimental	E16	dispreferred	1
filler	F16	na	na
experimental	E8	dispreferred	1
filler	F24	na	na
filler	F29	na	na
filler	F10	na	na
experimental	E15	dispreferred	1
filler	F35	na	na
filler	F22	na	na
practice	T21	na	na
practice	T9	na	na
practice	T7	na	na
practice	T19	na	na

2.1.3. Professional translators

TABLE 15: SUBSET 1.1			
sentence_type	**ID**	**aspectual_version**	**hypothesis**
practice	T1	na	na
practice	T8	na	na
practice	T23	na	na
practice	T2	na	na
filler	F9	na	na
experimental	E2	preferred	1
filler	F18	na	na
filler	F27	na	na
experimental	E10	preferred	1
filler	F31	na	na
experimental	E6	preferred	1
filler	F3	na	na
filler	F14	na	na
filler	F6	na	na

experimental	E19	dispreferred	2
filler	F43	na	na
filler	F12	na	na
experimental	E21	dispreferred	2
filler	F44	na	na
experimental	E4	dispreferred	1
filler	F8	na	na
filler	F1	na	na
experimental	E13	dispreferred	1
filler	F13	na	na
filler	F21	na	na
practice	T6	na	na
practice	T11	na	na
practice	T5	na	na
practice	T12	na	na

TABLE 16: SUBSET 1.2 (MIRROR)

sentence_type	ID	aspectual_version	hypothesis
practice	T8	na	na
practice	T2	na	na
practice	T23	na	na
practice	T1	na	na
experimental	E2	dispreferred	1
filler	F18	na	na
filler	F14	na	na
experimental	E6	dispreferred	1
filler	F27	na	na
filler	F21	na	na
experimental	E21	preferred	2
filler	F8	na	na
experimental	E4	preferred	1
filler	F31	na	na
experimental	E10	dispreferred	1
filler	F6	na	na
filler	F44	na	na
filler	F13	na	na
experimental	E19	preferred	2

filler	F1	na	na
filler	F9	na	na
filler	F3	na	na
experimental	E13	preferred	1
filler	F43	na	na
filler	F12	na	na
practice	T6	na	na
practice	T5	na	na
practice	T11	na	na
practice	T12	na	na

TABLE 17: SUBSET 2.1			
sentence_type	**ID**	**aspectual_version**	**hypothesis**
practice	T17	na	na
practice	T3	na	na
practice	T10	na	na
practice	T4	na	na
experimental	E3	preferred	1
filler	F36	na	na
filler	F10	na	na
experimental	E16	preferred	1
filler	F22	na	na
filler	F17	na	na
experimental	E1	dispreferred	1
filler	F4	na	na
filler	F16	na	na
experimental	E8	preferred	1
filler	F30	na	na
filler	F35	na	na
experimental	E24	dispreferred	2
filler	F15	na	na
filler	F34	na	na
experimental	E11	preferred	1
filler	F24	na	na
filler	F23	na	na

filler	F29	na	na
experimental	E23	preferred	2
filler	F42	na	na
practice	T19	na	na
practice	T7	na	na
practice	T21	na	na
practice	T9	na	na

TABLE 18: SUBSET 2.2 (MIRROR)			
sentence_type	**ID**	**aspectual_version**	**hypothesis**
practice	T3	na	na
practice	T10	na	na
practice	T4	na	na
practice	T17	na	na
experimental	E3	dispreferred	1
filler	F34	na	na
experimental	E8	dispreferred	1
filler	F36	na	na
filler	F24	na	na
filler	F30	na	na
experimental	E1	preferred	1
filler	F29	na	na
filler	F23	na	na
filler	F42	na	na
experimental	E23	dispreferred	2
filler	F16	na	na
filler	F10	na	na
experimental	E11	dispreferred	1
filler	F4	na	na
experimental	E24	preferred	2
filler	F35	na	na
filler	F22	na	na
experimental	E16	dispreferred	1
filler	F17	na	na
filler	F15	na	na

practice	T21	na	na
practice	T9	na	na
practice	T19	na	na
practice	T7	na	na

2.2. Forced-choice task

2.2.1. Set 1

TABLE 19: SUBSET 1.1		
ID	**sentence_type**	**hypothesis**
T4	practice	na
T2	practice	na
T3	practice	na
T1	practice	na
F13	filler	na
E5	experimental	1
F21	filler	na
F8	filler	na
E32	experimental	2
F2	filler	na
E23	experimental	1
F3	filler	na
F14	filler	na
E28	experimental	2
F15	filler	na
E9	experimental	1
F6	filler	na
F7	filler	na
E8	experimental	1
F24	filler	ns
F16	filler	na
E17	experimental	1
F17	filler	na
F5	filler	na
F22	filler	na
E29	experimental	2
F12	filler	na
E1	experimental	1
F9	filler	na
E16	experimental	2

F23	filler	na
F1	filler	na
E4	experimental	1
F20	filler	na
F4	filler	na
F11	filler	na
F19	filler	na
F18	filler	na
E22	experimental	1
F10	filler	na
T5	practice	na
T6	practice	na
T8	practice	na
T7	practice	na

TABLE 20: SUBSET 1.2		
ID	**sentence_type**	**hypothesis**
T1	practice	na
T4	practice	na
T3	practice	na
T2	practice	na
F18	filler	na
F1	filler	na
E28	experimental	2
F23	filler	na
E5	experimental	1
F9	filler	na
E17	experimental	1
F11	filler	na
F13	filler	na
F4	filler	na
F19	filler	na
F8	filler	na
E32	experimental	2
F21	filler	na
E4	experimental	1
F10	filler	na
F16	filler	na
E8	experimental	1
F12	filler	na
F22	filler	na
E23	experimental	1
F6	filler	na
F2	filler	na
E22	experimental	1
F3	filler	na
F15	filler	na

E16	experimental	2
F14	filler	na
F17	filler	na
F5	filler	na
E9	experimental	1
F7	filler	na
E29	experimental	2
F20	filler	na
F24	filler	na
E1	experimental	1
T7	practice	na
T5	practice	na
T8	practice	na
T6	practice	na

TABLE 21: SUBSET 1.3		
ID	**sentence_type**	**hypothesis**
T2	practice	na
T3	practice	na
T4	practice	na
T1	practice	na
F1	filler	na
E8	experimental	1
F21	filler	na
F7	filler	na
E17	experimental	1
F22	filler	na
E16	experimental	2
F5	filler	na
F9	filler	na
F20	filler	na
E4	experimental	1
F11	filler	na
F12	filler	na
E9	experimental	1
F16	filler	na
F2	filler	na
E1	experimental	1
F13	filler	na
F23	filler	na
F15	filler	na
E5	experimental	1
F8	filler	na
F14	filler	na
F18	filler	na
E23	experimental	1
F17	filler	na

E32	experimental	2
F19	filler	na
F4	filler	na
F24	filler	na
E29	experimental	2
F6	filler	na
E22	experimental	1
F3	filler	na
E28	experimental	2
F10	filler	na
T6	practice	na
T7	practice	na
T5	practice	na
T8	practice	na

TABLE 22: SUBSET 1.4		
ID	**sentence_type**	**hypothesis**
T3	practice	na
T4	practice	na
T2	practice	na
T1	practice	na
F1	filler	na
E16	experimental	2
F14	filler	na
E23	experimental	1
F9	filler	na
F7	filler	na
E29	experimental	2
F20	filler	na
E32	experimental	2
F6	filler	na
F8	filler	na
E9	experimental	1
F18	filler	na
F17	filler	na
E4	experimental	1
F13	filler	na
F23	filler	na
E17	experimental	1
F16	filler	na
F22	filler	na
E8	experimental	1
F24	filler	na
F19	filler	na
F5	filler	na
E5	experimental	1
F12	filler	na

F15	filler	na
E22	experimental	1
F10	filler	na
E1	experimental	1
F4	filler	na
F21	filler	na
E28	experimental	2
F3	filler	na
F2	filler	na
F11	filler	na
T5	practice	na
T6	practice	na
T7	practice	na
T8	practice	na

2.2.2. Set 2

TABLE 23: SUBSET 2.1		
ID	**sentence_type**	**hypothesis**
T3	practice	na
T2	practice	na
T4	practice	na
T1	practice	na
F13	filler	na
E10	experimental	1
F24	filler	na
E31	experimental	2
F12	filler	na
F22	filler	na
F16	filler	na
E30	experimental	2
F4	filler	na
E24	experimental	1
F17	filler	na
F10	filler	na
F15	filler	na
E18	experimental	1
F6	filler	na
F2	filler	na
E6	experimental	1
F18	filler	na
E15	experimental	2
F5	filler	na
E2	experimental	1
F9	filler	na
E21	experimental	1
F3	filler	na

F7	filler	na
E3	experimental	1
F19	filler	na
F1	filler	na
F8	filler	na
F20	filler	na
E27	experimental	2
F23	filler	na
E7	experimental	1
F14	filler	na
F21	filler	na
F11	filler	na
T8	practice	na
T6	practice	na
T5	practice	na
T7	practice	na

TABLE 24: SUBSET 2.2

ID	sentence_type	hypothesis
T4	practice	na
T1	practice	na
T3	practice	na
T2	practice	na
F7	filler	na
E3	experimental	1
F3	filler	na
F14	filler	na
E10	experimental	1
F19	filler	na
F11	filler	na
F22	filler	na
F15	filler	na
E7	experimental	1
F12	filler	na
F9	filler	na
E15	experimental	2
F21	filler	na
F2	filler	na
E27	experimental	2
F24	filler	na
E6	experimental	1
F13	filler	na
F10	filler	na
E30	experimental	2
F8	filler	na
E2	experimental	1
F5	filler	na
F4	filler	na
F16	filler	na

E18	experimental	1
F20	filler	na
F6	filler	na
F1	filler	na
E21	experimental	1
F23	filler	na
E31	experimental	2
F17	filler	na
E24	experimental	1
F18	filler	na
T6	practice	na
T8	practice	na
T5	practice	na
T7	practice	na

TABLE 25: SUBSET 2.3		
ID	**sentence_type**	**hypothesis**
T3	practice	na
T4	practice	na
T1	practice	na
T2	practice	na
F7	filler	na
F15	filler	na
E6	experimental	1
F6	filler	na
F8	filler	na
F1	filler	na
F17	filler	na
E24	experimental	1
F18	filler	na
F24	filler	na
E3	experimental	1
F14	filler	na
F13	filler	na
E21	experimental	1
F10	filler	na
F20	filler	na
E15	experimental	2
F16	filler	na
F19	filler	na
E2	experimental	1
F2	filler	na
E30	experimental	2
F12	filler	na
F11	filler	na
E31	experimental	2
F4	filler	na

F3	filler	na
E27	experimental	2
F5	filler	na
E10	experimental	1
F21	filler	na
F9	filler	na
E18	experimental	1
F22	filler	na
F23	filler	na
E7	experimental	1
T6	practice	na
T8	practice	na
T7	practice	na
T5	practice	na

TABLE 26: SUBSET 2.4		
ID	**sentence_type**	**hypothesis**
T2	practice	na
T1	practice	na
T4	practice	na
T3	practice	na
F7	filler	na
F14	filler	na
E24	experimental	1
F17	experimental	na
F15	filler	na
F16	filler	na
E6	experimental	1
F13	filler	na
F24	filler	na
E21	experimental	1
F11	filler	na
E30	experimental	2
F21	filler	na
E7	experimental	1
F10	filler	na
F19	filler	na
E15	experimental	2
F9	filler	na
E10	experimental	1
F23	filler	na
E3	experimental	1
F5	filler	na
F22	filler	na
E27	experimental	2
F4	filler	na
F3	filler	na

F6	filler	na
E2	experimental	1
F12	filler	na
F20	filler	na
E31	experimental	2
F18	filler	na
F2	filler	na
E18	experimental	1
F8	filler	na
F1	filler	na
T5	practice	na
T7	practice	na
T8	practice	na
T6	practice	na

Appendix 3. R code

1. Data preparation

1.1. Judgement task (Section 7.1.5.1)

```
library(mixtools)

ratings_all <- read.table(file.choose(), header=T, sep="\t")
#load file 'ratings_all.txt'

attach(ratings_all)
ratings = ratings_all$rating
ratings_model<-normalmixEM(x=ratings, k=2)
x<-ratings
index.lower <- which.min(ratings_model$mu)
find.cutoff <- function(proba=0.5, i=index.lower) {f <- function(x) {proba
- (ratings_model$lambda[i]*dnorm(x, ratings_model$mu[i],
ratings_model$sigma[i]) /(ratings_model$lambda[1]*dnorm(x,
ratings_model$mu[1], ratings_model$sigma[1]) +
ratings_model$lambda[2]*dnorm(x, ratings_model$mu[2],
ratings_model$sigma[2])))}
return(uniroot(f=f,lower=as.numeric(quantile(ratings_all$rating,0.05)),
upper=as.numeric(quantile(ratings_all$rating,0.95)))$root)}

cutoffs <- c(find.cutoff(proba=0.5), find.cutoff(proba=0.75))  # Around
c(1.8, 1.5)

hist(x)
abline(v=cutoffs, col=c("red", "blue"), lty=2)
cutoffs

ratings_all$rating <- ifelse(ratings_all$rating >=67.57894, 1, 0)
#re-codes all ratings above or equal to 67.57894 as 1, and all ratings belo
w the threshold as 0
```

1.2. Self-paced reading task (Section 7.1.5.2)

1.2.1. Modal chunk

```
library(lattice)
library(car)

rt_all <- read.table(file.choose(), header=T, sep="\t")
#load file 'rt_all.txt'

dip.test(rtr_all$RT, simulate.p.value = FALSE)
#checks for unimodality

qqPlot(rt_all$RT)
plot(density(rt_all$RT))
#to visualise the data
rt_all1=rt_all[rt_all$RT<3.5,]
#removes the outliers above 3.5
```

```
shapiro.test(rt_all1$logRT)
#test for normality; distribution not normal so log-transformation is requi
red

rt_all1$logRT=log(rt_all1$RT)
#log-transforms the data

qqPlot(rt_all1$logRT)
#to visualise the data


rt_all2=rt_all1[rt_all1$logRT>-1.5,]
#removes the outliers below -1.5

shapiro.test(rt_all2$logRT)
#test for normality
```

1.2.2. Post-modal chunk (spillover effect)

```
spillover_all <- read.table(file.choose(), header=T, sep="\t")
#load file 'spillover_all.txt'

dip.test(rtr_all$RT, simulate.p.value = FALSE)
#checks for unimodality

qqPlot(spillover_all$RT)
plot(density(spillover_all$RT))
#to visualise the data
spillover_all1=spillover_all[spillover_all$RT<3.4,]
#removes the outliers above 3.4

shapiro.test(spillover_all1$RT)
#test for normality; distribution not normal so log-transformation is requi
red

spillover_all1$logRT=log(spillover_all1$RT)
#log-transforms the data

qqPlot(spillover_all1$logRT)
#to visualise the data

spillover_all2=spillover_all1[spillover_all1$logRT>-1,]
#removes the outliers below -1
shapiro.test(spillover_all2$logRT)
#test for normality
```

2. Regression

2.1. Analogical mapping (Section 5.2)

```
library(MASS)
library(rms)

t_musiec <- read.table(file.choose(), header=T, sep=" ")
#load file 't_musiec.txt'

t_musiec1 = glm(aspect ~ modality, data = t_musiec, family = "binomial")
summary(t_musiec1)
```

2.2. Corpus study, general chunking hypothesis (Section 6.2.1)

```
library(MASS)
library(rms)

all_musiec <- read.table(file.choose(), header=T, sep=" ")
#load file 'all_musiec.txt'
```

2.2.1. General chunk model

```
all_glm1 = glm(aspect ~ cg_pref, data = all_musiec, family = "binomial")
summary(all_glm1)

all_lrm1 = lrm(aspect ~ cg_pref, data = all_musiec, x=T, y=T)
all_lrm1
```

2.2.2. General unigram model

```
all_glm2 = glm(aspect ~ ug_pref, data = all_musiec, family = "binomial")
summary(all_glm2)

all_lrm2 = lrm(aspect ~ ug_pref, data = all_musiec, x=T, y=T)
all_lrm2
```

2.3. Corpus study, genre-specific chunking hypothesis (Section 6.2.2)

2.3.1. Non-translated data

```
library(MASS)
library(rms)

nt_musiec <- read.table(file.choose(), header=T, sep=" ")
#load file 'nt_musiec.txt'
```

General chunk model

```
nt_glm1 = glm(aspect ~ cg_pref, data = nt_musiec, family = "binomial")
summary(nt_glm1)

nt_lrm1 = lrm(aspect ~ cg_pref, data = nt_musiec, x=T, y=T)
nt_lrm1
```

Legal chunk model

```
nt_glm2 = glm(aspect ~ cl_pref, data = nt_musiec, family = "binomial")
summary(nt_glm2)

nt_lrm2 = lrm(aspect ~ cl_pref, data = nt_musiec, x=T, y=T)
nt_lrm2
```

Legal unigram model

```
nt_glm3 = glm(aspect ~ ul_pref, data = nt_musiec, family = "binomial")
summary(nt_glm3)

nt_lrm3 = lrm(aspect ~ ul_pref, data = nt_musiec, x=T, y=T)
nt_lrm3
```

2.3.2. Translated data

```
nt_musiec <- read.table(file.choose(), header=T, sep=" ")
#load file 'nt_musiec.txt'
```

General chunk model

```
t_glm1 = glm(aspect ~ cg_pref, data = t_musiec, family = "binomial")
summary(t_glm1)

t_lrm1 = lrm(aspect ~ cg_pref, data = t_musiec, x=T, y=T)
t_lrm1
```

Legal chunk model

```
t_glm2 = glm(aspect ~ cl_pref, data = t_musiec, family = "binomial")
summary(t_glm2)

t_lrm2 = lrm(aspect ~ cl_pref, data = t_musiec, x=T, y=T)
t_lrm2
```

Legal unigram model

```
t_glm3 = glm(aspect ~ ul_pref, data = t_musiec, family = "binomial")
summary(t_glm3)

t_lrm3 = lrm(aspect ~ ul_pref, data = t_musiec, x=T, y=T)
t_lrm3
```

2.4. Corpus study, bootstrap validation (Section 6.2.3)

Non-translated data, legal chunk model

```
validate(nt_lrm2, bw=T, B=200)
```

Translated data, general chunk model

```
validate(t_lrm1, bw=T, B=200)
```

2.5. Experimental studies, general chunking hypothesis (Section 7.2.1)

2.5.1. Judgement task

```
attach(ratings_all)
#this is the binarised data

ratings_hyp1<-subset(ratings_all, condition<3)
#creates a new set with observations that test the general chunking
hypothesis
```

Model (1)

```
ratings1<-glmer(rating ~ version + (1 | participant), data=ratings_hyp1,
family=binomial)
summary(ratings1)
```

Model (2)

```
ratings2 <- glmer(rating ~ version + (1|participant) + (1|chunk), data=rati
ngs_hyp1, family=binomial)
summary(ratings2)
```

Comparison of model (1) and (2)

```
anova(ratings2, ratings1)
```

Group interaction

```
ratings2.1 <- glmer(rating ~ version*group + (1|participant) + (1|chunk), d
ata=ratings_hyp1, family=binomial)
summary(ratings2.1)
```

Individual groups: naive native speakers

```
ratings_ns_hyp1<-ratings_hyp1[which(ratings_hyp1$group=="ns"),]
#creates a new subset with observations from naive native speakers

ratings_ns <- glmer(rating ~ version + (1|participant) + (1|chunk), data=ra
tings_ns_hyp1, family=binomial)
summary(ratings_ns)
```

Individual groups: trainee translators

```
ratings_tt_hyp1<-ratings_hyp1[which(ratings_hyp1$group=="tt"),]
#creates a new subset with observations from trainee translators

ratings_tt <- glmer(rating ~ version + (1|participant) + (1|chunk), data=ra
tings_tt_hyp1, family=binomial)
summary(ratings_tt)
```

Individual groups: professional translators

```
ratings_pt_hyp1<-ratings_hyp1[which(ratings_hyp1$group=="pt"),]
#creates a new subset with observations from professional translators

ratings_pt <- glmer(rating ~ version + (1|participant) + (1|chunk), data=ra
tings_pt_hyp1, family=binomial)
```

```
summary(ratings_pt0
```

Age interaction

```
ratings2.2 <- glmer(rating ~ version*age + (1|participant) + (1|chunk), dat
a=ratings_hyp1, family=binomial)
summary(ratings2.2)
```

2.5.2. Self-paced reading task

```
rt_hyp1<-subset(rt_all2, condition<3)
#creates a new set with observations that test the general chunking
hypothesis
```

Model (1)

```
rt1 <- lmer(logRT~ version+sentence_position+chunk_position+chunk_length+(1
|participant),data=rt_hyp1,control=lmerControl(optimizer="bobyqa"))
summary(rt1)
```

Model (2)

```
rt2 <- lmer(logRT~ version+sentence_position+chunk_position+chunk_length+(1
|participant)+(1|chunk),data=rt_hyp1,control=lmerControl(optimizer="bobyqa"
))
summary(rt2)
```

Comparison of model (1) and (2)

```
anova(rt2, rt1)
```

Group interaction

```
rt1.1<-lmer(logRT~ version*group + sentence_position + chunk_position +
chunk_length + (1 | participant), data=rt_hyp1,
control=lmerControl(optimizer="bobyqa"))
summary(rt1.1)
```

Individual groups: naive native speakers

```
ratings_ns_hyp1<-rt_hyp1[which(rt_hyp1$group=="ns"),]
#creates a new subset with observations from naive native speakers

rt_ns<-lmer(logRT~ version +sentence_position +chunk_position +chunk_length
+(1|participant), data=ratings_ns_hyp1, control =
lmerControl(optimizer="bobyqa"))
summary(rt_ns)
```

Individual groups: trainee translators

```
ratings_tt_hyp1<-rt_hyp1[which(rt_hyp1$group=="tt"),]
#creates a new subset with observations from trainee translators

rt_tt<-lmer(logRT~ version + sentence_position + chunk_position +
chunk_length + (1|participant), data=ratings_tt_hyp1, control =
lmerControl(optimizer="bobyqa"))
summary(rt_tt)
```

Individual groups: professional translators

```
ratings_pt_hyp1<-rt_hyp1[which(rt_hyp1$group=="pt"),]
#creates a new subset with observations from professional translators

rt_pt<-lmer(logRT~ version +sentence_position +chunk_position +chunk_length
+(1|participant), data=ratings_pt_hyp1,
control=lmerControl(optimizer="bobyqa"))
summary(rt_pt)
```

Age interaction

```
rt_1.2<-lmer(logRT~ version*age +sentence_position +chunk_position
+chunk_length +(1|participant), data=rt_hyp1,
control=lmerControl(optimizer="bobyqa"))
summary(rt_1.2)
```

Spillover effect; model (1)

```
spillover_hyp1<-subset(spillover_all2, condition<3)
#creates a new set with observations that test the general chunking
hypothesis
```

```
spillover1<-lmer(logRT~ version+ sentence_position+ chunk_position+ chunk_l
ength+ (1|participant), data= spillover_hyp1, control=lmerControl(optimizer
="bobyqa"))
summary(spillover1)
```

Spillover effect; model (2)

```
spillover2<-lmer(logRT~ version+ sentence_position+ chunk_position+ chunk_l
ength+ (1|participant)+ (1|chunk), data= spillover_hyp1, control=lmerContro
l(optimizer="bobyqa"))
summary(spillover2)
```

Spillover effect; comparison of model (1) and (2)

```
anova(spillover2, spillover1)
```

Spillover effect; group interaction

```
spillover1.1<-lmer(logRT~ version*group+ sentence_position+ chunk_position+
chunk_length+ (1|participant), data= spillover_hyp1, control=lmerControl(op
timizer="bobyqa"))
summary(spillover1.1)
```

Spillover effect; individual groups: naive native speakers

```
spillover_ns_hyp1<- spillover_hyp1[which(spillover_hyp1$group=="ns"),]
#creates a new subset with observations from naive native speakers

spillover_ns <- lmer(logRT~ version+sentence_position+chunk_position+chunk_
length+(1|participant),data=spillover_ns_hyp1,control=lmerControl(optimizer
="bobyqa"))
summary(spillover_ns)
```

Spillover effect; individual groups: trainee translators

```
spillover_tt_hyp1<- spillover_hyp1[which(spillover_hyp1$group=="tt"),]
#creates a new subset with observations from trainee translators

spillover_tt <- lmer(logRT~ version+sentence_position+chunk_position+chunk_
length+(1|participant),data=spillover_tt_hyp1,control=lmerControl(optimizer
="bobyqa"))
summary(spillover_tt)
```

Spillover effect; individual groups: professional translators

```
spillover_pt_hyp1<- spillover_hyp1[which(spillover_hyp1$group=="pt"),]
#creates a new subset with observations from professional translators

spillover_pt <- lmer(logRT~ version+sentence_position+chunk_position+chunk_
length+(1|participant),data=spillover_pt_hyp1,control=lmerControl(optimizer
="bobyqa"))
summary(spillover_pt)
```

Spillover effect; age interaction

```
spillover1.2<-lmer(logRT~ version*age+sentence_position+chunk_position+chun
k_length+(1|participant), data= spillover_hyp1, control=lmerControl(optimiz
er="bobyqa"))
summary(spillover1.2)
```

2.5.3. Forced-choice task

```
fc_hyp1<-subset(fc_all, condition<3)
#creates a new set with observations that test the general chunking
hypothesis
```

Model (1)

```
choice1 <- glmer(choice ~ cg_pref + (1|participant), data = fc_hyp1, family
= binomial)
summary(choice1)
```

Model (2)

```
choice2 <- glmer(choice ~ cg_pref + (1|participant) + (1|sentence), data =
fc_hyp1, family = binomial)
summary(choice2)
```

Comparison of model (1) and (2)

```
anova(choice2, choice1)
```

Group interaction

```
choice2.1 <- glmer(choice ~ cg_pref*group + (1|participant) + (1|sentence),
data = fc_hyp1, family = binomial)
summary(choice2.1)
```

Individual groups: naive native speakers

```
fc_ns_hyp1<- fc_hyp1[which(fc_hyp1$group=="ns"),]
#creates a new subset with observations from naive native speakers

choice_ns <- glmer(choice ~ cg_pref + (1|participant) + (1|sentence), data
= ns_hyp1, family = binomial)
summary(fc_choice_ns)
```

Individual groups: trainee translators

```
fc_tt_hyp1<- fc_hyp1[which(fc_hyp1$group=="tt"),]
#creates a new subset with observations from trainee translators

choice_tt <- glmer(choice ~ cg_pref + (1|participant) + (1|sentence), data
= fc_tt_hyp1, family = binomial)
summary(choice_tt)
```

Individual groups: professional translators

```
fc_pt_hyp1<- fc_hyp1[which(fc_hyp1$group=="pt"),]
#creates a new subset with observations from professional translators

choice_pt <- glmer(choice ~ cg_pref + (1|participant) + (1|sentence), data
= fc_pt_hyp1, family = binomial)
summary(choice_pt)
```

Age interaction

```
choice2.2 <- glmer(choice ~ cg_pref*age + (1|participant) + (1|sentence), d
ata = fc_hyp1, family = binomial)
summary(choice2.2)
```

2.6. Experimental studies, genre-specific chunking hypothesis (Section 7.2.2)

2.6.1. Judgement task

```
ratings_hyp2<-subset(ratings_all, condition>2)
#creates a new set with observations that test the genre-specific chunking
hypothesis
```

Model (1)

```
ratings1<-glmer(rating ~ version + (1|participant), data=ratings_hyp2,
family=binomial)
summary(ratings1)
```

Model (2)

```
ratings2 <- glmer(rating ~ version + (1|participant) + (1|chunk), data=rati
ngs_hyp2, family=binomial)
summary(ratings2)
```

Comparison of model (1) and (2)

```
anova(ratings2, ratings1)
```

Group interaction

```
ratings2.1 <- glmer(rating ~ version*group + (1|participant) + (1|chunk), d
ata=ratings_hyp2, family=binomial)
summary(ratings2.1)
```

Individual groups: naive native speakers

```
ratings_ns_hyp2<-ratings_hyp2[which(ratings_hyp2$group=="ns"),]
#creates a new subset with observations from naive native speakers

ratings_ns <- glmer(rating ~ version + (1|participant) + (1|chunk), data=ra
tings_ns_hyp2, family=binomial)
summary(ratings_ns)
```

Individual groups: trainee translators

```
ratings_tt_hyp2<-ratings_hyp2[which(ratings_hyp2$group=="tt"),]
#creates a new subset with observations from trainee translators

ratings_tt <- glmer(rating ~ version + (1|participant) + (1|chunk), data=ra
tings_tt_hyp2, family=binomial)
summary(ratings_tt)
```

Individual groups: professional translators

```
ratings_pt_hyp2<-ratings_hyp2[which(ratings_hyp2$group=="pt"),]
#creates a new subset with observations from professional translators

ratings_pt <- glmer(rating ~ version + (1|participant) + (1|chunk), data=ra
tings_pt_hyp2, family=binomial)
summary(ratings_pt)
```

Age interaction

```
ratings2.2 <- glmer(rating ~ version*age + (1|participant) + (1|chunk), dat
a=ratings_hyp2, family=binomial)
summary(ratings2.2)
```

2.6.2. Self-paced reading task

```
rt_hyp2<-subset(rt_all2, condition>2)
#creates a new set with observations that test the genre-specific chunking
hypothesis
```

Model (1)

```
rt1<-lmer(logRT ~ version + sentence_position + chunk_position +
chunk_length +(1 | participant), data=rt_hyp2,
control=lmerControl(optimizer="bobyqa"))
summary(rt1)
```

Model (2)

```
rt2 <- lmer(logRT~ version+sentence_position+chunk_position+chunk_length+(1
|participant)+(1|chunk),data=rt_hyp2,control=lmerControl(optimizer="bobyqa"
))
summary(rt2)
```

Comparison of model (1) and (2)

```
anova(rt2, rt1)
```

Group interaction

```
rt1.1<-lmer(logRT~ version*group+ sentence_position + chunk_position + chun
k_length + (1 | participant), data=rt_hyp2, control=lmerControl(optimizer="
bobyqa"))
summary(rt1.1)
```

Individual groups: naive native speakers[27]

```
rt_ns_hyp2<-rt_hyp2[which(rt_hyp2$group=="ns"),]
#creates a new subset with observations from naive native speakers
```

```
rt_ns<-lm(logRT~ version+sentence_position+chunk_position+chunk_length, dat
a=rt_ns_hyp2)
summary(rt_ns)
```

Individual groups: trainee translators[28]

```
rt_tt_hyp2<-rt_hyp2[which(rt_hyp2$group=="tt"),]
#creates a new subset with observations from professional translators
```

```
rt_tt<-lm(logRT~ version+sentence_position+chunk_position+chunk_length, dat
a=rt_tt_hyp2)
summary(rt_tt)
```

Individual groups: professional translators

```
rt_pt_hyp2<-rt_hyp2[which(rt_hyp2$group=="pt"),]
#creates a new subset with observations from professional translators
```

```
rt_pt<-lmer(logRT~ version+sentence_position+chunk_position+chunk_length +
(1|participant), data=rt_pt_hyp2, control=lmerControl(optimizer="bobyqa"))
summary(rt_pt)
```

Age interaction

```
rt1.2<-lmer(logRT~ version*age+ sentence_position + chunk_position + chunk_
length + (1 | participant), data=rt_hyp2, control=lmerControl(optimizer="bo
byqa"))
summary(rt1.2)
```

Spillover effect; model (1)

```
spillover_hyp2<-subset(spillover_all2, condition>2)
#creates a new set with observations that test the genre-specific chunking
hypothesis
spillover1<-lmer(RT~ version+sentence_position+chunk_position+chunk_length+
(1|participant), data=spillover_hyp2, control=lmerControl(optimizer="bobyqa
"))
summary(spillover1)
```

Spillover effect; model (2)

```
spillover2<-lmer(logRT~ version+sentence_position+chunk_position+chunk_leng
th+(1|participant)+(1|chunk), data=spillover_hyp2, control=lmerControl(opti
mizer="bobyqa"))
summary(spillover2)
```

Spillover effect; comparison of model (1) and (2)

```
anova(spillover2, spillover1)
```

Spillover effect; group interaction

```
spillover1.1<-lmer(logRT~ version*group+sentence_position+chunk_position+ch
unk_length+(1|participant), data=spillover_hyp2, control=lmerControl(optimi
zer="bobyqa"))
```

[27] Naive native speakers saw one item per participant so random effect for participant cannot be included. A model with only fixed effects was fitted.

[28] Trainee translators saw one item per participant so random effect for participant cannot be included. A model with only fixed effects was fitted.

```
summary(spillover1.1)
```

Spillover effect; individual groups: naive native speakers[29]

```
spillover_ns_hyp2<- spillover_hyp2[which(spillover_hyp2$group=="ns"),]
#creates a new subset with observations from naive native speakers

spillover_ns<-lm(logRT~ version+sentence_position+chunk_position+chunk_leng
th, data=spillover_ns_hyp2)
summary(spillover_ns)
```

Spillover effect; individual groups: trainee translators[30]

```
spillover_tt_hyp2<-spillover_hyp2[which(spillover_hyp2$group=="tt"),]
#creates a new subset with observations from trainee translators

spillover_tt<-lm(logRT~ version+sentence_position+chunk_position+chunk_leng
th, data=spillover_tt_hyp2)
summary(spillover_tt)
```

Spillover effect; individual groups: professional translators

```
spillover_pt_hyp2<-spillover_hyp2[which(spillover_hyp2$group=="pt"),]
#creates a new subset with observations from professional translators

spillover_pt<-lmer(logRT~ version+sentence_position+chunk_position+chunk_le
ngth+(1|participant), data=spillover_pt_hyp2, control=lmerControl(optimizer
="bobyqa"))
summary(spillover_pt)
```

Spillover effect; age interaction

```
spillover1.2<-lmer(logRT~ version*age+sentence_position+chunk_position+chun
k_length+(1|participant), data=spillover_hyp2, control=lmerControl(optimize
r="bobyqa"))
summary(spillover1.2)
```

2.6.3. Forced-choice task

```
fc_hyp2<-subset(fc_all, condition>2)
#creates a new set with observations that test the genre-specific chunking
hypothesis
```

Model (1)

```
choice1<- glmer(choice ~ cg_pref + (1|participant), data = fc_hyp2, family
= binomial)
summary(choice1)
```

Model (2)

```
choice2<- glmer(choice ~ cg_pref + (1|participant) + (1|sentence), data = f
c_hyp2, family = binomial)
summary(choice2)
```

Comparison of model (1) and (2)

```
anova(choice2, choice1)
```

Group interaction

```
choice2.1<- glmer(choice ~ cg_pref*group + (1|participant) + (1|sentence),
data = fc_hyp2, family = binomial)
summary(choice2.1)
```

Individual groups: naive native speakers

```
fc_ns_hyp2<- fc_hyp2[which(fc_hyp2$group=="ns"),]
#creates a new subset with observations from naive native speakers
```

[29] Naive native speakers saw one item per participant so random effect for participant cannot be included. A model with only fixed effects was fitted.

[30] Trainee translators saw one item per participant so random effect for participant cannot be included. A model with only fixed effects was fitted.

```
choice_ns <- glmer(choice ~ cg_pref + (1|participant) + (1|sentence), data
= fc_ns_hyp2, family = binomial)
summary(choice_ns)
```

Individual groups: trainee translators

```
fc_tt_hyp2<- fc_hyp2[which(fc_hyp2$group=="tt"),]
#creates a new subset with observations from trainee translators

choice_tt <- glmer(choice ~ cg_pref + (1|participant) + (1|sentence), data
= fc_tt_hyp2, family = binomial)
summary(choice_tt)
```

Individual groups: professional translators

```
fc_pt_hyp2<- fc_hyp2[which(fc_hyp2$group=="pt"),]
#creates a new subset with observations from professional translators

choice_pt <- glmer(choice ~ cg_pref + (1|participant) + (1|sentence), data
= fc_pt_hyp2, family = binomial)
summary(choice_pt)
```

Age interaction

```
choice2.2<- glmer(choice ~ cg_pref*age + (1|participant) + (1|sentence), da
ta = fc_hyp2, family = binomial)
summary(choice2.2)
```