

IS PLAIN LEGAL LANGUAGE EASY TO TRANSLATE?

Plain English features in machine translation of a contract into Polish

Anna SETKOWICZ-RYSZKA*
PhD student (University of Łódź, Poland)

Abstract:

This article examines two machine translations of the same English contract (contract of supply) made using DeepL Pro in March 2022 and March 2023. The contract was used in my practice as a legal translator trainer. It proved easy to understand but difficult to translate into Polish, where plain legal language is much less developed than in English. The history of plain language in the UK, the US and Poland is briefly presented. The analysis of features of plain English covers: expressions from general language when mixed with legal or technical language, the pronoun you referring to one party to the contract, and other features, like active voice and short sentences. Additionally, renditions of expressions denoting obligation or permission are analysed separately as an important element of contracts. The conclusion is that pronoun you in this context constituted a major source of difficulty and was translated in many ways. The translations are also uneven in terms of register (formal legal language vs language of instructions) or forms of address (degree of politeness), which would make post-editing them demanding. Other problems that a post-editor would face include terminological inconsistency and increased number of nominalizations compared to source text. The active voice or length of sentences are unproblematic in MT. Both legal translators and MT providers should refer to existing plain Polish contracts for solutions.

Keywords: legal translation, contract translation, machine translation of legal texts, plain language, deontic modality

* ORCID: <https://orcid.org/0000-0003-0057-3477>

1. Introduction

Since I started teaching legal translation, I have had many eye-opening moments. One of them was discovering that a plain English contract was difficult to translate for trainees. The problems did not concern source text (ST) comprehension but rendering it in Polish. It should not have been a surprise. Plain English appeared in contracts in the 1980s in response to critiques of their incomprehensibility for an average consumer ('gobbledygook', 'small print') and has had much success. Plain language contracts were found to be more intelligible than traditional legal writing. They are written in a more colloquial and reader-oriented style. Meanwhile, plain Polish only appeared in official texts around 2010 and did not make its way to contracts until 2020. Plain legal style is still developing, and plain Polish contracts are rare, so there is little reference material. Most Polish contracts still contain low-frequency vocabulary, rather complex sentences, and other features like those that champions of plain English criticized.

But what happens when a plain-language contract is machine-translated? Current machine translation (MT) engines are trained on large volumes of parallel texts, i.e., source texts and their human translations. How successful is DeepL Pro, a generic MT engine claiming to outperform competitors¹, in dealing with the plain language features that proved difficult for trainee translators? How well does it cope with a change of register towards a more informal one? And finally, has anything changed between March 2022 and March 2023, when the MT was done?

This article analyses translation solutions applied by DeepL Pro in a contract with some plain English features. Unlike in the case of trainees' translations, it is not a holistic assessment. The focus is on Polish renditions of specific plain language features. Drawing on my own experience in translating contracts, I try to assess how useful MT output would be if I were to post-edit it. But before embarking on a detailed analysis, it is useful to remind a few facts about plain language in general and in contracts in particular. The following section looks briefly at the history of plain language in the UK and the US, contrasting it with the much shorter history of plain Polish. Section 3 reviews some literature on the use of MT in legal texts, including in contracts, while section 4 looks at the features of plain contract language in two DeepL translations into Polish, trying to assess which ones cause problems in MT and which do not. Finally, section 5 summarises the findings and presents conclusions.

2. Plain language: from official communication to contracts

As defined by the International Plain Language Federation, "[a] communication is in plain language if its wording, structure, and design are so clear that the intended

¹ Results of blind tests conducted in 2021 with the participation of professional translators, <https://www.deepl.com/en/whydeepl>, for more information see <https://translatepress.com/deepl-translator-review/>

readers can easily find what they need, understand what they find, and use that information”.² The ISO standard on plain language adds to this definition that “plain language focuses on how successfully readers can use the document rather than on mechanical measures such as readability formulas”, that it “saves time or money (or both) for readers and organizations”, and even that “the process of translating is more efficient for plain language documents than for documents that are difficult to understand”.³ The last claim is what this article tries to examine in the context of contracts.

Legal English is often referred to as legalese. The name suggests a negative assessment and indeed, it is described as “a peculiarly obscure and convoluted variety of English”⁴ or even “a form of prose so jumbled, dense, verbose, and overloaded that it confuses and frustrates most everyday readers and even many lawyers”⁵. Its syntactic complexity was even found to be harder for readers to process than low-frequency vocabulary.⁶ So plain language and legalese can be seen as the opposite ends of a spectrum.

2.1. Plain English

Plain language was in fact a response to critiques of the level of difficulty or incomprehensibility of official documents. In his 1940 *Brevity* memo⁷, Winston Churchill complained about the time and effort needed to find information in reports, asking for “short, conversational phrases”, no “woolly phrases” or “padding”, as well as better organisation of documents – with headings, shorter paragraphs and detailed information in appendices rather than in the body. In the US, John O’Hayre published stories showing how members of the administration struggled to understand documents they worked with.⁸ Legal language and bad legal drafting were criticised even earlier, including by lawyers themselves⁹, and when given a choice between

² <https://www.iplfederation.org/plain-language/>

³ ISO 24495-1:2023 *Plain language – Part 1: Governing principles and guidelines*, <https://www.iso.org/obp/ui/en/#iso:std:iso:24495:-1:ed-1:vl:en:ref:3>

⁴ María José MARÍN: Legalese as seen through the lens of corpus linguistics: an introduction to software tools for terminological analysis. *International Journal of Language & Law (JLL)*, vol. 6., August (2017) 19.

⁵ Joseph KIMBLE: *Lifting the Fog of Legalese. Essays on Plain Language*. Durham, Carolina Academic Press, 2006. XII.

⁶ Eric MARTÍNEZ – Francis MOLLICA – Edward GIBSON: Poor writing, not specialized concepts, drives processing difficulty in legal language. *Cognition*, vol. 224. (2022) 105070. <https://doi.org/10.1016/j.cognition.2022.105070>

⁷ Available at: bit.ly/44TS0pk; Churchill’s other calls for brevity reported at: <https://blog.nationalarchives.gov.uk/churchills-call-for-brevity/>

⁸ John O’HAYRE: *Gobbledygook Has Gotta Go*. U.S. Bureau of Land Management, 1966.

⁹ Tom GOLDSTEIN – Jethro K. LIEBERMAN: *The Lawyer’s Guide to Writing Well* (3rd ed.). Oakland, University of California Press, 2016. 3–4., 15–19.; KIMBLE op. cit. 175–179.

legalese and plain language versions of the same documents, a majority opted for plain language.¹⁰

In the 1980s, the efforts to introduce plain language in the UK focused on forms, leaflets and standard letters used by the government and local authorities. Simplified official forms were found to generate savings.¹¹ Several insurance companies followed suit with plain English insurance policies and discovered that not only customers but also their own staff found them “easier to work with”.¹² The next step was regulations aiming to ensure availability of certain information in consumer credit agreements to make it “easily legible”. Even before they entered into force, the Minister for Consumer Affairs urged trade associations and credit companies to go further and make such information easy to understand too.¹³

The analysis of several standard contracts used by trade associations, undertaken by the Plain English Campaign, revealed that they were slightly or much less readable than the least readable newspapers and much less readable than popular newspapers. A more fine-grained analysis of the language used in contracts produced a list of five major difficulties:

- vocabulary that did not appear in everyday speech (e.g., *hereinafter*, *aforsaid*, *notwithstanding*, *to determine* meaning ‘to terminate’, *to distraint*, *lien*, etc.);
- long and complex sentences (often with 100 words) with many clauses, as opposed to the standard of 15-20 words in popular newspapers;
- passive voice and avoiding you/we;
- “lack of normal punctuation”, especially commas;
- cross-references to other parts of the same document and to laws.¹⁴

The authors also criticized, citing examples, two related phenomena:

- repetition (up to ~~but not exceeding~~ the cost of...; full ~~and absolute~~ authority); and
- redundancy (strictly in accordance with the terms of this contract ~~without deduction~~; any loss or damage ~~occasioned~~).¹⁵

In the US, the idea of plain language received much publicity thanks to the plain language promissory note introduced by Citibank in 1975. The bank’s team “stripped the prior version, a dense and essentially unreadable document, of many substantive provisions and cleansed the remaining verbiage”, sparking interest in rephrasing

¹⁰ KIMBLE op. cit. 7–8.

¹¹ <https://www.plainenglish.co.uk/about-us/history/timeline.html>

¹² Martin CUTTS – Chrissie MAHER: *Small Print: The Language and Layout of Consumer Contracts*. National Consumer Council, 1983. 2.

¹³ Ibid. 4–5.

¹⁴ Ibid. 5.

¹⁵ Ibid. 11.

contracts to be clearer.¹⁶ After a period of federal-level initiatives in the field of plain language being promoted and withdrawn (though often continued at state level and promoted by law professors), there has been steady progress in its implementation since the late 1990s, when federal administration was requested to write plainly (Figure 1). Plain language subsequently became mandatory in legislation, reports or on websites thanks to the Plain Writing Act of 2010 and several Executive Orders. Importantly, in both UK and US, glossaries¹⁷ were prepared with explanations of and suggested replacements for more formal words and legal terms, including Latin expressions and words of Latin origin.

Washington
June 1, 1998

MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

SUBJECT: Plain Language in Government Writing

The Vice President and I have made reinventing the Federal Government a top priority of my Administration. We are determined to make the Government more responsive, accessible, and understandable in its communications with the public.

The Federal Government's writing must be in plain language. By using plain language, we send a clear message about what the Government is doing, what it requires, and what services it offers. Plain language saves the Government and the private sector time, effort, and money. Plain language requirements vary from one document to another, depending on the intended audience. Plain language documents have logical organization, easy-to-read design features, and use:

- common, everyday words, except for necessary technical terms;
- "you" and other pronouns;
- the active voice; and
- short sentences.

To ensure the use of plain language, I direct you to do the following:

- By October 1, 1998, use plain language in all new documents, other than regulations, that explain how to obtain a benefit or service or how to comply with a requirement you administer or enforce. For example, these documents may include letters, forms, notices, and instructions. By January 1, 2002, all such documents created prior to October 1, 1998 must also be in plain language.
- By January 1, 1999, use plain language in all proposed and final rulemakings published in the Federal Register, unless you proposed the rule before that date. You should consider rewriting existing regulations in plain language when you have the opportunity and resources to do so.

The National Partnership for Reinventing Government will issue guidance to help you comply with these directives and to explain more fully the elements of plain language. You should also use customer feedback and common sense to guide your plain language efforts.

I ask the independent agencies to comply with these directives. This memorandum does not confer any right or benefit enforceable by law against the United States or its representatives. The Director of the Office of Management and Budget will publish this memorandum in the Federal Register.



Figure 1. President Clinton's 1998 memorandum requiring administration to use plain language

¹⁶ Carl FELSENFELD: The Plain English Movement in the United States. *Canadian Business Law Journal*, no. 6 (1981–1982), 409.

¹⁷ Plain English Campaign prepared *The A to Z guide to legal phrases* <https://tinyurl.com/4wx7fhxn>, while a list of alternatives is presented in KIMBLE op. cit. 165–174.

2.2. Plain Polish

Plain Polish remained no more than a topic of discussions and debates for a long time,¹⁸ echoing similar debates in English-speaking countries.¹⁹ The 2002 Regulation on the Principles of Legislative Technique introduced a requirement to use clear, communicative, and adequate language in legislation, but left the assessment to drafters themselves. The 2014 Act on Consumer Rights went a step further and provided examples of information for customers in understandable language.

It was not until around 2010 that any practical recommendations were prepared and any efforts at their implementation made. The need for change became apparent when incomprehensibility of documents on EU funds was blamed for their poor uptake in Poland. Thanks to efforts of the Plain Polish Lab (PPL) and the willingness of local and central authorities, official communication became much more accessible. Banks and insurers, like their UK and US counterparts in the past, embraced plain language as a device that saves time and money. It was banks that made a commitment to introduce plain Polish in communication with their customers and developed the first contracts in plain Polish together with PPL experts.²⁰

In defense of Polish contracts, the vocabulary is not as different from everyday language as certain lexical items from legalese compared to general English. There are no pronominal adverbs like the English *hereinafter*, *thereof*, *whereupon*, etc., doublets/triplets are not a standard feature, and – due to different drafting conventions²¹ – there is no need to list all possible cases to make contracts self-sufficient²², all-encompassing, and precise²³, so there are no strings of quasi-synonyms.

The language of Polish contracts is often based on provisions of the 1964 Civil Code, which is not archaic yet, even if not very recent. Importantly, the Civil Code applies to any matters not regulated in contracts or replaces contractual provisions that are inconsistent with it (Article 58(1) of the Code), while interpretation of contracts cannot be based on the explicit wording only but should take into account the ‘parties’ intentions and the aim of the contract (Article 65(2)). This certainly

¹⁸ For a summary in legislative context, see Marta ANDRUSZKIEWICZ: Problem jasności w języku prawnym – aspekty lingwistyczne i teoretycznoprawne. [The problem of clarity in legal language – linguistic and legal theoretical aspects] *Comparative Legilinguistics*, vol. 31. (2017) 7–25. <http://dx.doi.org/10.14746/cl.2017.31.1>

¹⁹ Joseph KIMBLE: A Curious Criticism of Plain Language. *Legal Communications and Rhetoric: JALWD*, vol. 13. (2016) 181–192.

²⁰ For a comparison of readability indices of a plain bank account contract and typical Polish contracts, see Anna SETKOWICZ-RYSZKA: Why can plain English in contracts cause difficulties in translation into Polish? *Lingua Legis*, vol. 30 (2022), 53–57.

²¹ Distinction between “concise” drafting style in civil law countries and “precise” style in common law is discussed in Deborah CAO: *Translating Law*. Clevedon, Multilingual Matters, 2007. 28–29.

²² Giuditta CORDERO-MOSS (ed.): *Boilerplate Clauses, International Commercial Contracts and the Applicable Law*. Cambridge University Press, 2011. 116.

²³ Christopher WILLIAMS: Legal English and Plain Language: an introduction. *ESP Across Cultures*, no. 1. (2004) 111–124., 121.

reduces the responsibility compared to that of drafters of common law contracts, who need to ensure that the contract contains all information needed to understand its meaning²⁴, so old expressions that stood out in court are preferred as safer than untested plain English alternatives.²⁵

That said, there is still much room for improvement in the clarity of Polish contracts. The advice from plain Polish experts as to the choice of linguistic means includes avoiding passive voice, nominalizations, impersonal forms, long words, and complex sentences, while providing paraphrases of specialist terms, using everyday words, addressing the reader directly, as well as making sure conditional sentences start with the condition (“if... then...”).²⁶

3. Using MT to translate legal texts

The ISO defines post-editing (PE) as “edit[ing] and correct[ing] machine translation output”, either “to obtain a product comparable to a product obtained by human translation” (full PE) or “to obtain a merely comprehensible text without any attempt to produce a product comparable to a product obtained by human translation” (light PE).²⁷ It is stressed that MT errors differ from human errors. Frequent MT errors include transfer errors (omissions, additions, distortions), wrong terminology, lack of terminological consistency and textual coherence, as well as inadequate – too colloquial or too formal – register and wrong forms of address.²⁸

In general, contracts are repetitive texts, so MT of contracts – especially the standard provisions – should be fairly successful. Back in 2014, a study on the MT of four contracts (in the PL-EN direction) showed that although approximately 90% of sentences needed PE, the share of nonsense sentences was low – 8% in total - with the highest share of nonsense sentences in a contract type without an English equivalent: *umowa o dzieło* [contract for a specific task]. Most errors concerned wrong choice of words or phrases – including imprecise legal terms – or missing words or phrases. Unsurprisingly, terminological incongruity between the two legal systems was not addressed in MT output.²⁹ Yet, more recent literature on post-editing neural MT

²⁴ Peter M. TIERSMA: *Parchment, Paper, Pixels: Law and the Technologies of Communication*. University of Chicago Press, 2010. 126.

²⁵ WILLIAMS op. cit. 118–120.

²⁶ Tomasz PIEKOT – Grzegorz ZARZECZNY – Ewelina MOROŃ: Standard ‘plain language’ w polskiej sferze publicznej. [Plain language’ standard in the Polish public sphere] In: Monika ZAŚKO-ZIELIŃSKA – Krzysztof KREDENS (ed.): *Lingwistyka kryminalistyczna. Teoria i praktyka*. [Forensic Linguistics. Theory and practice.] Wrocław, Quaestio, 2019. 202–203.

²⁷ ISO 18587:2017 *Translation services – Post-editing of machine translation output – Requirements*, definitions available at: <https://www.iso.org/obp/ui/en/#iso:std:iso:18587:ed-1:vl:en>

²⁸ Łucja BIEL: Postedycja tłumaczeń maszynowych. [Post-dispatch of machine translations.] *Lingua Legis*, vol. 29. (2021) 11–34., 20., 25.

²⁹ Joanna SYCZ-OPOŃ: Machine Translation – Can It Assist in Professional Translation of Contracts? *International Journal of Legal Discourse*, vol. 20. (2014) 181–200.

(NMT) output in other language pairs stresses that post-editing machine-translated legal documents leads to gains in both quality of the translations and productivity.³⁰

However, in EN-PL language pair, an evaluation of the effort needed to correct NMT output in the context of European Union texts revealed that most mistakes concerned accuracy and terminology, including inconsistent use of terminology, though a modest productivity gain was reported.³¹ Also for other languages (e.g., Slovene, Dutch), EU translators using NMT complained about “polysemic misrepresentations” leading to terminology errors, “complete semantic blunders”, “neural neologisms”, problems with proper names, abbreviations, and even additions or omissions that sometimes changed the meaning to the opposite.³² That seems to confirm findings about accuracy errors being more demanding to correct than fluency errors.³³ It should also be noted that Polish is an inflectional language, so a change of one element in a sentence can require changes to other elements to maintain agreement.

4. Features of plain English contracts in MT

This section seeks to analyse how successful DeepL MT engine was in translating in March 2022 and March 2023 a contract which, although not fully adherent to the rules of plain English, is closer to plain English than to legalese. It is a standard contract of supply entered into via a UK website which matched wood pellet buyers and suppliers, governed by Scottish law. (The website no longer exists, the contract was retrieved for teaching purposes in 2019.) The Gunning Fog index for a sample of this contract is 9.3 on a scale from 0 to 20, the Flesch reading ease index is 65.1, so it should be easily understood by 13- to 15-year-old students, while the average sentence length is 15 words.³⁴ Naturally, readability indices are a crude measure of how plain a text is, as they focus on surface-level features, such as sentence or

³⁰ See special issue of *Revista de Llengua i Dret, Journal of Language and Law*, vol. 78. (2022), including Jeffrey KILLMAN – Mónica RODRÍGUEZ-CASTRO: Post-editing vs. translating in the legal context: Quality and time effects from English to Spanish. *Revista de Llengua i Dret, Journal of Language and Law*, 78, 56–72. <https://www.doi.org/10.2436/rld.i78.2022.3831> ; Vilemini SOSONI – John O’SHEA – Maria STASMIOTI: Translating law: A comparison of human and post-edited translations from Greek to English. *Revista de Llengua i Dret, Journal of Language and Law*, 78, 92–120. <https://doi.org/10.2436/rld.i78.2022.3704>

³¹ Karolina STEFANIAK; Evaluating the usefulness of neural machine translation for the Polish translators in the European Commission. *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation*, 2020. 263–269.

³² Mateja ARNEJŠEK – Alenka UNK: Multidimensional assessment of the eTranslation output for English-Slovene. *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation*, 2020. 383–392. 386.

³³ For an overview see Valentina RAGNI – Lucas NUNES VIEIRA: What has changed with neural machine translation? A critical review of human factors. *Perspectives*, vol. 30., no. 1. (2022) 137–158. 142–149.

³⁴ SETKOWICZ-RYSZKA op. cit. 53. Measures from free online tools: <https://tinyurl.com/4w8rmff2>

word length, while neglecting textual coherence or content.³⁵ A closer examination revealed that there are almost no binomials (except for *any and all* twice), no quasi-synonyms (except for *valid and enforceable* and *enforce or execute*, each used once), and no pronominal adverbs (except for one occurrence of *hereof*). The whole 3,426-word contract was analysed using memoQ Live Docs feature, which enables creating a parallel corpus.³⁶ The “find” function enables finding the relevant features and translation solutions can be examined manually.

The following subsections focus on Polish translations of plain English features listed in President Clinton’s memorandum, namely passages where “common, everyday words” are mixed with legal or technical terminology and phraseology (4.1 and 4.2), pronouns instead of names of parties (4.3). Due to the importance of linguistic expressions of obligation, prohibition or permission in contracts, modal verbs and other markers of modality are also analysed (4.4), while other, mostly unproblematic, plain language features are discussed jointly (4.5). Original spelling from ST is retained, while backtranslations are provided in square brackets.

4.1. “Common, everyday words, except for necessary technical terms” – scenario 1: legal language

From the point of view of keeping the right register the fact that colloquial expressions appear next to formal ones seems the main difficulty for a translator dealing with a plain(er) English original. Given the scarcity of plain contracts in Polish and their concentration in banking, there is little reference material for translating a contract of supply. In DeepL renditions of mixed passages where legal terms appear into Polish there are no colloquial expressions at all, so the style is closer to the English legalese. The more colloquial English expressions in the examples below, such as:

- *do not pay* [formal equivalent: *fail to pay*],
- *to chase the payment* [*to collect the payment*],
- *in line with* [*in accordance with/pursuant to*],
- *to change* [*to amend*],
- *as they relate to* [*to the extent they concern*],
- *You are not happy with* [*you disagree/do not consent*]

are changed to a more formal style, so the effort made by the original drafters to include informal expressions where possible is lost. Interestingly, certain Polish expressions, like *w przypadku* [*in the event of*], must be followed by a noun, which increases the degree of formality with a nominalization³⁷ used where the ST may contain a verb, as in example (1). Apart from nominalizations, formality is increased by the use of words like *niniejsza*, *powiadomić*, *wszelkie*, *skutkować*, all of which have

³⁵ Kathy CONKLIN – Richard HYLE – Fabio PARENTE: Assessing plain and intelligible language in the Consumer Rights Act: a role for reading scores? *Legal Studies*, vol. 39. (2018) 378– 397. <http://dx.doi.org/10.1017/lst.2018.25>

³⁶ <https://docs.memoq.com/current/en/Places/resconsole-livedocs.html>

³⁷ Additional nominalizations are italicized.

less formal equivalents. Moreover, examples (2) and (3) also demonstrate confusion with ‘parties’ names (see subsection 4.3).

- (1) ST: In the event that You do not pay any due sums via XXX, Supplier reserves the right to chase the payment directly.

DeepL-2022, DeepL-2023: W przypadku *nieuiszczenia* przez Klienta należnych kwot za pośrednictwem XXX, Dostawca zastrzega sobie prawo do bezpośredniego *dochodzenia* zapłaty.

[In the event of non-payment by the Client of due sums via XXX, Supplier reserves the right of direct collection of the payment.]

- (2) ST: You may not withhold payment in the event of a dispute and any late payments may accrue interest and other charges in line with government late payment guidelines.

DeepL-2022: W przypadku sporu Dostawca nie może wstrzymać płatności, a wszelkie opóźnienia w płatnościach mogą skutkować *naliczeniem* odsetek i innych opłat zgodnie z rządowymi wytycznymi.

[In the event of a dispute Supplier may not suspend payment and any delay in payment may result in the accrual of interest and other charges in accordance with governmental guidelines.]

DeepL-2023: Nie wolno wstrzymywać płatności w przypadku sporu, a od wszelkich opóźnionych płatności mogą być naliczane odsetki i inne opłaty zgodnie z rządowymi wytycznymi dotyczącymi opóźnień w płatnościach.

[It is not permitted to suspend payment in the event of a dispute and on any late payments there may be accrued interest and other charges in accordance with governmental guidelines concerning delays in payments.]

- (3) ST: If Supplier change this Contract of Supply as they relate to your Delivery, Supplier will give you reasonable advance notice of the changes and provide details of how to cancel if You are not happy with them.

DeepL-2022: Jeżeli Dostawca zmieni niniejszą Umowę Dostawy w zakresie dotyczącym Dostawy Dostawcy, powiadomi Dostawcę o zmianach z odpowiednim wyprzedzeniem i poda szczegóły dotyczące *sposobu odstąpienia* od Umowy, jeżeli Dostawcy nie odpowiadają te zmiany.

[If Supplier amends this Contract of Supply to the extent concerning Supplier’s Supply, it will notify Supplier about the amendments appropriately in advance and provide details concerning the manner of withdrawal from the Contract if the amendments do not suit Supplier.]

DeepL-2023: Jeśli Dostawca zmieni niniejszą Umowę Dostawy w zakresie dotyczącym Dostawy, powiadomi o zmianach z odpowiednim wyprzedzeniem i poda szczegóły dotyczące *sposobu odstąpienia* od Umowy, jeśli nie będą one satysfakcjonujące dla Dostawcy.

[If Supplier amends this Contract of Supply to the extent concerning the Supply, it will notify about the amendments appropriately in advance and provide details concerning the manner of withdrawal from the Contract if they are not satisfactory for Supplier.]

As observed by Sycz-Opoń, system-bound legal terms, like *waiver of breach*, *reasonable*, *legal duty of care* or *fraudulent misrepresentation*, are not provided with explanations for recipients from a different legal system or accompanied by the original terms in brackets. They are translated literally, which may cause misunderstandings, especially with *waiver of breach*, which should be explained as, e.g., *waiver of the rights available to a party in case of the other party's breach of contract*. This does not concern plain language features in the ST, but since plain language is also concerned with substance, lack of explanations makes the translation more difficult to understand.

4.2. “Common, everyday words, except for necessary technical terms” – scenario 2: technical language

What happens with fragments where there appear technical terms concerning wood pellet delivery is quite different. In translation into Polish, the language remains less formal and resembles that used in instructions (which may explain the word *User*, discussed in the next subsection). It is in these fragments that verbs in 2nd person singular appear most often in the 2022 rendition. The following examples illustrate this (and inconsistent translations of *Wood Pellets*, italicized).

- (4) ST: If Wood Pellets are left in the delivery pipe when Supplier is unable to blow any more into your store, Supplier will have to clear the pellets from the pipe, onto the ground where the pipes lay, if you do not provide a more suitable receptacle.
 DeepL-2022: Jeżeli *peletki drzewne* pozostaną w rurze dostawczej, a Dostawca nie będzie w stanie wdmuchać ich do magazynu, Dostawca będzie musiał usunąć *peletki* z rury na ziemię, na której leżą rury, jeżeli nie dostarczysz innego odpowiedniego pojemnika. [backtranslation not needed, *do not provide* retained in 2nd person singular]
 DeepL-2023: Jeśli *pelety drzewne* pozostaną w rurze dostawczej, gdy Dostawca nie będzie w stanie wdmuchiwać ich więcej do Państwa magazynu, Dostawca będzie musiał usunąć *pelety* z rury na ziemię, na której leżą rury, jeśli nie zapewnią Państwo bardziej odpowiedniego pojemnika. [backtranslation not needed, *do not provide* changed to 3rd person plural with the honorific *Państwo*]
- (5) ST: It will be Your responsibility to dispose of these pellets, at your cost.
 DeepL-2022: Klient będzie odpowiedzialny za pozbycie się tych *peletek* na własny koszt.
 DeepL-2023: Klient będzie odpowiedzialny za pozbycie się tego *granulatu* na swój koszt.
 [both versions: Client will be responsible for getting rid of these pellets at their own cost.]
- (6) ST: You must provide your own means of moving the bags from kerbside to where they will be stored.
 DeepL-2022: Użytkownik musi zapewnić sobie własny środek do przenoszenia worków z krawężnika do miejsca ich składowania.

[User must provide their own means for carrying bags from the kerb to the place of their storage.]

DeepL-2023: Należy zapewnić sobie środki do przenoszenia worków z krawężnika do miejsca ich składowania.

[One should provide one's own means for carrying bags from the kerb to the place of their storage.]

4.3. You and other pronouns

For many types of contracts regulated in the Polish Civil Code, the names of parties are mostly standardized. For instance, in *umowa dostawy* [contract of supply], the parties are called *odbiorca* [recipient] and *dostawca* [supplier], except for cases where pronouns, especially *it* are used to avoid repeating the party's name. This influences verb forms: verbs are in 3rd person, singular or plural. There are already some contracts where 'we' and 'you' are used to refer to parties, but most of them relate to banking products. Therefore, translating *you* requires human translators to adopt a top-down strategy and either use the parties' names as in the Civil Code, or *ty* [*you*] needs to be retained. Somewhat in between the two solutions, it is possible to use honorific forms of address with a verb in 3rd person singular – for *Pan/Pani* – or plural – for *Państwo*. This form of address is used in the Schedule to the Act on Consumer Rights of 2014,³⁸ containing examples of information about consumers' right to withdraw from a contract of sale.

The ST is mixed in that one of the parties is referred to as *Supplier*, while the other as *you*. The pronoun *we* is probably not used to avoid confusion, because the contract is made via an intermediary. In both DeepL versions, *Supplier* is translated consistently and correctly as *Dostawca*. What happens with *you* is that its direct equivalent, *ty*, appears in the definitions and some other provisions in both 2022 and 2023 versions, while the honorific *Państwo* features quite often in the 2023 version. We also find various parties' names as translations of *you*. Some are safe general equivalents, like *Klient* [Client] or *Kontrahent* [Counterparty], some are wrong, like *Subskrybent* [Subscriber], *Dystrybutor* [Distributor] or *Użytkownik* [User]. There are even cases of *you* mistranslated as *Dostawca* [Supplier] (see examples (2) and (3) in subsection 4.1). The only correct name of the party receiving the supply – *Odbiorca* – appears in the 2022 version, but as a translation of *Customer* in the ST. Additionally, in Polish, the pronoun can be omitted, as the verb form alone denotes it. Verbs in 2nd person singular appear in both MT versions.

The full set of solutions and numbers of instances they are used is provided in Table 1. Only nominative forms of the names are provided, but they appear in the appropriate cases in the text. Instances when *you* appears in the same sentence as Supplier and without it are grouped separately, to see what happens when a contextual cue is available or not. Judging by various names used to refer to that party, this

³⁸ Kancelaria Sejmu: USTAWA z dnia 30 maja 2014 r. o prawach konsumenta:
<https://isap.sejm.gov.pl/isap.nsf/download.xsp/WDU20140000827/U/D20140827Lj.pdf>

cue is not decoded. Different total numbers of solutions result from the fact that sometimes the part containing *you* can be dropped, like when *the terms under which Supplier will supply and deliver Wood Pellets to You* are translated as *warunki, na jakich Dostawca będzie dostarczał Pellet drzewny* [terms under which Supplier will supply Wood Pellets]. Some sentences were reformulated to contain the adjective *własny* or pronoun *swój* [both meaning *one's own*], eliminating the need to choose any equivalent.

Table 1. Renditions of pronoun 'you' in 2022 and 2023 DeepL translations of the contract of supply

Type	Example(s)	DeepL-2022	DeepL-2023
Definition	“You/Your/Yours” – Means the person or company...	“Ty/Twoja/Twojego” - oznacza osobę lub firmę	“Ty/Twoja/Twój” - Oznacza osobę lub firmę...
You/your +Supplier	“The price payable by You to the Supplier will be the price as set out on [...] website...” “If your Order Confirmation does not include a delivery date Supplier will schedule your delivery...”	Ty (4) Twoje (1) Verb in 2 nd sg (6) - - *Dostawca (4) *Dystrybutor (1) Klient (29) Kontrahent (6) *Użytkownik (5)	Ty (2) Twoje (1) Verb in 2 nd sg (4) Verb in 2 nd pl (1) Państwo (11) *Dostawca (1) - Klient (23) Kontrahent (1) *Użytkownik (6)
You/your -Supplier	“In such an event You would be liable for a reasonable failed delivery fee.” “This provision does not affect your other statutory rights as a consumer.”	Ty (1) Twoje (5) Verb in 2 nd sg (5) - Klient (8) Kontrahent (2) - *Użytkownik (5)	Ty (4) Twoje (5) Verb in 2 nd sg (10) Państwo (3) Klient (6) - *Subskrybent (2) *Użytkownik (5)

* non-recommended solutions

The diverse solutions would make post-editing these translations challenging: even if most are acceptable, they cannot be used interchangeably in the same text and any change of the name may entail other changes, especially in verb forms. Counterintuitively, a simple pronoun causes major difficulty in MT when used in a type of text where pronouns rarely appear in the target legal culture.

4.4. Deontic modality markers

Modal verbs and other constructions expressing deontic modality (obligation, permission, prohibition) are an important feature of texts which establish rights and impose obligations, including contracts. In English this was typically achieved by modal verbs: *shall*, *must*, *can* and *may*, with *not* where necessary, or non-modal verbs

to agree or *to undertake* followed by a verb can also be used to express obligation.³⁹ Supporters of plain legal language criticized especially *shall* as a confusing and often misused verb and suggested replacing it with *must*.⁴⁰ One of the victories of plain language movements may be *must* in statutes enacted in the UK now.

The Polish linguistic means used to express deontic modality are completely different: the main ways are verbs in present or future tense indicative, and constructions such as *być (z)obowiązany do* [be obliged to], while the verb *musieć* [must] is rare in legislation or contracts. Permission is often expressed by the verb *móc* [can/may] in the present or future tense, though other constructions are possible, e.g., *mieć prawo* [to have the right].⁴¹

In the ST, the most frequent modal verbs are *may*, *shall* and *can*, but *must* appears as well, along with several other expressions. The distribution of deontic modality markers and their Polish renditions by DeepL is presented in Table 2. Less felicitous solutions are marked with asterisks, but they are not errors.

Table 2. Renditions of deontic modality markers in 2022 and 2023 DeepL translations of the contract of supply

Deontic modality marker	DeepL-2022	DeepL-2023
may (15)	<i>móc</i> in present tense (15)	<i>móc</i> in present tense (15)
shall (12)	present tense of the main verb (8) future tense of the main verb (4)	present tense of the main verb (4) future tense of the main verb (8)
can (5)	<i>móc</i> in present tense (3) <i>można</i> – impersonal form of <i>móc</i> (1) <i>możliwość</i> [possibility] (1)	<i>móc</i> in present tense (3) <i>można</i> – impersonal form of <i>móc</i> (1) <i>możliwość</i> [possibility] (1)
must (3)	<i>musieć</i> in present tense (2) <i>mieć obowiązek</i> [have a duty] (1) -	<i>musieć</i> in present tense (1) - <i>*należy</i> – modal verb, impersonal form of <i>should</i> (2)
have the right to + verb (4)	<i>mieć prawo</i> in present tense + verb (1) <i>mieć prawo</i> in present tense + deverbal noun (3)	<i>mieć prawo</i> in present tense + verb (2) <i>mieć prawo</i> in present tense + deverbal noun (2)
agree (3)	<i>zobowiązać się</i> in present tense (2) <i>*zgadzać się</i> in present tense (1)	<i>zobowiązać się</i> in present tense (2) <i>*zgadzać się</i> in present tense (1)
would be liable (1)	<i>być zobowiązany</i> in future tense (1)	<i>być zobowiązany</i> in future tense (1)

³⁹ Aleksandra MATULEWSKA: Deontic modality and modals in the language of contracts. *Comparative Legilinguistics*, vol. 2. (2010) 75–92.

⁴⁰ KIMBLE (2006) op. cit. 42.; Natalia ZYCH: Idea plain language a teksty prawne. *Przegląd Legislacyjny*, no. 3. (97) (2016), 65–90., 74.

⁴¹ MATULEWSKA op. cit.; Łucja BIEL: The textual fit of translated EU law: a corpus-based study of deontic modality. *The Translator*, vol. 20., no. 3. (2014) 332–355.
<http://dx.doi.org/10.1080/13556509.2014.909675>

cannot (1)	<i>nie można</i> – impersonal negated form (1)	<i>nie może</i> – finite negated form (1)
can not be held liable (1)	<i>ponosić odpowiedzialność</i> [be liable] in present tense + negation (1)	<i>ponosić odpowiedzialność</i> [be liable] in present tense + negation (1)
may not (1)	<i>móc</i> in present tense + negation (1)	* <i>nie wolno</i> – impersonal form + negation (1)
will have to (1)	* <i>musieć</i> in future tense (1)	* <i>musieć</i> in future tense (1)
is not obliged to (1)	<i>być zobowiązany</i> in future tense (1)	<i>być zobowiązany</i> in future tense (1)
will be under no obligation (1)	<i>być zobowiązany</i> in future tense + negation (1)	<i>być zobowiązany</i> in future tense + negation (1)

There are also numerous instances of the future tense, most of which are rendered as the future tense in both Polish versions, with two exceptions – both in the same places in the 2022 and 2023 renditions. The first one is example (6) in subsection 4.2. In both Polish versions a nominalization – is introduced. In the other one, the adverb *usually* seems to cause the use of the present tense:

ST: The Delivery Date will usually be within the window of dates in Your Order Request...

DeepL-2022: Data Dostawy zwykle mieści się w przedziale dat podanym w Zamówieniu Klienta...

[The Delivery Date is usually within the date range stated in the Client's Order...]

DeepL-2023: Data Dostawy zazwyczaj mieści się w przedziale dat podanych w Zamówieniu...

[same as above, but: stated in the Order...]

Summarizing, what happens with deontic modality markers in both Polish translation by DeepL is quite typical for Polish contracts and, as can be seen in Table 2, there are small differences between them. Unlike parties' names, deontic modality markers need not be standardized, so replacing *shall* with other modal verbs or expressions in the ST would not make post-editing very demanding.

4.5. Other plain language features

This section deals with the active voice, nominalizations, and short sentences. The active or passive voice is retained in most cases and the only changes are those required by rules of the target language. As an example, in English a *contract is governed* by a certain law, while in Polish the active voice is used: *umowa podlega prawu...* [*contract submits to the law of...], or someone *is held liable* in English, while in Polish this person *ponosi odpowiedzialność* [bears responsibility]. Sometimes, an English passive construction is rendered differently in the two MT versions: *Payment will be deemed to have been received* was translated in 2022 as *Uznaje się, że płatność została otrzymana* [*One regards that payment has been received], so the

passive form became an impersonal one, while in 2023 as *Płatność będzie uważana za otrzymaną* [Payment will be considered as received], so the passive form was retained.

Less adherence to plain language rules can be seen in the translations of infinitives, which are often nominalized. For example, the phrase *has the right to cancel the Delivery* becomes *ma prawo do anulowania Dostawy* [has the right of cancellation of the Delivery] in two out of three instances in the 2022 version and in all three instances in the 2023 version, even though in this context an infinitive can be used in Polish. Throughout the contract, many verbs, both infinitives and finite forms are replaced with deverbal nouns, including deverbal nouns (gerunds), which is a typical feature of the Polish official register. Some nominalizations were also marked in earlier examples.

A special case of nominalization is when the verb appears in a heading. Headings are not a key part of a contract (and that they do not influence its interpretation), but they are an important feature of the layout. The difficulty in dealing with plain English lies with the preference for finite verb forms rather than deverbal nouns. In Polish contracts, headings usually contain nouns, not sentence-like structures. The examples provided below show that both DeepL renditions are very similar, though with only two examples no generalizations can be made.

The first heading like that is *How the contract is formed between You and Supplier*, translated with a finite verb form in both versions: *W jaki sposób powstaje umowa pomiędzy Klientem a Dostawcą* (DeepL-2022) and *Jak powstaje umowa pomiędzy Użytkownikiem a Dostawcą* (DeepL-2023). Both verbs are in the active voice and can be back translated as **How the contract emerges between the Client/User and Supplier*. This structure is more typical of manuals or instructions, and I would use a deverbal noun in translation of a contract. The second interesting heading is *Cancelling if You Change Your Mind*, which is nominalized: *if you change your mind becomes w przypadku zmiany zdania* [in the case of a change of mind] in both DeepL versions.

Finally, short sentences are not a problem in MT, as the text is segmented and it is only human translators that might be tempted to join sentences to make them longer, like in typical contracts. Again, it is rather the register that causes problems. The sentence *It is your responsibility to ensure Delivery can be made* was translated in 2022 as *Obowiązkiem Klienta jest upewnienie się, że dostawa może zostać zrealizowana* [The duty of Client is making sure that the delivery can be made], so an extra nominalization was introduced, and in 2023 as *Na Tobie spoczywa odpowiedzialność za zapewnienie możliwości dostawy* [**On You rests the responsibility for ensuring the possibility of delivery*], with two deverbal nouns added. At times, a sentence grows longer when a single word is translated as a multi-word expression, like in:

ST: You A_ *acknowledge* that B_ *airborne dust* may be created during Delivery.

DeepL-2022, DeepL-2023: Klient A_ *przyjmuje do wiadomości*, że podczas dostawy może powstać B_ *pył unoszący się w powietrzu*.

As demonstrated by the above examples, verbs in the active voice or short sentences are not problematic in MT. It is rather the register that often shifts towards more formal and heavier, probably in keeping with the register of typical contracts used to train the engine. Interestingly, this tendency is stronger in the 2023 version, despite the gradual progress of plain Polish in general.

5. Conclusions

As we have seen, some features of plain English contracts cause problems in MT (and subsequent PE) and others do not. The latter group includes especially markers of modality, the active voice, and short sentences. The surprising top difficulty in MT is the pronoun *you* referring to one party. The mixing of colloquial expressions with legal or technical vocabulary results in different registers in different parts of the contract. Post-editing of both versions would involve levelling out the patchwork of registers – from legalese to the language of user manuals to colloquial language – and references to the customer – from informal to polite forms of address to completely impersonal constructions. Even if individual segments are acceptable in isolation, they may be unacceptable in context⁴² and part of translation competence in the MT era is making translations compiled from various sources read like coherent texts.⁴³ The analysis confirms some of the findings concerning typical MT errors, including in legal texts. Although this analysis was limited to the English-Polish language pair, some of the problems described above can be expected to appear in translation into other languages with less developed plain registers.

Naturally, there is more to plain language than just register and certain linguistic devices. Legal texts are not read for pleasure but in order to quickly obtain information, so intelligibility is the paramount concern.⁴⁴ Text layout and content, including the order in which information is provided, matter a lot.⁴⁵ One cannot expect a plain language contract from one legal culture to become easily accessible in another just because its drafters refrained from using words that are difficult in the source legal culture. Besides, plain language is never the product of a single person: it is developed through cooperation or negotiation between various experts, and tested

⁴² Sheila CASTILHO: *Context-aware MT evaluation: what have we learned?* 2023, talk available at bit.ly/47073zC; Elena VOITA – Rico SENNRICH – Ivan TITOV (2019). When a Good Translation is Wrong in Context: Context-Aware Machine Translation Improves on Deixis, Ellipsis, and Lexical Cohesion. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, (2019) 1198–1212. <http://dx.doi.org/10.18653/v1/P19-1116>

⁴³ Anthony PYM: Translation Skill-Sets in a Machine-Translation Age. *Meta*, vol. LVIII., no. 3. (2013) 487–503. 496. <http://dx.doi.org/10.7202/1025047ar>

⁴⁴ Justyna ZANDBERG-MALEC: Prosty język w komunikacji prawniczej – okiem redaktora językowego. *Poznańskie Studia Polonistyczne*. [Simple language in legal communication - through the eyes of a language editor. Poznan Studies in Polonistics] *Seria Językoznawcza*, vol. 28. no. 1. (2021) 191–204. 193. <http://dx.doi.org/10.14746/pspsj.2021.28.1.13>

⁴⁵ FELSENFELD op. cit. 419–420.; KIMBLE (2006) op. cit. 69–72.; ZANDBERG-MALEC op. cit. 194–197.

on potential readers.⁴⁶ So, although translators face the challenge of translating plain language contracts before this style of drafting gains ground in Polish, one day our solutions and intuitions must be adjusted to emerging local practice. The new register of plain language contracts in Polish merits also attention from providers of MT services, who need to include them in training data.

⁴⁶ FELSENFELD op. cit. 419.; ZYCH op. cit. 75., HADRYAN (2009) 23.