

dr hab. Konrad Klimkowski, prof. UMCS
Instytut Językoznawstwa i Literaturoznawstwa
Katedra Lingwistyki Stosowanej
Uniwersytet Marii Skłodowskiej-Curie w Lublinie

Lublin, 16 listopada 2025 r.

Recenzja rozprawy doktorskiej autorstwa Pani mgr Anny Setkowicz-Ryszki pt. „Post-editing of machine translation output in the training of legal translators: Comparing translation processes and products before and after a period of post-editing tasks”, napisanej pod kierunkiem prof. dr hab. Łukasza Boguckiego

Przedmiotem recenzji jest rozprawa doktorska mgr Anny Setkowicz-Ryszka, zatytułowana *Post-editing of machine translation output in the training of legal translators: Comparing translation processes and products before and after a period of post-editing tasks*. Praca, złożona we wrześniu 2025 roku w Szkole Doktorskiej Nauk Humanistycznych Uniwersytetu Łódzkiego koncentruje się na badaniu potencjalnego wpływu ćwiczeń z postedycji (PE) na rozwój „kompetencji eksperckiej” w tłumaczeniu tekstów prawnych. Struktura rozprawy obejmuje wprowadzenie (s. 1-2), dziewięć głównych rozdziałów, pięć załączników (m.in. teksty źródłowe, ankiety) oraz wykazy tabel i rysunków. Bibliografia obejmuje 202 pozycje i zajmuje 41 stron rozprawy (360-401).

Część pierwsza rozprawy (Introduction) to ściślej rzecz biorąc wprowadzenie (Foreword), w którym co prawda pojawia się motywacja autorki do podjęcia studium oraz określone są obserwowane przez nią problemy badawcze w polu przedmiotu badań, ale cele badawcze zarysowane są stosunkowo ogólnie (trzeci akapit w Introduction). Nie pojawia się tu stwierdzenie, w jaki sposób te – istotne badawczo, ale bardzo ogólne – cele mają zostać zoperacjonalizowane. Zasadniczo autorka nie udziela też odpowiedzi na sformułowane tu pytania w części podsumowującej pracę (Conclusions).

Po części wprowadzającej rozpoczyna się imponująca ilościowo i jakościowo część teoretyczna pracy, na którą składają się Rozdziały 1–4. Rozdział 1 (Tytuł: *Translation – revision – post-editing*, s. 3) wprowadza definicje kluczowych pojęć pracy: tłumaczenie od podstaw (TfS), tłumaczenie maszynowe (MT), postedycja (PE w podziale na automatyczną, statyczną i interaktywną) oraz weryfikacja tekstu docelowego (Revision). W dalszych podrozdziałach (1.2–1.4) autorka przeprowadza analizę porównawczą kompetencji tłumaczeniowej (TC) i kompetencji postedycyjnej (PEC) oraz dokonuje przeglądu modeli kompetencji tłumaczeniowej (TC), w tym kompetencji tłumacza tekstów prawniczych (LTC). Prezentowane są modele EMT, TransComp (Susanne Göpferich), a także te oparte na standardach ISO 17100 i ISO 20771:2020. Autorka omawia również kompetencje wymagane od postedytorów. Rozdział kończy się dyskusją na temat eksperckiej kompetencji tłumaczeniowej (*translation expertise*) – oraz relacji pomiędzy pojęciami *competence* i *expertise*.

Rozdziały 2–4 stanowią szczegółowy kontekst dla omówienia zagadnień tłumaczenia tekstów prawnych/prawniczych. Rozdział 2 koncentruje się na języku prawnym i prawniczym, omówieniu cech, takich jak skomplikowana składnia, nadmierna „formułowość” (boilerplate clauses), czy zjawisko Plain English. Rozdział ten zawiera także szczegółową sekcję dotyczącą terminologii prawnej/prawniczej (Legal terminology), jej systemowości oraz metod poszukiwania ekwiwalentów. Poruszane są także zagadnienia różnic w systemach prawa jako źródła wyzwań dla tłumacza. Pojawia się także zagadnienie typologii błędów tłumaczeniowych, co ma istotne znaczenie dla dalszych, empirycznych, etapów studium. Rozdział 3 porusza kwestie związane z dydaktyką tłumaczenia tekstów prawnych i prawniczych, porównując modele LTC i omawiając, w jaki sposób można badać i wzmacniać kompetencje studentów. Podrozdział 3.2. podejmuje zagadnienia oceny (percepcji) własnej kompetencji tłumaczeniowej przez tłumaczy tekstów prawnych/prawniczych. Rozdział 4 szczegółowo analizuje specyfikę umów, koncentrując się na ich makrostrukturze, klauzulach kluczowych, użyciu operatorów modalnych, frazeologii i tzw. „boilerplate provisions”.

Część teoretyczna rozprawy (Rozdziały 1–4) wyróżnia się szerokim spektrum omawianych zagadnień, niezbędnym do osadzenia badań nad postedycją w kontekście rozwoju kompetencji eksperckich tłumaczy tekstów prawnych/prawniczych. Autorka dokonała skutecznego uporządkowania prezentowanych teorii, wychodząc od definicji kluczowych dla pracy koncepcji (tłumaczenia TfS, MT, PE, weryfikacji), a następnie płynnie przechodząc do złożonych modeli kompetencji tłumaczeniowej (TC) i kompetencji tłumacza tekstów prawnych i prawniczych (LTC). Szczególnie cenne jest obszerne porównanie różnorodnych modeli LTC (m.in. Kelly, PACTE, EMT, TransComp, Soriano-Barabino, ISO, QUALETRA), które pozwoliło na identyfikację wspólnych elementów i potencjalnych empirycznych eksponentów kompetencji eksperckiej w produkcji i procesie tłumaczenia. Na uwagę zasługuje również merytoryczna dyskusja z prezentowanymi teoriami. Autorka nie ogranicza się do pasywnego ich przeglądu, lecz aktywnie wiąże abstrakcyjne składowe

kompetencji (np. kompetencję strategiczną, metakognitywną, domenową) z konkretnymi wyzwaniami LT, takimi jak trudności w wykrywaniu błędów adekwatności w pozornie poprawnym tłumaczeniu maszynowym, czy kwestie związane z formułami prawnymi (formułami deontycznymi, binomials/trinomials). Dokładna analiza cech języka umów i terminologii prawnej, z uwzględnieniem różnic między prawem cywilnym a *common law* stanowi solidną bazę wiedzy domenowej dla skutecznego przeprowadzenia części empirycznej studium.

Rozdział 5 (*Study design*, s. 158) definiuje sześć szczegółowych pytań i hipotez badawczych (RQs 1–6), dotyczących poprawy jakości tłumaczenia (RQ1, RQ2), zmian w procesach tłumaczenia (RQ3, RQ4) i deklarowanej postawy wobec postedycji (RQ5, RQ6). Autorka przyjęła triangulację metod badawczych. Metody badawcze obejmują:

1. Ocenę produktu (Translation product evaluation) przy użyciu metody PIE (Preselected Items Evaluation), skupiającą się na 25 predefiniowanych elementach charakterystycznych dla języka prawnego/prawniczego w każdym tekście.
2. Analizę procesu (Process analysis) z wykorzystaniem Keylogging (Inputlog), Task Segment Framework (TSF) i rejestracji ekranu w celu identyfikacji wyznaczników rozwijającej się kompetencji eksperckiej (expertise).
3. Badanie postaw za pomocą anonimowej ankiety z pytaniami zamkniętymi i otwartymi.

Rozdział 5 prezentuje ambitny projekt badań empirycznych, który celuje w zrozumienie nie tylko co się zmienia (produkt), ale i *jak* się zmienia (proces i postawy) w wyniku interwencji dydaktycznej opartej na postedycji.

Kluczową zaletą metodologii jest jej oparcie na zasadzie triangulacji. Autorka słusznie zauważa, że kompetencji nie można zmierzyć bezpośrednio, lecz jedynie za pomocą jej manifestacji w produkcie i procesie. Aby uzyskać pogłębiony, maksymalnie obiektywny wgląd, autorka połączyła:

1. Metody zorientowane na produkt (ocena wybranych elementów PIE).
2. Metody zorientowane na proces (Keylogging, nagrania ekranu, Task Segment Framework).
3. Metody badające postawy (ankieta ilościowa i jakościowa).

Projekt badawczy jest również ambitny pod względem organizacyjnym. Uwzględnia zróżnicowaną grupę uczestników (studenci różnych programów, prawnicy, tłumacze praktycy). Szczegółowo opisane metody analizy (np. PIE z modyfikacją klasyfikacji rozwiązań) oraz techniki rejestracji procesu (np. Keylogging i TSF) świadczą o wysokim rygorze metodycznym badań i dogłębnym zaplanowaniu studium.

Rozdział 6 (*Analysis of translation solutions*, s. 175) to ilościowa analiza przekładu jako produktu. Autorka przedstawia wyniki dla zadań 1 i 2 dla 34 uczestników, porównując wyniki ogólne oraz rozkład w poszczególnych, siedmiu kategoriach badanych itemów.

Wykorzystano zaawansowane testy statystyczne (ANOVA) do oceny istotności różnic między grupami (studenci MA, podyplomowi, grupa mieszana) i wpływu doświadczenia.

Rozdział 7 (*Analysis of translation processes*, s. 223) zawiera szczegółową analizę procesu przekładu, choć autorka przyznaje, że zakres zebranych tu informacji (Keylogging) jest bardzo ograniczony. Autorka wskazuje, że udało się zbadać 22 z 34 uczestników badania, co nie pomniejsza roli tej próby i chyba niestuszenie autorka opisuje liczbę tych badań zawężającym określeniem „few”. Oczywiście ta mniejsza liczba danych ogranicza możliwości porównawcze czy korelacyjne, ale daje pewien obraz obserwacji procesów przekładu. Rozdział przedstawia ogólne obserwacje dotyczące czasu pracy, środowiska tłumaczeniowego (screen layout), wykorzystania wiedzy oraz strategii rozwiązywania problemów i użycia źródeł informacji (np. słowników, PTs, korpusów, ustawodawstwa). Centralnym elementem są porównania zmian zachodzących u poszczególnych uczestników (within-participant comparisons), oparte na Keylogging, rejestracji ekranu i wywiadach, opisujące styl pracy i domniemany rozwój kompetencji 22 tłumaczy. Analiza dostarcza ilościowych wskaźników zmian w zachowaniach tłumaczy. Wskaźniki takie jak czas spędzony na zadaniu, czas pauzy, udział czasu wyszukiwania w całkowitym czasie procesu oraz analiza segmentów (ADD, CHG, SEARCH, FILLER) umożliwiają precyzyjny opis zmian w stylu pracy (np. redukcja całkowitego czasu pracy, wzrost wydajności u niektórych uczestników). Analiza ta jest kluczowa dla obiektywizacji spostrzeżeń dotyczących rozwijającej się kompetencji (od competence do expertise) (np. płynność przechodzenia między wyszukiwaniem a pisanem, czy wzrost świadomości metakognitywnej).

Rozdział 8 (*Questionnaire survey*, s. 308) analizuje postawy uczestników wobec tłumaczenia maszynowego i postędy. Autorka przeprowadza analizę ilościową pytań zamkniętych (dotyczących tła, stylu pracy i opinii) oraz analizę jakościową odpowiedzi na pytania otwarte. W analizie ilościowej badano również powiązania statystyczne (np. korelacja pomiędzy posiadanym doświadczeniem zawodowym a percepcją użyteczności procesów postędy). Rozdział przedstawia niezbędną kategoryzację zebranych danych ankietowych, uzupełniając obraz empiryczny o perspektywę subiektywną uczestników.

Analiza jakościowa odpowiedzi na pytania otwarte (sekcja 8.2) jest szczególnie ważna, ponieważ ujawnia poszczególne aspekty postawy danego respondenta, które nie mogły zostać uchwycone w pytaniach zamkniętych (gdzie brakowało opcji „to zależy”). Autorka skutecznie identyfikuje i kategoryzuje powtarzające się wątki, takie jak:

- Aspekt czasu: Postrzeganie PE jako szybszej niż TfS (21 odpowiedzi), choć z zastrzeżeniem, że czasami poprawa zajmuje więcej czasu niż tłumaczenie od nowa.
- Aspekt wysiłku: Uznanie, że PE jest wysiłkowa, wymaga doświadczenia, wiedzy merytorycznej i kontroli (14 odpowiedzi), zwłaszcza ze względu na „łańcuchy edycji” w języku fleksyjnym, takim jak polski.

- Jakość MT: Krytyka MT za niską jakość, błędy terminologiczne, kalki językowe, czy niespójność, ale też dostrzeżenie jej zalet (np. zachowanie składni, „szkieletu” zdania).
- Wpływ priming’u: Problemy z odrzuceniem sugestii aplikacji MT.
- Preferencje TfS: Uczestnicy preferują tłumaczenie „klasyczne” ze względu na kreatywność i większą satysfakcję.

Ta dogłębna analiza jakościowa nadaje „ludzki wymiar” danym ilościowym, wyjaśniając, między innymi, dlaczego, mimo faktu, że PE przyspiesza pracę, jest jednocześnie postrzegana jako działanie wymagające wysokiego poziomu skupienia i często tak samo (lub bardziej) wysiłkowe niż TfS.

Rozdział 9 (*Answers to research questions. Conclusions*, s. 330) dostarcza syntetycznych odpowiedzi na 6 pytań badawczych rozprawy. Autorce nie udaje się wykazać silnych korelacji między badanymi zmiennymi, zatem możemy mówić o obserwowanych przez nią tendencjach w interpretowanych danych (lista punktowana i ostatni akapit na str. 335):

- RQ1 (Produkt): Nastąpiła umiarkowana poprawa ogólnych wyników w zakresie jakości tłumaczenia (0,74 punktu na 50 możliwych), głównie w kategoriach wiedzy o świecie (KNOWL) oraz terminologii i konwencji lokalnych (TERM+LOCAL).
- RQ2 (Tło zawodowe): Nie stwierdzono istotnych statystycznie różnic w poprawie jakości między trzema głównymi grupami (TRAN, IPSKT, MIX). Jednak doświadczenie zawodowe było powiązane z lepszymi wynikami w kluczowych kategoriach zadań (PHRAS i KNOWL).
- RQ3 (Proces): Odnotowano skrócenie całkowitego czasu pracy u większości uczestników. Badania sugerują, że dzięki ćwiczeniom PE uczestnicy nabyli bardziej zaawansowane umiejętności wyszukiwania informacji, w tym korzystania z ustawodawstwa i tekstów paralelnych, oraz wykazali się większą świadomością metakognitywną (np. identyfikacją problemów i wiedzy o lukach).
- RQ4 (Proces a akceptowalność produktu): Stwierdzono, że zachowania świadczące o świadomości problemu (pauzy, oznaczanie wątpliwości, intensywne wyszukiwania) nie zawsze skutkowały akceptowalnymi rozwiązaniami, co sugeruje, że rozwój kompetencji eksperckiej w zakresie integracji umiejętności wymaga więcej czasu niż interwencja dydaktyczna.
- RQ5 (Postawy): Większość respondentów uważa jakość MT za niską, wymagającą dużego wysiłku w postędy. Jednocześnie postędy jest postrzegana jako pomocna w nauce tłumaczenia tekstów prawnych/prawniczych, zwłaszcza w rozwijaniu umiejętności wyszukiwania informacji i weryfikacji tekstów.

Ostatni akapit na stronie 335 podkreśla, że główny wpływ ćwiczeń PE zaobserwowano w zmianach w procesie tłumaczenia, a niekoniecznie w jakości samego produktu w krótkim okresie. PE aktywnie promuje ekspertyzę adaptacyjny model kompetencji eksperckiej

(*adaptive expertise*) poprzez stawianie tłumaczom dobrze określonych zadań w zakresie identyfikacji i rozwiązywania problemów tłumaczeniowych, krytycznej analizy źródeł (zwłaszcza w kontekście MT) oraz doskonalenia strategii podejmowania decyzji.

Obraz, który uzyskuje autorka w wyniku przeprowadzonych badań jest bardzo skomplikowany i próba budowania prostych syntez wynikowych właściwie nie mogła się powieść. Różnice pomiędzy 34 uczestnikami badania w różnych wymiarach badań ciężko jest sprowadzić do jednoznacznej odpowiedzi o tym, czy zajęcia poświęcone postedycji to „dobry” czy „zły” pomysł edukacyjny, dla kogo i kiedy. Być może jest tak dlatego, że zastosowanie tak złożonego aparatu statystycznego dla stosunkowo niewielkiej grupy badanych nie mogło dać wyników „populacyjnych”, umożliwiających ich operacyjną generalizację. Da się też założyć w granicach dociekań przednaukowych, nieempirycznych, że cechy indywidualne poszczególnych respondentów (poziom wiedzy, poziom kompetencji językowej w obu językach ogólnym i specjalistycznym, zaangażowanie/motywacja w wykonanie zadania) mogły mieć dobry albo zły wpływ na poszczególne wyniki (fragmentaryczne czy całościowe) w stopniu równym a może większym niż przyjęte kategorie pomiaru. Wyraźniejszy obraz pojawia się dla pytania drugiego w odniesieniu do kategorii doświadczenia zawodowego w korelacji z wybranymi zmiennymi w ramach zadań 1 i 2, zaś największy stopień generalizacji wyników napotkać można na s. 334-335 przy syntezie dotyczącej pytania nr 6, gdzie autorka przedstawia słabe korelacje zmiennych wskazujących np. na to, że badani posiadający ograniczone doświadczenie zawodowe doceniają rolę postedycji w rozwoju kompetencji w zakresie tłumaczenia tekstów prawnych i prawniczych.

Rolę konkluzji wydaje się pełnić ostatni akapit pracy, w którym autorka stwierdza, że techniki (i technologie) postedycji będą przynosiły korzyści rozwojowe i zawodowe tym, którzy mają zaawansowaną „niecyfrową” kompetencję w zakresie korekty i weryfikacji tekstu przekładu, ale mogą mieć charakter bariery rozwojowej dla osób, które zaczęły pracę z narzędziami postedycyjnymi przed wykształceniem kompetencji „klasycznych”. Przy czym sama autorka zauważa, że nawet ta konstatacja nie może mieć charakteru uniwersalnego podsumowania pracy, gdyż: „each individual participant has their own story that explains why their progres was as it was...”

Zdaniem autora niniejszej recenzji istnieją dwa różne sposoby odczytania wyników badań autorki. Można oceniać je z perspektywy pozytywistycznej, z naciskiem na uzyskanie klarownych danych i relacji między nimi, jako podstawy oceny sukcesu projektu badawczego. W tej perspektywie, mimo bezdyskusyjnie bardzo wysokiej jakości metodologii badań, złożona struktura danych zasadniczo wymyka się „obiektywnym generalizacjom”. Widzimy co najwyżej tendencje, ale i ich wypadku autorka przyznaje, że nie do końca wiadomo, do jakiego stopnia uzyskane dane nie są „historią indywidualnych przypadków”.

Możliwe jest jednak odczytanie alternatywne, które można określić jako relatywistyczne i konstruktywistyczne, zadowalające się prawdopodobieństwem możliwości wystąpienia pewnych faktów i wykorzystania ich do programowania interwencji dydaktycznych, które mogą – nigdy nie muszą – dać pozytywne efekty w oparciu o kontekstowe (nie uniwersalne) wykorzystanie odkrytych przez autorkę tendencji. Przykładowo, autor recenzji dostrzega znaczącą wartość spostrzeżenia autorki, że – najprawdopodobniej – w przypadku postępcy sprawności procesowe rosną szybciej niż efektywność produktowa, co może być strategiczną wskazówką do planowania ewaluacji zajęć obejmujących postępcę.

Na koniec recenzji autor pozwala sobie na kilka uwag krytycznych wobec rozprawy. Pierwsza uwaga zasadniczo została już zarysowana przy omawianiu części wstępnej (Introduction). W odbiorze recenzenta części tej brakuje klasycznej dla anglosaskiego dyskursu naukowego struktury: przedmiot – problem – cel badań. Formułowany tu zarzut nie ma charakteru stricte formalnego: ostatecznie przedmiot, problem i cele (jako ogólne pytania nurtujące autorkę) można „jakoś” wyczytać ze wstępu. Intencją autora jest raczej wskazanie, że sugerowana tu struktura wstępu (a co za tym idzie zakończenia – Conclusions) pomogłaby autorce lepiej zbudować perswazyjną warstwę przekazu i wykazać, w jaki sposób można odpowiedzieć na nurtujące ją pytania mimo komplikacji wyników badań empirycznych.

Nieoptymalne – zdaniem recenzenta – rozplanowanie wstępu prowadzi do słabo wyrażonych konkluzji. Nie chodzi oczywiście o konkluzje dotyczące części empirycznej, gdzie przedstawione przez autorkę syntezy są w pełni przekonujące. Chodzi o to, że po lekturze pracy nie wiemy np. w jaki sposób odkryte tendencje wpłyną na prowadzone przez autorkę zajęcia czy to tłumaczenia czy postępcy tekstów prawnych/prawniczych. Oczywiście jest, że autorka nadała pracy charakter teoretyczno-empiryczny, a nie koniecznie praktyczny (aplikatywny), ale stawiane w części wstępnej pytania mają wymiar praktyczny, który nie powinien zupełnie zniknąć na poziomie konkluzji rozprawy.

Praca wykazuje nieliczne niedociągnięcia techniczno-edytorskie: kilka błędów w odsyłaczach do literatury (reference errors np. s. 16); fragment „...compares the aggregate results.. s. 176 wydaje się niezakończony w żadnym poprzedzającym zdaniu nadrzędnym.

Język pracy jest klarowny, a autorka dobrze wykorzystuje narzędzia strukturyzacji tekstu (paragraphing), aby uzyskać precyzję wypowiedzi. Stylistyka tekstu zdradza czasem lekkie interferencje z języka polskiego, ale bez przekraczania granicy błędu stylistycznego (np. each individual participant, s. 335). Można także odnaleźć kilka błędów natury gramatycznej (One can say that Plain Polish was also been born out of the difficulties, s.72), ale są one ewidentnym „przeoczeniem”, a nie brakiem kompetencji językowo-tekstowej autorki.

W podsumowaniu niniejszej recenzji należy bezdyskusyjnie stwierdzić, że przedłożona przez kandydatkę do stopnia doktora rozprawa spełnia wszelkie wymogi formalne

niezbędne do sformułowania recenzji pozytywnej. Autor niniejszej recenzji stwierdza także, iż rozprawa doktorska autorstwa Pani magister Anny Setkowicz-Ryszki spełnia wszystkie warunki określone odpowiednimi przepisami prawa Rzeczypospolitej Polskiej. Dlatego też recenzent wnosi do o dopuszczenie kandydatki do dalszych etapów postępowania o nadanie stopnia doktora w dyscyplinie językoznawstwo. Dodatkowo, ze względu na fakt, że część teoretyczna rozprawy posiada wyjątkowo szeroki zakres dociekań, oraz że metodologia badań jest ambitna, precyzyjnie zaplanowana i rygorystycznie zrealizowana, recenzent przedkłada wniosek o wyróżnienie rozprawy.

Kound Klisulubi