

MICHAŁ HRABIA

Uniwersytet Śląski w Katowicach, Wydział Humanistyczny, Instytut Językoznawstwa

michal.hrabia@us.edu.pl

ORCID: 0000-0001-5477-8211

## **Sztuczna inteligencja w służbie tłumaczenia: Google Translate i ChatGPT-3.5 wobec regułowego przekładu automatycznego. Analiza porównawcza na materiale języka francuskiego i polskiego**

**Artificial Intelligence in the Service of Translation: Google Translate  
and ChatGPT-3.5 versus Rule-Based Machine Translation.  
A Comparative Analysis on French and Polish Language Data**

### **Abstract**

The aim of this article is to evaluate the performance of modern translation systems powered by Artificial Intelligence (AI). The opening section outlines the evolution of machine translation technologies, from statistical and rule-based methods to AI-driven models. It provides a detailed overview of two contemporary, freely available systems supported by AI – Google Translate, and ChatGPT-3.5 – as well as a traditional rule-based translation model founded on the object-oriented approach by Wiesław Banyś. Subsequently, the author compares translations produced by AI systems with those that could be generated based on linguistic research conducted within the object-oriented theoretical framework. The analysis focuses on Polish translations of the polysemous French verbs *monter*, *entrer*, *remettre*, and the adjective *farouche*, with particular attention given to the systems' ability to perform word-sense disambiguation. Drawing on this analysis, the author proposes a typology of translation errors observed in AI-generated outputs, identifying cardinal semantic errors, errors of semantic excess or deficiency, lexical errors, inflectional-syntactic errors and register-related errors.

**Keywords:** machine translation, word-sense disambiguation, translation errors, Artificial Intelligence (AI), ChatGPT, Google Translate, object-oriented approach

**Słowa kluczowe:** tłumaczenie automatyczne, odwieloznacznianie słów, błędy tłumaczeniowe, sztuczna inteligencja, ChatGPT, Tłumacz Google, ujęcie zorientowane obiektowo

**Open Access.**  Published by Committee for Philology of the Polish Academy of Sciences, Wrocław Branch and College for Interdisciplinary Studies, University of Wrocław. This work is licensed under the Creative Commons Attribution-ShareAlike 3.0 Unported License.

Niniejszy artykuł, pisany u progu „zupełnie nowej ery”, eksploruje zagadnienie tłumaczenia automatycznego wspieranego przez sztuczną inteligencję. W pierwszej części prześledzimy rozwój technologii wykorzystywanych w przekładzie maszynowym, koncentrując się na najnowszych rozwiązaniach opartych na architekturze transformatorów tj. ChacieGPT i Google Translate<sup>1</sup>. Następnie porównamy tłumaczenia generowane przez systemy AI z efektami, jakie można by uzyskać za pomocą programów opartych na klasycznych metodach regułowych. Jako przykład paradygmatu regułowego posłuży nam ujęcie zorientowane obiektowo, zaproponowane na początku obecnego stulecia przez Wiesława Banysia (2002ab, 2005).

### **Tłumaczenie automatyczne: od *invisible idiot* do sztucznej inteligencji**

Choć w powszechnej świadomości tłumaczenie automatyczne jest wynalazkiem stosunkowo młodym, badania nad tą technologią trwają już od lat 60. Za symboliczną datę wyznaczającą narodziny lingwistyki komputerowej uznaje się bowiem rok 1962, kiedy to założone zostało *The Association for Machine Translation and Computational Linguistics*, którego głównym celem miała być praca na rzecz rozwoju metodologii badań językoznawczych wykorzystujących moce obliczeniowe komputera. Przez wiele lat do tłumaczenia automatycznego podchodzono jednak sceptycznie, a nawet anegdotycznie. Znana jest historia pewnego Amerykanina, który przybywszy do Związku Radzieckiego, przodującego wówczas w językoznawstwie komputerowym, poprosił zaprojektowaną przez Rosjan maszynę o przetłumaczenie zdania *Out of sign, out of mind* (pl. *Co z oczu, to z serca*). Ponieważ nie znał rosyjskiego, nie mógł ocenić jakości wygenerowanego przekładu. Poprosił więc o jego retranslację na język angielski i ku powszechnej konsternacji maszyna wydrukowała tekst: *invisible idiot* (pl. *niewidzialny idiota*) (zob. Lipiński 2004: 98). Od tamtych czasów zmieniło się oczywiście wiele, a rewolucja cyfrowa XXI wieku przyniosła rozwiązania i możliwości, o których 80 lat temu nikt nawet nie mógł pomyśleć. Jak słusznie zauważa Beata Śmigielska (2023: 13), pomimo pewnych niedociągnięć „[...] współczesne translatory automatyczne są już zazwyczaj w stanie zapewnić bardzo dobre tłumaczenia” [tłum. własne]<sup>2</sup>.

W historii rozwoju tłumaczenia automatycznego zwykło wyróżniać się cztery etapy (por. Bogacki 2023: 9–10): tłumaczenie regułowe (oparte na regułach sformułowanych przez językoznawców i zintegrowanych ze słownikami elektronicznymi), tłumaczenie statystyczne (wykorzystujące modele statystyczne i probabilistyczne), tłumaczenie neuronalne (wykorzystujące algorytmy oparte na sieciach neuronowych) oraz tłumaczenie hybrydowe (integrujące w mniejszym lub większym stopniu wcześniejsze paradygmaty). Ostatni, piąty etap, który należy bez wątpienia dodać do powyższej listy, to tłumaczenie wspierane przez modele sztucznej inteligencji (*Artificial Intelligence [AI]*), wykorzystujące powstałą

1 Zgodnie z wiedzą autora tego rodzaju opracowanie podejmowane jest w polskojęzycznym językoznawstwie po raz pierwszy.

2 „[...] les traducteurs automatiques actuels sont déjà capables de donner généralement de très bonnes traductions”.

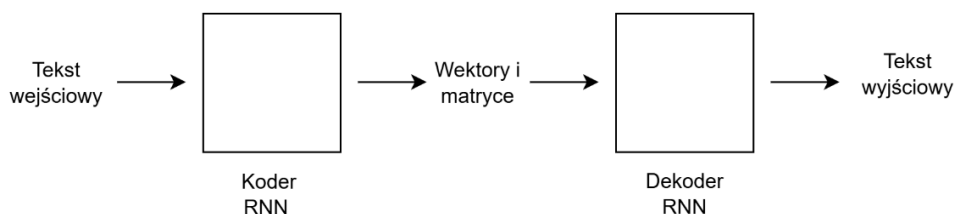
w 2017 r. technologię transformatorów. W kolejnych rozdziałach przyjrzymy się bliżej dwóm narzędziom translacyjnym bazującym na AI: Tłumaczowi Google (Google Translate) oraz Chatowi GPT 3.5.

## Technologie AI w tłumaczeniu automatycznym. Google Translate: od statystyki do transformatorów

Google Translate to darmowa, łatwa w obsłudze i bez wątpienia najbardziej znana aplikacja do tłumaczenia automatycznego, używana przez miliony ludzi na całym świecie. Obsługuje aktualnie ponad 130 języków, w tym tak egzotyczne jak krio (język kreolski oparty na angielskim i lokalnych językach Afryki Zachodniej), shona (język z rodziny bantu) czy dogri (język indoirañski).

Początki Tłumacza Google sięgają roku 2006. Aplikacja wykorzystywała wówczas metody statystyczne (*Statistical Methods SMT*), generując przekład w oparciu o analizę dostępnych on-line korpusów paralelnych zbudowanych z teksów uprzednio przetłumaczonych przez człowieka. Proces przekładu rozpoczynał się od fragmentowania tekstu wyjściowego na mniejsze sekwencje, po którym następowało wyszukiwanie sekwencji zbliżonych w korpusie języka docelowego (por. Kozłowski 2004: 71, Hrabia 2018a: 339–340).

Pierwszą rewolucję przyniósł rok 2014, w którym to Tłumacz Google zaczął korzystać z architektury kodera-dekodera (*Encoder-Decoder*) opartej na rekurencyjnych sieciach neuronowych (*Recurrent Neural Networks RNN*) (zob. Kalchbrenner, Blunsom 2013; Cho *et al.* 2014; Sharma *et al.* 2023: 13–14). Sieci takie używają do przesyłania informacji sprzężenia zwrotnego, co oznacza, że dane wyjściowe i stan komórki z każdego etapu stają się danymi wejściowymi w następnym kroku przetwarzania. Sieć neuronowa przechowuje więc informacje z wcześniejszych faz działania i wykorzystuje je przy analizie kolejnych sekwencji, co stanowi pewną formę sztucznej „pamięci”. Łatwo można porównać taki proces z czytaniem tekstu przez człowieka: gdy czytamy zdanie zachowujemy w pamięci pewną dawkę informacji i używamy jej jako kontekstu niezbędnego do zrozumienia kolejnego zdania w tekście. W modelu Google RNN zastosowano dwie sieci neuronowe: koder, który przekształcał tekst w języku wejściowym na format pośredni, niejęzykowy tj. wektory i matryce, oraz dekodery, który zamieniał wektory na tekst w języku docelowym (rys. 1).

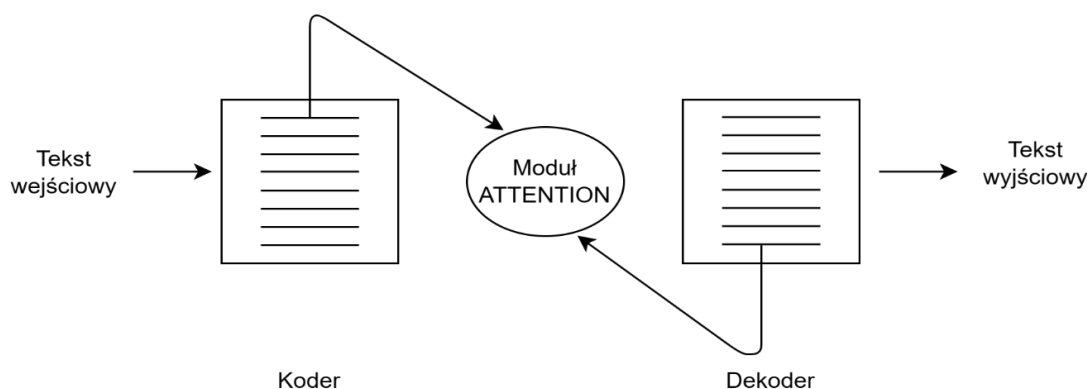


Rys. 1. Architektura kodera-dekodera RNN<sup>3</sup>

3 Wszystkie rysunki i schematy zostały sporządzone przez autora.

Sama umiejętność „zapamiętywania” informacji przez sieci neuronowe szybko okazała się niewystarczająca. Niedługo później Google wprowadził więc do swojej aplikacji dodatkowy mechanizm – *Attention* – imitujący ludzką uwagę (zob. Bahdanau *et al.* 2014, Luong *et al.* 2015). Mechanizm ten pozwalał maszynie skoncentrować się na informacjach ważnych (np. pewnych fragmentach zdania, kluczowych syntagmach), a zignorować te mniej istotne. W kolejnych latach okazało się, że to właśnie „uwaga” będzie odgrywać zasadniczą rolę w tłumaczeniu automatycznym.

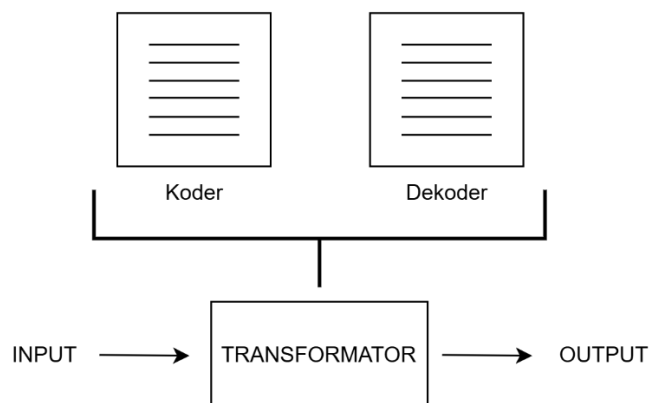
W listopadzie 2016 r. Google zaproponował ulepszoną wersję modelu z 2014 r.: *Google Neural Machine Translation GMNT* (Wu *et al.* 2016). Koder i dekodek zostały znacznie rozbudowane (w każdym z nich znalazło się 8 warstw sieci neuronowych), a mechanizm *Attention* zaczął odgrywać rolę modułu łącznika między nimi (rys. 2). Umożliwiło to efektywne sprawdzanie danych wychodzących z kodera pod kątem ich istotności dla bieżących obliczeń po stronie dekodek. Wprowadzono również inne udoskonalenia, z których na szczególną uwagę zasługuje *Zero-Shot Translation* – mechanizm bazujący na maszynowym uczeniu głębokim (ang. *deep learning*), umożliwiający tłumaczenie automatyczne między parami językowymi, dla których nie są dostępne paralelne korpusy (np. między językiem polskim a dogri) (zob. Johnson *et al.* 2016).



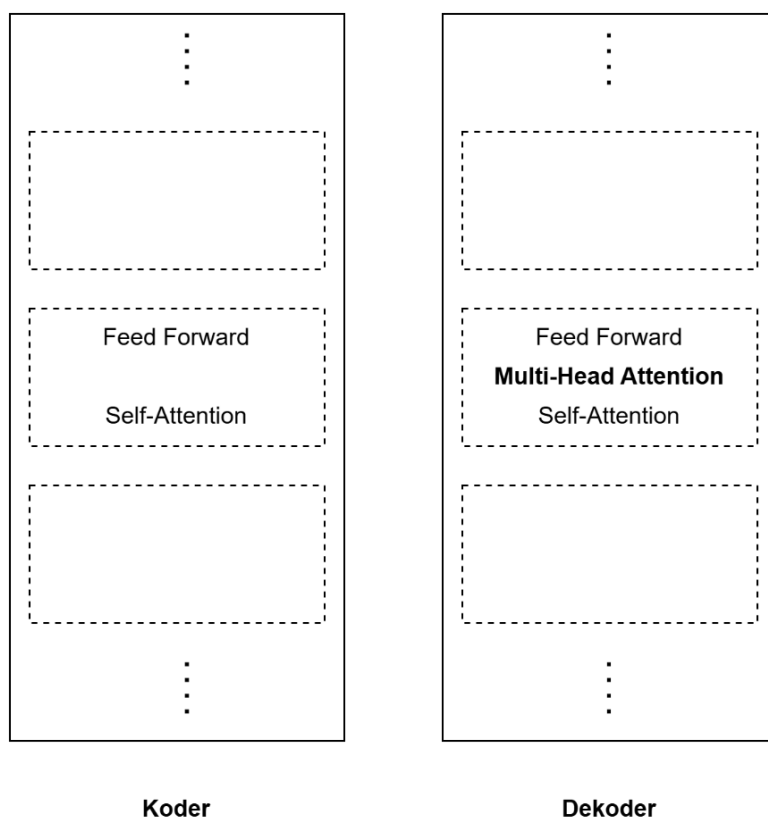
Rys. 2. Architektura *Google Neural Machine Translation GMNT*

W roku 2017 ukazał się przełomowy, nie tylko dla tłumaczenia automatycznego, ale dla rozwoju całej sztucznej inteligencji, artykuł „Attention is all you need” (Vaswani *et al.* 2017), napisany przez naukowców współpracujących z Google. Zaproponowano w nim nową architekturę systemu opartą na modelu transformatora, rezygnując całkowicie z wykorzystywania rekurencyjnych czy konwolucyjnych sieci neuronowych. Architektura ta wykorzystywana jest przez Tłumacza Google do dziś.

Model transformatora zbudowany jest, podobnie jak GMNT, z dwóch modułów: kodera i dekodek. Każda z sześciu warstw kodera składa się z dwóch podwarstw: *Self-Attention* oraz *Feed Forward*. W strukturze dekodek znajduje się ponadto podwarstwa *Attention* (a dokładniej: *Multi-Head Attention*) umieszczona między *Self-Attention* a *Feed Forward* (zob. Sharma 2023 : 17–20, Alammar 2018) (rys. 3 i 4).



Rys. 3. Architektura transformatora



Rys. 4. Budowa warstw kodera i dekodera w architekturze transformatora

Kluczową rolę w tym modelu, zarówno po stronie kodera, jak i dekodera, odgrywa mechanizm uwagi (*Attention*) oraz autouwagi (*Self-Attention*). *Self-Attention* ocenia stopień relacji między wszystkimi słowami w tekście wyjściowym, co umożliwia ich odpowiednie skontekstualizowanie w odniesieniu do całości. *Attention* ocenia z kolei wagę relacji między każdym słowem w języku wejściowym a każdym słowem w języku docelowym, nakreślając tym samym zależności między danymi wejściowymi a wyjściowymi. W przeciwieństwie do architektury opartej na RNN w modelu transformatora przypisywanie wagi poszczególnym słowom odbywa się w sposób równoczesny, co w zauważalny sposób zwiększa jakość przekładu.

### ChatGPT: od czatowania do tłumaczenia

ChatGPT 3.5, stworzony przez naukowców z Open AI – amerykańskiej organizacji badawczej założonej m.in. przez Elona Muska i Sama Altmana, udostępniony został za darmo szerokiemu gronu odbiorców w listopadzie 2022 r., wywołując natychmiast żywe reakcje na całym świecie. Ten chatbot to przykład sztucznej inteligencji opartej na języku naturalnym, która przeszkolona na olbrzymiej bazie korpusowej może generować koherentny i kontekstualnie poprawny tekst w wielu językach. Wykorzystuje architekturę transformatora, podobną w swych ogólnych zarysach do tej zaproponowanej przez Vaswani *et al.* (2017) dla Tłumacza Google.

Open AI nieustannie rozwija swój projekt. Pierwsza wersja GPT (*Generative Pretrained Transformer*) powstała w 2018 r., ale już w 2019 r. model został udoskonalony dzięki wykorzystaniu procesu uczenia głębokiego nienadzorowanego (ang. *unsupervised deep learning*). W 2022 r. udostępniono zmodyfikowaną wersję trzeciej generacji chatu (GPT-3.5), a zaledwie rok później pojawił się GPT-4, który z pewnością nie będzie ostatnim.

W przeciwieństwie do Google Translate ChatGPT nie został pierwotnie pomyślany jako narzędzie do tłumaczenia automatycznego. Jak jednak słusznie zauważa Stéphane Trinh (2019), „sztuczna inteligencja i tłumaczenie ewoluują wspólnie” [tłum. własne]<sup>4</sup>. Nie sposób zatem nie zauważyć i nie docenić ogromnego potencjału translatorycznego GPT (por. np. Nazir, Wang 2023: 5; Jiao *et al.* 2023; Kalyan 2024: 15–16, Śmigielska 2023).

Ciągłe rozwijanie umiejętności tłumaczeniowych modelu możliwe jest dzięki zaawansowanemu procesowi „szkolenia”. GPT, jak sama nazwa wskazuje, to model jedynie „wstępnie przeszkolony” (ang. *pre-trained*), co oznacza, że na podstawie analizy dużych zbiorów danych tekstowych poznał ogólne wzorce danego języka. Proces uczenia się modelu na tym się jednak nie kończy. Kolejnym etapem jest *fine-tuning* (pol. dostrajanie), czyli dostosowywanie GPT do wypełniania specyficznych zadań (np. tłumaczenia automatycznego) poprzez trenowanie go na bazie mniejszego, odpowiednio dobranego zbioru danych (por. Nazir, Wang 2023: 2–3, Ray 2023: 126). Można np. dostarczać listy przykładów tych samych zdań lub tekstów w różnych językach, dzięki czemu model nauczy się lepiej przewidywać w przyszłości najbardziej odpowiednie tłumaczenia (zob. Jungha, Boyoung 2023: 15–16; Rybnikova 2024)<sup>5</sup>.

4 „[l]’intelligence artificielle et la traduction évoluent ensemble »

5 Zdaniem autora na etapie *fine-tuningu* możliwa i wskazana wydaje się współpraca między programistami a językoznawcami.

ChatGPT jest też bez wątpienia przydatnym narzędziem dla tłumacza-człowieka. Według Sai Cheng Siu (2023: 4–19) pomaga on m.in. w kontekstowym lub kulturowym wyjaśnianiu wyrażen, objaśnianiu terminów specjalistycznych, wyszukiwaniu błędów gramatycznych lub redakcyjnych itp. Wśród mocnych stron chatu Siu (2023: 19–22) wymienia przede wszystkim: szybkość działania i cenę, wszechstronność i elastyczność, kreatywność i zdolność adaptacji tłumaczenia oraz możliwość dostosowywania przekładu do kontekstu tłumaczeniowego. Warto jednak pamiętać o istotnych ograniczeniach modelu, do których należą m.in. (por. Siu 2023: 22–28):

- jakość przekładu – chat generuje zasadniczo poprawny przekład tekstów ogólnych, ale zdecydowanie gorsze są efekty tłumaczenia specjalistycznego np. z zakresu nauk biomedycznych (zob. Jiao *et al.* 2023);
- halucynacje ChatuGPT – czyli generowanie tłumaczenia, które nie ma oparcia w tekście źródłowym; oznacza to, że model wprowadza do przekładu dodatkowe informacje lub pomija ważne szczegóły;
- ograniczone wsparcie dla niektórych języków – najlepsze rezultaty daje tłumaczenie z i na angielski, gdyż to na tekstach w tym języku model był wstępnie szkolony;
- zmienność wyjściowa – GPT ma tendencję do generowania różnych tłumaczeń dla tego samego zdania/tekstu w różnych odstępach czasu;
- obawy dotyczące bezpieczeństwa – tekst w języku wyjściowym wprowadzany jest do zewnętrznych serwerów, co wiąże się z ryzykiem naruszenia poufności danych.

ChatGPT zapytany, czy jest lepszym tłumaczem od Google Translate, odpowiada, że „ogromny rozmiar pozwala [mu] na bardziej złożone przetwarzanie języka, co często przekłada się na lepszą dokładność tłumaczeń”. W dalszej części artykułu spróbujemy skonfrontować to stwierdzenie z rzeczywistością.

### **Klasy obiektowe: od programowania do regułowego tłumaczenia automatycznego**

Jak już wspominaliśmy, związki lingwistyki i translatoryki z naukami obliczeniowymi były dostrzegalne na długo przed powstaniem sztucznej inteligencji. Przenikanie się w tych dyscyplinach pewnych koncepcji i terminów jest zjawiskiem nierzadkim w zasadzie już od czasów Noama Chomskiego i jego słynnych prac nad gramatykami formalnymi (zob. Chomsky 1956). Za jeden z przykładów stosunkowo niedawnego „transferu wiedzy” między informatyką a językoznawstwem uznać można koncept „klasy obiektowej”. Jest to termin zaczerpnięty z teorii programowania obiektowego, jednego z głównych paradygmatów programistycznych, rozpowszechnionych dzięki językom takim jak C++ czy późniejszy C#. Na gruncie francuskiego językoznawstwa pojęcie to zaszczerpił Gaston Gross (1994, 2008, 2012), a w Polsce – Wiesław Banyś (2002ab).

W ujęciu zorientowanym obiektowo Banyśa (fr. *approche orientée objets AOO*) klasa obiektowa to zespół obiektów językowych (prototypowo rzeczowników), charakteryzujących się tym samym zestawem operatorów tj. operacji (czynności, które obiekt może wykonać lub które można wykonać na obiekcie) i atrybutów (cech obiektu) (zob. Banyś 2002a: 22). W rezultacie możliwe jest wyróżnienie np. klasy obiektowej <kot>, dlatego że wszystkie obiekty, które do niej przynależą (tj. instancje takie jak: *kot*,

*pers, mruczek, kocur* itp.), współdzielą pewien, im tylko właściwy, zbiór operatorów (np. *mruczy, miauczy, liże się* itd.)<sup>6</sup>.

Możliwość wykorzystania klas obiektowych w przekładzie maszynowym zauważył już Gross (1994: 29), ale szczególnie wybrzmiała ona w metodologii dezambiguizacji (odwieloznaczniania) słów polisemicznych opracowanej w ramach AOO przez Banysia (2005). Baniś zaproponował mianowicie w pełni sformalizowany, zorientowany obiektowo opis znaczeń leksemów za pomocą schematów składniowo-semantycznych, które mogą zostać łatwo zaimplementowane w systemach tłumaczenia automatycznego. Jako przykład niech posłuży nam wycinek deskrypcji francuskiego czasownika *fermer* (Hrabia 2018b: 141):

1. fr. X – [ANM hum] – *fermer* – Y – [CONC inanimé ; <partie du corps>]      **pl. zamykać**
2. fr. X – [ANM hum] – *fermer* – Y – [CONC <rideaux>]      **pl. zasłaniać**

W schemacie 1 na pozycji podmiotu (X) znajduje się klasa obiektowa [obiekt ożywiony - człowiek], a na pozycji dopełnienia (Y) – klasy [obiekt nieożywiony] lub <część ciała>. Jeżeli zdanie francuskie realizuje opisywany schemat, czasownik *fermer* będzie tłumaczony na język polski za pomocą ekwiwalentu *zamykać* (np. fr. *Jean a fermé la porte*, pl. *Jan zamknął drzwi*). Schemat 2 formalizuje z kolei sytuację, w której pozycja dopełnienia (Y) wypełniona jest przez klasę <zasłona>. W takim przypadku właściwym ekwiwalentem *fermer*, który powinien zostać wygenerowany przez system, jest czasownik *zasłaniać* (np. fr. *Jean a fermé les rideaux dans le salon*, pl. *Jan zasłonił zasłony w salonie*).

W kierowanym przez Wiesława Banysia nieistniejącym już Zakładzie Językoznawstwa Stosowanego i Translatoryki Uniwersytetu Śląskiego powstało wiele opracowań dezambiguizacyjnych (zob. np. Drzazga, Żłobińska-Nowak 2014; Gabrysiak 2009; Hrabia 2011, 2018c; Żłobińska-Nowak 2008, 2014). Wpisują się one w nurt rzadko już prowadzonych badań nad regułowym tłumaczeniem automatycznym. Pomimo wad, takich jak czasochłonność analiz korpusowych czy konieczność „ręcznego” wyliczania elementów tworzących klasy obiektowe (tzw. *listing*), badania te mają jedną niepowtarzalną zaletę: są wyjątkowo dokładne, a ich wyniki zapewniają dużą adekwatność przekładu.

W kolejnym rozdziale porównamy efekty dezambiguizacji zorientowanej obiektowo, wykonanej siłą myśli ludzkiej, z aktualnymi możliwościami tłumaczeniowymi AI.

### **Człowiek *versus* sztuczna inteligencja: analiza porównawcza przekładu. Korpus badawczy i metodologia badań**

Za wyjściowy materiał korpusowy posłużyły nam opracowania dezambiguizacyjne dotyczące trzech francuskich predykatów czasownikowych: *monter* (Żłobińska-Nowak 2008), *entrer* (Żłobińska-Nowak 2014), *remettre* (Hrabia 2018c) oraz przymiotnika *farouche* (Hrabia 2011). We wskazanych pracach wyłonione i sformalizowane z użyciem klas obiektowych zostały następujące znaczenia/ekwiwalenty analizowanych jednostek<sup>7</sup>:

- 6 W ujęciu zorientowanym obiektowo bardzo ważną rolę odgrywa hierarchia klas, zgodnie z którą podklasy (klasy hiponimiczne, np. <kot>) dziedziczą operatory z nadklas (klas hiperonimicznych, np. <zwierzę>) (zob. Śmigielska 2004, Hrabia 2018b). Szczegółowe omówienie tego zagadnienia wykracza jednak poza ramy niniejszego artykułu.
- 7 W metodologii dezambiguizacji zaproponowanej przez Banysia znaczenie w języku A i ekwiwalent w języku B są pojęciami niejako tożsamymi, gdyż z punktu widzenia operacyjnego „dane słowo ma tyle znaczeń w języku źródłowym, ile



- dla czasownika *monter*: wejść/wchodzić, wjechać/wjeżdżać, wspinać się, płynąć w górę, wnieść się/wznosić się, iść/pójść, jechać/jeździć (na zwierzęciu), wsiąść/wsiadać, dosiąść/dosiadać, pokryć/pokrywać, pojechać (do), awansować, dojść/dochodzić, napłynąć/napływać, wezbrać/wzbierać, wejść/wschodzić, piąć się w górę, prowadzić (do, w kierunku), rósć/rosnąć, podnieść/podnosić (się), sięgnąć/sięgać, wzrósć/wzrastać, iść w górę, drożyć/podrożyć, wynieść/wynosić, stroić, założyć/zakładać, wystawić/wystawiać, ustawić/ustawiać, (z)montować, wprawić/wprawiać (do), zaopatrzyć/zaopatrywać, unieść/unosić się, po którym (ławo, trudno) się wchodzi;
- dla czasownika *entrer*: wejść/wchodzić (do, przez, na), wpływać/wpływać (do), nadejść/nadchodzić (od), wstąpić/wstępować (do), wjechać/wjeżdżać;
- dla czasownika *remettre*: włożyć/wkładać z powrotem (ponownie), założyć/zakładać z powrotem (ponownie), położyć z powrotem (ponownie), zanurzyć/zanurzać ponownie (z powrotem), wsadzić/wsadzać z powrotem (ponownie), dolać/dolewać, dosypać/dosypywać, dołożyć/dokładać, odłożyć/odkładać, odprowadzić/odprowadzać, włączyć/włączać z powrotem (ponownie), nastawić/nastawiać, zamknąć/zamykać, dostarczyć/dostarczać, oddać/oddawać, powierzyć/powierzać, przyznać/przyznawać, darować/darowywać, przywrócić/przywracać, pogodzić/godzić, poznać/poznawać;
- dla przymiotnika *farouche*: płochliwy, dziki, zagorzały, okrutny, wrogi, przeraźliwy, groźny, niepohamowany, niewzruszony, zaciekły, burzliwy, pierwotny, nieprzystępny, nietowarzystki, nieokrzesany, nieprzyjazny.

Z każdego opracowania wybraliśmy reprezentatywne przykłady użyc (zdania lub rzadziej zwroty) egzemplifikujące wszystkie wyżej wymienione znaczenia (74 przykłady dla *monter*, 11 dla *entrer*, 50 dla *remettre* i 45 dla *farouche*). Następnie wygenerowaliśmy ich tłumaczenie na język polski, używając aplikacji Google Translate (<https://translate.google.pl/> [data dostępu: 05.07.2024]) i ChatuGPT-3.5 (<https://chatgpt.com/> [data dostępu: 05.07.2024]). ChatGPT otrzymał prostą komendę: „Przetłumacz poniższe zdania lub zwroty z języka francuskiego na język polski”.

W zdecydowanej większości efekty pracy wykonanej przez AI można uznać za zadowalające, jeżeli jako kryterium sukcesu przyjmiemy ogólną zrozumiałość komunikatu. Mniejsze lub większe błędy lub niezręczności tłumaczeniowe były jednak dość liczne. Po dokonaniu ich analizy pogrupowaliśmy je w 4 kategorie:

- błędy semantyczne: kardynalne, nieadekwatności semantycznej, ubytku i naddatku semantycznego;
- błędy leksykalne;
- błędy fleksyjne i składniowe;
- błędy stylistyczne.

## Typologia i ocena błędów AI

174

O błędach w komunikacji interlingwalnej napisano już wiele (zob. np. Dzierżanowska 1990; Pisarska, Tomaszewicz 1996; Kozłowska 2002; Hejwowski 2004, 2009; Paprocka 2005; Delisle *et al.* 2006; Matulewska 2014). Wszystkie typologie stworzone na gruncie tak rodzimego, jak i zagranicznego przekładoznawstwa dotyczyły – rzecz jasna – błędów popełnianych przez tłumaczy-ludzi. Niektóre z wyróżnionych kategorii, jak np. błędy metodologiczne (metatranslacyjne), a więc takie, które wynikają z niezastosowania właściwych zasad, reguł czy technik tłumaczeniowych (por. Delisle *et al.* 2006: 29, Hejwowski 2004: 158), nie mogą znaleźć zastosowania przy ocenie przekładu wykonanego przez AI, gdyż sztuczna inteligencja żadnych zasad czy reguł translatorskich nie stosuje „świadomie”. Absurdalne też byłoby stwierdzenie, że wadliwość przekładu wykonanego przez programy bazujące na AI wynika np. z ich niedostatecznej znajomości języka wyjściowego lub docelowego.

Przy tworzeniu typologii błędów na potrzeby niniejszego opracowania najbardziej pomocny okazał się ogólny podział nieprawidłowości przekładowych na błędy językowe (fr. *erreurs de langue*) tj. takie, które są widoczne bez porównywania z tekstem źródłowym oraz błędy tłumaczeniowe sensu stricto (błędy transferu, fr. *erreurs de transfert*), które można wykryć dopiero przy komparacji przekładu z oryginałem (por. np. Kozłowska 2002: 198, Paprocka 2005: 55, Delisle *et al.* 2006: 29–30)<sup>8</sup>. Do kategorii błędów językowych zaliczają się wydzielone przez nas błędy nieadekwatności semantycznej, błędy fleksyjno-składniowe, leksykalne oraz błędy rejestru. Błędy tłumaczeniowe zaś to błędy semantyczne kardynalne oraz błędy zubożenia lub nadmiaru semantycznego.

### Błędy semantyczne kardynalne

Pierwsza wydzielona kategoria obejmuje błędy kardynalne, które w istotnym stopniu wypaczają sens oryginału lub uniemożliwiają zrozumienie przekładu. Choć odnaleźliśmy je w niewielkiej liczbie analizowanych przykładów, zaskakujący wydaje się sam fakt ich występowania w tłumaczeniach wykonanych przez AI.

Przyjrzyjmy się kilkorgu z nich:

|   | Słowo<br>odwieloznaczane<br>w AOO | Ekwiwalent<br>wyznaczony<br>w AOO | Przykład frazy/<br>zdania w języku<br>źródłowym | Adekwatne<br>tłumaczenie<br>w języku<br>docelowym | Tłumaczenie<br>wygenerowane<br>przez AI                  |
|---|-----------------------------------|-----------------------------------|-------------------------------------------------|---------------------------------------------------|----------------------------------------------------------|
| 1 | monter                            | sięgnąć /<br>sięgać               | Une robe qui<br>monte jusqu'au<br>cou.          | Sukienka, która sięga<br>(aż) po szyję.           | <b>Google:</b><br>Sukienka <u>zapinana</u><br>pod szyję. |

8 Jak zauważa Paprocka (2005: 55–57), przyporządkowanie niektórych błędów tłumaczeniowych do jednej z tych dwóch ogólnych kategorii może być problematyczne. Postuluje ona odejście od sztywnych podziałów i wyróżnienie kategorii pośredniej, obejmującej błędy dostrzegalne bez porównania z tekstem źródłowym (co jest charakterystyczne dla *erreurs de langue*), które jednocześnie – podobnie jak *erreurs de transfert* – prowadzą do zniekształcenia sensu oryginału (Paprocka 2005: 93). To przenikanie się różnych typów błędów uwidacznia się w pewnym stopniu także w naszej typologii.

|    | Słowo<br>odwieloznaczane<br>w AOO | Ekwiwalent<br>wyznaczony<br>w AOO | Przykład frazy/<br>zdania w języku<br>źródłowym                                     | Adekwatne<br>tłumaczenie<br>w języku<br>docelowym                                            | Tłumaczenie<br>wygenerowane<br>przez AI                                                        |
|----|-----------------------------------|-----------------------------------|-------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------|
| 2  | monter                            | iść w górę                        | Si les fruits montent, tu manges par exemple plus de yaourts et moins de fruits.    | Jeśli owoce idą w górę, jesz na przykład więcej jogurtów, a mniej owoców.                    | <b>Google:</b> <u>Jeśli na przykład jest dużo owoców, jesz więcej jogurtu, a mniej owoców.</u> |
| 3  | monter                            | stroić                            | monter une guitare, une harpe                                                       | stroić gitarę, harfę                                                                         | <b>GPT:</b> <u>składać gitarę, harfę.</u><br><b>Google:</b> <u>zbuduj gitarę, harfę</u>        |
| 4  | se monter                         | zaopatrzyć /<br>zaopatrywać       | se monter en matériel de jardin                                                     | Zaopatrzyć się w sprzęt ogrodniczy.                                                          | <b>Google:</b> <u>korzystać ze sprzętu ogrodniczego</u>                                        |
| 5  | entrer                            | wstąpić do                        | Un pasteur protestant entre chez un coiffeur italien et se fait couper les cheveux. | Protestancki pastor wstępuje do włoskiego fryzjera i obcina włosy (każe sobie obciąć włosy). | <b>Google:</b> <u>Protestancki pastor przychodzi do włoskiego fryzjera i obcina mu włosy.</u>  |
| 6  | monter                            | wznieść się /<br>wznosić się      | Une hirondelle monte dans l'air.                                                    | Jaskółka wznosi się w powietrze.                                                             | <b>GPT:</b> <u>Jeden jaskółka nie czyni wiosny.</u>                                            |
| 7  | remettre                          | odprowadzić /<br>odprowadzać      | On a réussi à remettre les vaches et les moutons au clos.                           | Udało się nam odprowadzić krowy i owce do zagrody.                                           | <b>GPT:</b> <u>Udało nam się ponownie wyprowadzić krowy lub owce na pastwisko.</u>             |
| 8  | remettre                          | darować /<br>darowywać            | remettre une partie de sa peine à un condamné                                       | darować część kary skazanemu                                                                 | <b>GPT:</b> <u>odroczyć część kary skazanemu</u>                                               |
| 9  | remettre                          | poznać /<br>poznawać              | Je le remets maintenant, c'est Durand.                                              | Teraz go rozpoznaję, to Durand.                                                              | <b>Google:</b> <u>Odłóż to teraz z powrotem, to Durand.</u>                                    |
| 10 | farouche                          | niewzruszony                      | Il prononça ces derniers mots avec une gravité farouche. (Bernanos)                 | Wypowiedział te ostatnie słowa z niewzruszoną powagą.                                        | <b>GPT:</b> <u>Wypowiedział te ostatnie słowa z <u>dziką</u> powagą.</u>                       |

1. *Sukienka, która sięga po szyję*, niekoniecznie musi być *zapinana*. Między sensem oryginału i przekładu istnieje więc znaczna rozbieżność, która skutkować będzie błędnym zrozumieniem komunikatu przez odbiorcę posługującego się językiem docelowym.
2. Fakt, że *owoce idą w górę* jednoznacznie wskazuje na wzrost ich ceny, a nie liczby. Przekład zmienia zatem całkowicie treść komunikatu.
3. Oba narzędzia – Google Translate i ChatGPT – podały błędną wersję tłumaczenia. Być może było to jednak spowodowane brakiem szerszego kontekstu.

4. *Korzystanie ze sprzętu a zaopatrywanie się w niego* denotują dwie różne sytuacje pozajęzykowe. Tłumaczenie zaproponowane przez AI jest zatem całkowicie niewłaściwe.
5. Ekwiwalent czasownika *entrer* – *przychodzić (do)* – uznać można za względnie akceptowalny. Przekład całego zdania jest jednak jednoznacznie niepoprawny, gdyż zostają w nim odwrócone role semantyczne pacjensa i agensa. System błędnie zinterpretował strukturę gramatyczną *se faire faire*, która wskazuje, że akcja jest wykonywana na podmiocie.
6. Tłumaczenie wydaje się przykładem halucynacji ChatuGPT. Po rozpoznaniu w pozycji inicjalnej zdania syntagmy *une hirondelle* system prawdopodobnie „przewidział” najbardziej typowe i frekwencyjnie najczęstsze zakończenie: *ne fait pas le printemps* i w rezultacie wygenerował przekład znanego powiedzenia.
7. Zaproponowane tłumaczenie jest semantycznie antonimiczne względem oryginału.
8. Istnieje znaczna różnica między *darowaniem* a *odroczeniem kary*. Zaproponowany przekład uznać należy więc jednoznacznie za błędny.
9. Tłumaczenie jest całkowicie niezrozumiałe. Google Translate nie rozpoznał mniej typowego znaczenia czasownika *remettre*.
10. Syntagma *dzika powaga* wypacza znacząco sens oryginału, gdyż w żaden sposób nie wskazuje na niewzruszoność denotowanego stanu.

### Błędy semantyczne: nieadekwatność semantyczna

Do tej kategorii proponujemy zaliczyć takie przykłady, w których zaproponowane przez AI tłumaczenie zmienia nieznacznie sens oryginału, ale umożliwia, po chwili zastanowienia, zasadniczo poprawne zrozumienie komunikatu jako całości. Wybrane dla francuskich polisemów polskie ekwiwalenty są jednak ewidentnie niewłaściwe z uwagi na ich semantyczną niekompatybilność z pozostałymi elementami zdania i brzmią w podanych kontekstach całkowicie nienaturalnie dla rodzimego użytkownika polszczyzny. Błędy tego typu, których uniknięcie okazuje się dużym wyzwaniem dla systemów tłumaczeniowych opartych na AI, są stosunkowo proste do wyeliminowania w sformalizowanych „ręcznie” dezambiguacyjnych opisach języka np. poprzez wskazanie odpowiednich klas obiektowych w pozycji podmiotu i dopełnienia.

Oto kilka przykładów zaczerpniętych z korpusu:

|   | Słowo odwieloznacz-<br>niane w AOO | Ekwiwalent<br>wyznaczony<br>w AOO | Przykład frazy/zdania<br>w języku źródłowym                                                                                                                                       | Adekwatne<br>tłumaczenie<br>w języku docelowym                                                                                                                                     | Tłumaczenie<br>wygenerowane przez<br>AI                                                                                                                                                                   |
|---|------------------------------------|-----------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 | entrer                             | wpływać (do)<br>(od)              | Le fleuve la Seine<br>passe au nord-est du<br>territoire, elle entre<br>par l'est au Coudray<br>Montceaux et<br>parcourt vingt-quatre<br>kilomètres jusqu'à<br>Vigneux-sur-Seine. | Sekwana płynie<br>na północny<br>wschód od (tego)<br>terytorium, wpływa<br>od wschodu do<br>Coudray Montceaux<br>i płynie dwadzieścia<br>cztery kilometry do<br>Vigneux-sur-Seine. | <b>GPT:</b> Rzeka<br>Sekwana przepływa<br>na północny wschód<br>od terytorium,<br>wchodzi od<br>wschodu w Coudray<br>Montceaux<br>i przepływa<br>dwadzieścia cztery<br>kilometry do<br>Vigneux-sur-Seine. |

|   | Słowo odwieloznacz-<br>niane w AOO | Ekwiwalent<br>wyznaczony<br>w AOO   | Przykład frazy/zdania<br>w języku źródłowym                                                                 | Adekwatne<br>tłumaczenie<br>w języku docelowym                                                   | Tłumaczenie<br>wygenerowane przez<br>AI                                                                  |
|---|------------------------------------|-------------------------------------|-------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------|
| 2 | entrer                             | nadejść (do)<br>od                  | Vendredi une<br>perturbation océanique<br>entre par l'ouest.                                                | W piątek od zachodu<br>nadchodzą zakłócenia<br>oceaniczne.                                       | <b>GPT:</b> W piątek<br>zakłócenie oceaniczne<br><u>wchodzi</u> od zachodu.                              |
| 3 | entrer                             | wjechać<br>(do) (od/<br>czym)       | On ne peut entrer dans<br>ce pays sans visa.                                                                | Nie można wjechać<br>do tego kraju bez<br>wizy.                                                  | <b>GPT:</b> Nie można<br><u>wejść</u> do tego kraju<br>bez wizy.                                         |
| 4 | remettre                           | położyć<br>z powrotem<br>(ponownie) | remettre un livre sur<br>son rayon                                                                          | położyć z powrotem/<br>odłożyć książkę na<br>półce                                               | <b>GPT :</b> <u>włożyć</u> książkę<br>z powrotem na półkę                                                |
| 5 | remettre                           | położyć<br>z powrotem<br>(ponownie) | remettre un enfant au lit                                                                                   | położyć dziecko<br>z powrotem do łóżka                                                           | <b>GPT:</b> <u>włożyć</u> dziecko<br>z powrotem do łóżka                                                 |
| 6 | remettre                           | odprowadzić<br>/<br>odprowadzać     | remettre un cheval à<br>l'écurie                                                                            | odprowadzić konia<br>do stajni                                                                   | <b>GPT:</b> <u>włożyć</u> konia<br>do stajni                                                             |
| 7 | farouche                           | okrutny                             | C'est qu'en effet ce roi<br>farouche règne surtout<br>par la terreur.                                       | Ten okrutny król<br>rzeczywiście rządzi<br>głównie poprzez<br>strach.                            | <b>GPT:</b> Faktycznie<br>ten <u>dziki</u> król panuje<br>głównie poprzez<br>strach.                     |
| 8 | farouche                           | nieprzystępny                       | Rosetta n'est pas une<br>femme farouche et elle<br>devient très rapidement<br>la maîtresse du<br>narrateur. | Rosetta nie jest<br>nieprzystępną kobietą<br>i bardzo szybko<br>staje się kochanką<br>narratora. | <b>GPT:</b> Rosetta nie jest<br><u>dziką</u> kobietą i bardzo<br>szybko staje się<br>kochanką narratora. |

- 1-3. We wszystkich tych przypadkach ChatGPT wybiera najbardziej neutralny znaczeniowo ekwiwalent *entrer* – *wchodzić*. Choć przekaz komunikatu jest zrozumiały, tłumaczeń zaproponowanych przez AI nie można uznać za właściwe. Wszak obiekt językowy należący do klasy <ciek wodny> (np. *rzeka*, *strumień*), wykonujący akcję predykowaną przez francuskie *entrer*, nie *wchodzi*, a *wpływa*, <zjawiska meteorologiczne> *nadchodzą*, *napływają*, a do obiektu z klasy <kraj> się *wjeżdża*, a nie *wchodzi*.
- 4-5. Tłumaczenie czasownika *remettre* jako *włożyć z powrotem* samo w sobie jest poprawne. W przykładzie 4 okolicznik miejsca *sur le rayon* (pl. *na półkę*) wyklucza jednak możliwość wyboru takiego ekwiwalentu, gdyż *włożyć* można tylko *coś w coś/do czegoś*. Z kolei w zdaniu 5 wybór ten jest blokowany przez obecność określonych klas obiektowych na pozycjach dopełnienia i okolicznika (odpowiednio: <dziecko> i <obiekt, na którym się śpi>).
6. Fraza *włożyć konia do stajni* może wywołać uśmiech. Ponadto w tłumaczeniu ztraca się wartość semantyczna powrotu do stanu uprzedniego (w tym przypadku: 'wcześniejszego przebywania konia w stajni') wprowadzana przez francuski prefiks *re-*.
- 7-8. Podobnie jak w przykładach 1-3 ChatGPT wybrał frekwencyjnie najczęstszy ekwiwalent. W obu zdaniach odbiorca jest w stanie odtworzyć sens oryginalnego komunikatu, niemniej syntagmy *dziki król* i *dzika kobieta* uznać trzeba za nienaturalne w podanym kontekście i/

lub nieprecyzyjne z uwagi chociażby na dużą polisemiczność przymiotnika *dziki*. Obiekt językowy z klasy <władca> określany jako *farouche* jest tak naprawdę *władcą okrutnym*, z kolei *femme farouche* to *kobieta nieprzystępna*.

### Błędy semantyczne: zubożenie i naddatek semantyczny

Do błędów zubożenia lub naddatku semantycznego zaliczymy takie nieprawidłowości w przekładzie, które powodują nieznaczną zmianę sensu oryginału, ale – w przeciwieństwie do błędów nieadekwatności semantycznej – są możliwe do zaobserwowania dopiero przy porównaniu tekstu źródłowego i docelowego. Niewłaściwie dobrany ekwiwalent jest bowiem w tym przypadku semantycznie kompatybilny z pozostałymi elementami zdania w tekście docelowym, a modyfikacja znaczenia polega wyłącznie na nieuzasadnionym użyciu hiponimu lub hiperonimu w tłumaczeniu, tj. zawężeniu (przykłady 1-4) lub rozszerzeniu intensji (przykład 5) słowa wyjściowego.

|   | Słowo odwieloznacz-<br>niane w AOO | Ekwiwalent<br>wyznaczony<br>w AOO | Przykład frazy/<br>zdania w języku<br>źródłowym                                           | Adekwatne<br>tłumaczenie w języku<br>docelowym                                                         | Tłumaczenie<br>wygenerowane przez<br>AI                                                                                            |
|---|------------------------------------|-----------------------------------|-------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------|
| 1 | monter                             | piąć się<br>w górę                | Le chemin monte<br>jusqu'aux alpages.                                                     | Ścieżka pnie się w górę<br>aż do hal.                                                                  | <b>Google:</b><br>Ścieżka prowadzi na<br>górskie pastwiska.                                                                        |
| 2 | monter                             | wzniesić się/<br>wznosić się      | Les flammes montent<br>de la maison en feu.                                               | Z płonącego domu<br>wznoszą się płomienie.                                                             | <b>GPT:</b> Płomienie<br>wydobywają się<br>z płonącego domu.                                                                       |
| 3 | entrer                             | wejść (do,<br>przez, na)          | Demain j'entrerais à<br>la cuisine, sous un<br>prétexte quelconque.                       | Jutro pod jakimś<br>pretekstem wejdę do<br>kuchni.                                                     | <b>Google:</b> Jutro <u>pójdę</u><br>do kuchni, pod jakimś<br>pretekstem.                                                          |
| 4 | remettre                           | dosypać /<br>dosypywać            | Aujourd'hui, ma<br>mère me demande de<br>remettre du sucre dans<br>la boîte bientôt vide. | Dzisiaj moja mama<br>prosi mnie, żebym<br>dosypała cukru do<br>pudełka, które wkrótce<br>będzie puste. | <b>Google:</b> Dzisiaj moja<br>mama prosi mnie,<br>żebym <u>wsypała</u> trochę<br>cukru do pudełka, które<br>wkrótce będzie puste. |

1. W przekładzie dokonany przez Tłumacza Google sens oryginału jest w dużej mierze zachowany, jednak pominięta została ważna dla czasownika *monter* cecha semantyczna 'podążania w górę'.
2. Podobnie jak w przykładzie poprzednim tłumaczenie nie uwzględnia wartości semantycznej wykonywania akcji 'w górę'.
3. Różnica między *pójściem* gdzieś a *wejściem* jest wyraźna i choć przekład brzmi w pełni naturalnie, nie akcentuje istotnej semantycznie opozycji 'być na zewnątrz' (przed akcją predykowaną przez *entrer*) vs 'być wewnątrz' (po akcji).
4. Wartość perlokucyjna przetłumaczonego komunikatu nie ulega zmianie, gdyż odbiorczyni słów zapewne wykona prawidłową akcję. Iteracyjny charakter prefiksu *re-*, którego

ekwiwalentem w tym kontekście jest przedrostek *do-* (por. Śmiech 1986: 92), zanika jednak w przekładzie całkowicie.

|   | Słowo odwieloznacz-<br>niane w AOO | Ekwiwalent<br>wyznaczony<br>w AOO | Przykład frazy/<br>zdania w języku<br>źródłowym                | Adekwatne<br>tłumaczenie w języku<br>docelowym                           | Tłumaczenie<br>wygenerowane przez<br>AI                                                |
|---|------------------------------------|-----------------------------------|----------------------------------------------------------------|--------------------------------------------------------------------------|----------------------------------------------------------------------------------------|
| 5 | monter                             | wspinać się                       | On monte une côte<br>assez raide pour<br>arriver à ce village. | Wspinamy się po<br>stromym wzniesieniu,<br>aby dotrzeć do tej<br>wioski. | <b>GPT:</b> <u>Wchodzimy</u><br>stromym<br>wzniesieniem, aby<br>dotrzeć do tej wioski. |

5. Z tłumaczenie zaproponowanego przez ChatGPT wprost wynika, że wykonawcą czynności musi być żywa istota. Tymczasem oryginał na to nie wskazuje – akcję wyrażoną czasownikiem *monter* (poprawnie tłumaczonym w tym kontekście jako *wspinać się*) może wykonywać zarówno istota żywa, jak i np. środek transportu.

### Błędy leksykalne

Ta niewielka kategoria obejmuje ekwiwalenty, które nie istnieją w języku polskim. Błędy tego typu były popełniane jedynie przez ChatGPT. Warto zaznaczyć, że nie są to barbarzyństwa, gdyż bez wątpienia nie powstały pod wpływem leksyki języka wyjściowego.

|   | Słowo odwieloznacz-<br>niane w AOO | Ekwiwalent<br>wyznaczony<br>w AOO | Przykład frazy/<br>zdania w języku<br>źródłowym                                                                                                | Adekwatne<br>tłumaczenie w języku<br>docelowym                                                                                             | Tłumaczenie<br>wygenerowane<br>przez AI                                                                                                                   |
|---|------------------------------------|-----------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 | farouche                           | zaciekły                          | L'Église romaine<br>menait en parallèle<br>un combat farouche<br>contre les théories<br>politiques modernes<br>(libéralisme et<br>socialisme). | Kościół rzymski<br>prowadził jednocześnie<br>zacieklą walkę<br>ze nowożytnymi<br>teoriami politycznymi<br>(liberalizmem<br>i socjalizmem). | <b>GPT:</b><br>Kościół rzymski<br>toczył równoległą<br><u>zmagę</u> przeciwko<br>nowoczesnym<br>teoriom politycznym<br>(liberalizmowi<br>i socjalizmowi). |
| 2 | monter                             | dosiąść/<br>dosiadać              | monter un beau<br>destrier                                                                                                                     | dosiadać pięknego<br>rumaka                                                                                                                | <b>GPT:</b><br><u>Wsiedzieć</u> na<br>pięknego rumaka.                                                                                                    |

### Błędy fleksyjne i składniowe

Błędy te często współlistnieją z błędami semantycznymi, ale ze względu na ich wyrazisty charakter postanowiliśmy umieścić je w osobnej kategorii. O ile pewne niezręczności natury syntaktycznej nie są szczególnie zaskakujące (przykład 4), o tyle problemy z polską deklinacją (przykłady 1-2) budzą zdziwienie, gdyż, jak mogłoby się wydawać, wyekstrahowanie reguł fleksji – licznych, ale jednak

skończonych – nie powinno nastroczać trudności sztucznej inteligencji przeszkolonej na ogromnej liczbie tekstów.

|   | Słowo odwieloznacz-<br>niane w AOO | Ekwiwalent<br>wyznaczony<br>w AOO | Przykład frazy/<br>zdania w języku<br>źródłowym                                                               | Adekwatne<br>tłumaczenie<br>w języku<br>docelowym                                                       | Tłumaczenie<br>wygenerowane przez AI                                                                              |
|---|------------------------------------|-----------------------------------|---------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------|
| 1 | farouche                           | dziki                             | Nous faisons la connaissance d'une tribu farouche.                                                            | Poznajemy dzikie plemię.                                                                                | <b>GPT:</b> Poznajemy <u>dziką</u> plemię.                                                                        |
| 2 | monter                             | wejść /<br>wchodzić               | Montez la rampe à droite et prenez le passage de droite.                                                      | Wejdź na rampę po prawej stronie i skręć w prawo.                                                       | <b>GPT:</b> Wejdź na rampę po prawej stronie i <u>weź</u> <u>prawą przejście</u> .                                |
| 3 | monter                             | wjechać /<br>wjeżdżać             | monter avec l'ascenseur                                                                                       | wjechać windą                                                                                           | <b>Google:</b> <u>jedź</u> windą na górę                                                                          |
| 4 | monter                             | prowadzić (do,<br>w kierunku)     | Sous les voûtes, les mêmes escaliers babyloniens montaient vers les temples qu'avaient abandonnés les nuages. | Pod sklepieniami te same babilońskie schody prowadziły do opuszczonych/porzuconych przez chmury świątyń | <b>Google:</b> Pod sklepieniami te same babilońskie schody prowadziły do świątyń, <u>które porzuciły chmury</u> . |
| 5 | entrer                             | wejść (do,<br>przez, na)          | Entrer en chaire, à l'autel.                                                                                  | Wejść na ambonę, na ołtarz                                                                              | <b>GPT/Google:</b> Wejść na ambonę, <u>do ołtarza</u> .                                                           |

1. ChatGPT potraktował *plemię* jak rzeczownik rodzaju żeńskiego.
2. Tłumaczenie *weź prawą przejście* jest oczywiście całkowicie niepoprawne pod względem semantycznym, ale nawet abstrahując od semantyki, uderzająca jest tutaj błędna odmiana przymiotnika *prawe*. GPT ponownie uzgodnił przymiotnik do rodzaju żeńskiego zamiast nijakiego.
3. Z niewyjaśnionych powodów Google zaproponował tłumaczenie w formie trybu rozkazującego, podczas gdy komunikat oryginalny podany jest w bezokoliczniku.
4. W przykładzie francuskim w zdaniu podrzędnym względnym użyto inwersji stylistycznej (*qu'avaient abandonnés les nuages* zamiast *que les nuages avaient abandonnés*). Google zaproponował tłumaczenie słowo w słowo, co doprowadziło do wieloznaczności składniowej w polskiej wersji.
5. Łatwo dostrzec w tłumaczeniu błędny wybór przyimka. *Wchodzimy* oczywiście *na ołtarz*, a nie *do ołtarza*.

### Błędy stylistyczne: błędy rejestru

W analizowanym korpusie odnaleźliśmy jeden przykład, który zaliczyć można do kategorii błędów rejestru.



|   | Słowo odwieloznacz-<br>niane w AOO | Ekwiwalent<br>wyznaczony<br>w AOO | Przykład frazy/<br>zdania w języku<br>źródłowym | Adekwatne<br>tłumaczenie<br>w języku docelowym | Tłumaczenie<br>wygenerowane przez<br>AI           |
|---|------------------------------------|-----------------------------------|-------------------------------------------------|------------------------------------------------|---------------------------------------------------|
| 1 | monter                             | kryć / pokryć                     | L'éta lon monte la<br>jument.                   | Ogier kryje klacz.                             | <b>GPT:</b><br>Ogier <u>kopuluje</u><br>z klaczą. |

Sens zdania źródłowego jest w przekładzie zaproponowanym przez ChatGPT w pełni zachowany. Wydaje się jednak, że właściwszym stylistycznie ekwiwalentem dla *monter* byłby należący do technolektu weterynaryjno-hodowlanego termin *kryć*, a nie czasownik *kopulować* (fr. *copuler*), charakterystyczny raczej dla standardowego języka naukowego<sup>9</sup>.

### Wnioski i perspektywy

Jak wykazała przeprowadzona analiza, błędy popełniane przez systemy tłumaczeniowe oparte na sztucznej inteligencji odnoszą się zasadniczo do wszystkich poziomów organizacji języka – od fleksji, przez składnię, semantykę, aż po stylistykę. W badanym korpusie odnaleźliśmy sporo uchybień, choć trzeba raz jeszcze podkreślić, że błędy kardynalne, uniemożliwiające zrozumienie komunikatu źródłowego bądź znacznie wypaczające jego sens, stanowiły wyraźną mniejszość. Wynikały one być może z ograniczonego kontekstu, który był przetwarzany przez AI. Ocenianie umiejętności przekładowych na poziomie pojedynczych zdań nie jest jednak bezzasadne, gdyż użytkownicy korzystający z programów do tłumaczenia automatycznego nie zawsze przecież oczekują przekładu całości koherentnego i kohezyjnego tekstu; często powierzają sztucznej inteligencji zadania tłumaczeniowe, które ograniczają się do jednej frazy lub krótkiej syntagmy.

Oba systemy popełniały błędy w zasadzie z równą częstotliwością. Wbrew swoim zapewnieniom ChatGPT 3.5 nie jest więc na pewno lepszym tłumaczem od Google Translate. Wykazywał się zresztą mniejszą umiejętnością odwieloznaczania słów, co częściej niż w przypadku Tłumacza Google prowadziło do błędów nieadekwatności semantycznej w wykonanym przekładzie.

Na podstawie naszego badania trudno kwantytatywnie ocenić efektywność i poprawność tłumaczenia AI w ujęciu globalnym. Korpus, na którym pracowaliśmy, nie był bowiem korpusem zrównoważonym, co oznacza, że jednakowo traktowaliśmy użycia o różnej frekwencyjności. Bez większych wątpliwości przyjąć można, że liczba błędów byłaby relatywnie mniejsza, gdyby korpus uwzględniał rzeczywistą częstość występowania poszczególnych znaczeń danego słowa w języku źródłowym. Użycia statystycznie dominujące, bardziej „typowe” stanowią bowiem mniejsze wyzwanie w tłumaczeniu automatycznym.

Warto oczywiście pamiętać, że sztuczna inteligencja, choć nie jest już melodią przyszłości, wciąż ku tej przyszłości dąży, rozwijając się nieustannie i zwiększając swoje możliwości niemalże z dnia na dzień. Badania, których wyniki zostały przedstawione w niniejszym artykule weryfikowały umiejętności przekładowe ChatuGPT 3.5, gdyż podczas ich przeprowadzania tylko wersja 3.5 była powszechnie i darmowo dostępna dla wszystkich użytkowników. Nowe odsłony GPT istnieją już teraz, a w momencie ukazania się tekstu drukiem, istnieć będą z pewnością kolejne, jeszcze bardziej efektywne

<sup>9</sup> Pomijamy tutaj dyskusję na temat zależności między językiem fachowym (technolektem) a językiem naukowym (zob. np. Jarosz 2018).

translacyjnie. Wskazane wydaje się nam obserwowanie zmian tej efektywności w dłuższej perspektywie i przeprowadzenie za jakiś czas podobnej analizy na tym samym materiale korpusowym.

## Bibliografia

- Banyś, Wiesław (2002a) „Bases de données lexicales électroniques – une approche orientée objets. Partie I: Questions de modularité”. [W:] *Neophilologica*. Vol. 15; 7-29.
- Banyś, Wiesław (2002b) „Bases de données lexicales électroniques – une approche orientée objets. Partie II: Questions de description”. [W:] *Neophilologica*. Vol. 15; 206–248.
- Banyś, Wiesław (2005) „Désambiguïsation des sens des mots et représentation lexicale du monde”. [W:] *Neophilologica*. Vol. 17; 57–76.
- Chomsky, Noam (1956) „Three Models for the Description of Language”. [W:] *IRE Transactions on Information Theory*. Vol. 2/3; 113–124.
- Delisle, Jean, Hannelore Lee-Jahnke, Monique C. Cormier [przekład i adaptacja: Teresa Tomaszewicz] (2006) *Terminologia tłumaczenia*. Poznań: Wydawnictwo Naukowe UAM.
- Drzazga, Anna, Aleksandra Żłobińska-Nowak (2014) „Word Sense Disambiguation according to the Object Oriented Approach by W. Banyś”. [W:] *Linguistica Silesiana*. Vol. 35; 191–205.
- Dzierżanowska, Halina (1990) *Przekład tekstów nieliterackich*. Warszawa: Wydawnictwo Naukowe PWN.
- Gabrysiak, Katarzyna (2009) „Quelques remarques sur la désambiguïsation du verbe *mettre*”. [W:] *Neophilologica*. Vol. 20 ; 71–90.
- Gross, Gaston (1994) „Classes d’objets et description des verbes”. [W:] *Langages*. Vol. 115; 15–30.
- Gross, Gaston (2008) „Les classes d’objets”. [W:] *Lalies*. Vol. 28; 111–165.
- Gross, Gaston (2012) *Manuel d’analyse linguistique*. Villeneuve d’Ascq: Septentrion.
- Hejwowski, Krzysztof (2009) „Klasyfikacja błędów tłumaczeniowych – teoria i praktyka”. [W:] Andrzej Kopczyński, Magdalena Kizeweter (red.) *Jakość i ocena tłumaczenia*. Warszawa: Wydawnictwo Szkoły Wyższej Psychologii Społecznej „Academia”; 141–161.
- Hrabia, Michał (2011) „Désambiguïsation des sens du prédicat adjectival *farouche* dans le cadre d’une approche orientée objets „. [W:] *Neophilologica*. Vol. 23; 150–167.
- Hrabia, Michał (2018a) „Kocie frazeologizmy w tłumaczeniu automatycznym francusko-polskim”. [W:] Ewa Borkowska, Andrzej Borkowski, Maria Długolecka-Pietrzak, Sławomir Sobieraj (red.) *Kot w literaturze, kulturze, języku i mediach*. Siedlce: Wydawnictwo Naukowe IKR[i]BL; 339–352.
- Hrabia, Michał (2018b) „Le rôle de l’héritage sémantique dans la traduction automatique”. [W:] *Romanica Cracoviensia*. Vol. 3/2018; 135–145.
- Hrabia, Michał (2018c) „Verbe *remettre* et ses équivalents polonais. Désambiguïsation et analyse sémantico-préfixale pour les besoins de la traduction automatique”. [W:] *Neophilologica*. Vol. 30; 109–120.
- Jarosz, Beata (2018) „O (nie)tożsamości pojęć język specjalistyczny, język specjalny, język fachowy, język profesjonalny, język zawodowy, profesjolekt, technolekt”. [W:] *Polonica*. Tom 38; 85–108.
- Kozłowska, Zofia (2002) „O błędach językowych w tekstach polskich przekładów”. [W:] Andrzej Kopczyński, Urszula Zaliwska-Okrutna (red.) *Język rodzimy a język obcy, komunikacja, przekład, dydaktyka*. Warszawa: Wydawnictwo Uniwersytetu Warszawskiego; 137–147.
- Lipiński, Krzysztof (2004) *Mity przekładoznawstwa*. Kraków: Wydawnictwo EGIS.
- Matulewska, Aleksandra (2014) „Horror tłumaczeniowy czy tłumacz z piekła rodem? Czyli kilka słów o efektywności komunikacji interlingwalnej”. [W:] *Scripta Neophilologica Posnaniensia*. Nr XIV; 101–118.

- Paprocka, Natalia (2005) *Erreurs en traduction pragmatique du français en polonais*. Łask: Oficyna Wydawnicza LEKSEM.
- Pisarska, Alicja, Teresa Tomaszewicz (1996) *Współczesne tendencje przekładoznawcze*. Poznań: Wydawnictwo Naukowe UAM.
- Śmiech, Witold (1986) *Derywacja prefiksalna czasowników polskich*. Wrocław: Ossolineum.
- Śmigielska, Beata (2004) „Approche orientée objets et hiérarchie linguistique de concepts. Questions d'application”. [W:] *Neophilologica*. Vol. 16; 106–126.
- Żłobińska-Nowak, Aleksandra (2008) *Désambiguïsation des expressions lexicales des opérateurs de l'espace dans le cadre d'une approche orientée objets: les verbes de mouvement „monter” et „sortir” et leurs équivalents polonais*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- Żłobińska-Nowak, Aleksandra (2014) „Entrer/entrar – différences et ressemblances dans leur emploi suivant l'approche orientée objets”. [W:] *Roczniki Humanistyczne*. Tom LXII, zeszyt 8; 133–148.

### Źródła internetowe

- Alammar, Jay (2018) „The Illustrated Transformer”. [Na:] <https://jalammar.github.io/illustrated-transformer/> (data dostępu: 06.06.2024).
- Bahdanau, Dzmitry, Kyunghyun Cho, Yoshua Bengio (2014) „Neural Machine Translation by Jointly Learning to Align and Translate”. [Presented at:] *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA*. [W:] *arXiv:1409.0473*. [Na:] <https://arxiv.org/abs/1409.0473> (data dostępu: 06.06.2024).
- Bogacki, Krzysztof (2023) „Analyse linguistique au service du TAL”. [W:] *Neophilologica*. Vol. 35; 1–13. [Na:] <https://journals.us.edu.pl/index.php/NEO/article/view/15742/12520> (data dostępu: 06.06.2024).
- Cho, Kyunghyun, Bart van Merriënboer, Dzmitry Bahdanau, Yoshua Bengio (2014) „On the Properties of Neural Machine Translation: Encoder–Decoder Approaches”. [W:] Dekai Wu, Marine Carpuat, Xavier Carreras, Eva Maria Vecchi (red.) *Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation*. Doha: Association for Computational Linguistics; 103–111. [Na:] <https://aclanthology.org/W14-4012/> (data dostępu: 06.06.2024).
- Jiao, Wenxiang, Wenxuan Wang, Jen-tse Huang, Xing Wang, Shuming Shi, Zhaopeng Tu (2023) „Is ChatGPT a good translator? Yes with GPT-4 as the engine”. [W:] *arXiv:2301.08745*. [Na:] <https://arxiv.org/abs/2301.08745> (data dostępu: 06.06.2024).
- Johnson, Melvin, Mike Schuster, Quoc V. Le, Maxim Krikun, Yonghui Wu, Zhifeng Chen, Nikhil Thorat, Fernanda Viégas, Martin Wattenberg, Greg Corrado, Macduff Hughes, Jeffrey Dean (2017) „Google's Multilingual Neural Machine Translation System: Enabling Zero-Shot Translation”. [W:] *Transactions of the Association for Computational Linguistics*. Vol. 5; 339–351. [Na:] <https://aclanthology.org/Q17-1024/> (data dostępu: 06.06.2024).
- Jungha, Son, Kim Boyoung (2023) „Translation Performance from the User's Perspective of Large Language Models and Neural Machine Translation Systems”. [W:] *Information*. Vol. 14, issue 10; 1–18. [Na:] <https://www.mdpi.com/2078-2489/14/10/574> (data dostępu: 06.06.2024).
- Kalchbrenner, Nal, Phil Blunsom (2013) „Recurrent Continuous Translation Models”. [W:] *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, Seattle, WA, USA, 10 October 2013*; 1700–1709. [Na:] <https://aclanthology.org/D13-1176/> (data dostępu: 06.06.2024).
- Kalyan, Katikapalli Subramanyam (2024) „A Survey of GPT-3 Family Large Language Models Including Chat GPT and GPT-4”. [W:] *Natural Language Processing Journal*. Vol. 6; 1–48. [Na:] <https://www.sciencedirect.com/journal/natural-language-processing-journal/vol/6/suppl/C> (data dostępu: 06.06.2024).

- Kozłowski, Sebastian (2004) „Co to jest tłumaczenie automatyczne?” [W:] Jakub Szymanik, Marcin Zajenkowski (red.) *Kognitywistyka: o umyśle umyślnie i nieumyślnie*. Warszawa: Koło Filozoficzne przy Kolegium MISH UW; 63–93. [Na:] [https://web.archive.org/web/20140125011838/http://kf.mish.uw.edu.pl:80/kog/kog\\_seb.pdf](https://web.archive.org/web/20140125011838/http://kf.mish.uw.edu.pl:80/kog/kog_seb.pdf) (data dostępu: 06.06.2024).
- Luong, Thang, Hieu Pham, Christopher D. Manning (2015) „Effective Approaches to Attention-Based Neural Machine Translation”. [W:] Lluís Màrquez, Chris Callison-Burch, Jian Su (red.) *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. Lisbon: Association for Computational Linguistics; 1412–1421. [Na:] <https://aclanthology.org/D15-1166/> (data dostępu: 06.06.2024).
- Nazir, Anam, Ze Wang (2023) „A Comprehensive Survey of ChatGPT: Advancements, Applications, Prospects, and Challenges”. [W:] *Meta-Radiology*. Vol. 1(2); 1–12. [Na:] <https://www.sciencedirect.com/journal/meta-radiology/vol/1/issue/2> (data dostępu: 06.06.2024).
- Ray, Partha Pratim (2023) „ChatGPT: A Comprehensive Review on Background, Applications, Key Challenges, Bias, Ethics, Limitations and Future Scope”. [W:] *Internet of Things and Cyber-Physical Systems*. Vol. 3; 121–154. [Na:] <https://www.sciencedirect.com/science/article/pii/S266734522300024X> (data dostępu: 06.06.2024).
- Rybnikova, Julia (2024) „Guide: How to Fine-Tune GPT-3.5 for Translation”. [Na:] <https://custom.mt/guide-how-to-fine-tune-GPT-35-for-translation/> (data dostępu: 06.06.2024).
- Sharma, Sonali, et al. (2023) „Machine Translation Systems Based on Classical-Statistical-Deep-Learning Approaches”. [W:] *Electronics*. Vol. 12/7; 1–29. [Na:] <https://www.mdpi.com/2079-9292/12/7/1716> (data dostępu: 06.06.2024).
- Siu, Sai Cheong (2023) „ChatGPT and GPT-4 for Professional Translators: Exploring the Potential of Large Language Models in Translation”. [W:] *Computational Linguistics & Natural Language Processing eJournal*. Vol. 5/ 24; 1–36. [Na:] [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4448091](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4448091) (data dostępu: 06.06.2024).
- Śmigielska, Beata (2023) „Traduction automatique et désambiguïsation des sens des mots. Le cas du verbe français louer”. [W:] *Neophilologica*. Vol. 35; 1–26. [Na:] <https://journals.us.edu.pl/index.php/NEO/article/view/16289/12729> (data dostępu: 06.06.2024).
- Trinht, Stéphane (2019) „L’intelligence artificielle et la traduction évoluent ensemble”. [W:] *Le Magazine d’Information des Langagiers Circuit*. [Na:] <https://www.circuitmagazine.org/dossier-141/l-intelligence-artificielle-et-la-traduction-evoluent-ensemble> (data dostępu: 06.06.2024).
- Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, Illia Polosukhin (2017) „Attention is All you Need”. [W:] Isabelle Guyon (red.) *Advances in Neural Information Processing Systems*. Vol. 30; 1–11. [Na:] [https://papers.nips.cc/paper\\_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html](https://papers.nips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html) (data dostępu: 06.06.2024).
- Wu, Yonghui, Mike Schuster, Zhifeng Chen, Quoc V. Le, Mohammad Norouzi, Wolfgang Macherey, Maxim Krikun, Yuan Cao, Qin Gao, Klaus Macherey, Jeff Klingner, Apurva Shah, Melvin Johnson, Xiaobing Liu, Łukasz Kaiser, Stephan Gouws, Yoshikiyo Kato, Taku Kudo, Hideto Kazawa, Keith Stevens, George Kurian, Nishant Patil, Wei Wang, Cliff Young, Jason Smith, Jason Riesa, Alex Rudnick, Oriol Vinyals, Greg Corrado, Macduff Hughes, Jeffrey Dean (2016) „Google’s Neural Machine Translation System: Bridging the Gap between Human and Machine Translation”. [W:] *Computing Research Repository (CoRR)*; 1–23. [Na:] <https://arxiv.org/abs/1609.08144> (data dostępu: 06.06.2024).