Joanna Grzybowska
ORCID: 0009-0009-2712-3397
Uniwersytet Wrocławski, Wrocław

Linguistische Treffen in Wrocław, Vol. 28, 2025 (II)
ISSN: 2084–3062, e-ISSN: 2657–5647
https://doi.org/10.23817/lingtreff.28-9
S. 165–174

Kontrastywna analiza językowa z wykorzystaniem modeli sztucznej inteligencji: porównanie języka niemieckiego i innych języków

Kontrastive Sprachanalyse unter Verwendung Künstlicher Intelligenz: Ein Vergleich der deutschen Sprache mit anderen Sprachen

Mit der dynamischen Entwicklung der künstlichen Intelligenz (KI) und der Verarbeitung natürlicher Sprache (Natural Language Processing, NLP) eröffnen sich neue Möglichkeiten in der sprachwissenschaftlichen Analyse, insbesondere im Bereich der kontrastiven und komparativen Forschung. Traditionelle Methoden der Sprachanalyse sind zeitaufwendig und ressourcenintensiv, sodass künstliche Intelligenz zu einem wertvollen Werkzeug in diesem Forschungsbereich wird. Ziel dieser Publikation ist es, einen innovativen Ansatz zur kontrastiven Analyse der deutschen Sprache im Vergleich zu anderen Sprachen unter Verwendung der neuesten KI-Modelle, wie z. B. Transformer-Modelle (BERT, GPT), vorzustellen. Im Rahmen der Analyse werden grammatikalische und semantische Strukturen der Sprachen verglichen sowie der Versuch unternommen, Unterschiede und Gemeinsamkeiten in großem Maßstab automatisiert zu analysieren. Im Vergleich zu bisherigen Studien zeichnet sich der vorgestellte Ansatz durch den Einsatz fortschrittlicher KI-Werkzeuge aus, die eine schnellere und präzisere Identifikation sprachlicher Unterschiede ermöglichen. Die durchgeführten Forschungsziele umfassten den Vergleich der Methoden zur Anwendung von KI in der kontrastiven Linguistik sowie die Bewertung ihrer Effektivität. Die Analyse ergab, dass KI strukturelle Muster in Sprachen effektiv erkennen kann, jedoch weiterhin von der Qualität der bereitgestellten Daten abhängig bleibt. Basierend auf den Forschungsergebnissen werden Schlussfolgerungen zur Rolle der KI in der sprachwissenschaftlichen Forschung präsentiert. Eine potenzielle praktische Anwendung der Ergebnisse liegt in der Beschleunigung von Forschungsprozessen und der Erhöhung der Präzision bei der Analyse von Sprachunterschieden und -ähnlichkeiten.

Schlüsselwörter: kontrastive Linguistik, künstliche Intelligenz, natürliche Sprachverarbeitung, KI-Modelle

Contrastive Linguistic Analysis Using Artificial Intelligence Models: A Comparison of German and Other Languages

With the rapid development of artificial intelligence (AI) and natural language processing (NLP), new opportunities are emerging in linguistic analysis, particularly in contrastive and comparative studies. Traditional methods of linguistic analysis are time-consuming and resource-intensive, making AI a valuable tool for such research. This paper aims to present an innovative approach to contrastive analysis of the German language in comparison with other languages, utilizing the latest AI models, such as transformer-based models (e.g., BERT, GPT). The analysis includes a comparison of grammatical and semantic structures, as well as an attempt at large-scale automated identification of differences and similarities. Compared to previous research, the discussed approach stands out due to the use of advanced AI tools that enable faster and more precise identification of linguistic differences. The research objectives included comparing

AI-based methods in contrastive linguistics and assessing their effectiveness. The analysis revealed that AI can successfully identify structural patterns in languages; however, it remains dependent on the quality of the provided data. Based on the study's results, conclusions regarding the role of AI in linguistic research will be presented. A potential practical application of these findings is the acceleration of research processes and the enhancement of precision in analyzing linguistic differences and similarities.

Keywords: contrastive linguistics, artificial intelligence, natural language processing, AI models

Author: Joanna Grzybowska, University of Wrocław, Pl. Nankiera 15b, 50-140 Wrocław, Poland, e-mail: 298975@uwr.edu.pl

Received: 15.2.2025 **Accepted**: 26.5.2025

W ostatnich latach rozwój technologii, zwłaszcza w dziedzinie sztucznej inteligencji, otworzył przed lingwistyką nowe możliwości analizy języków i ich porównywania. Dzięki nowoczesnym narzędziom przetwarzania języka naturalnego, jesteśmy w stanie badać złożone struktury gramatyczne oraz semantyczne różnice między językami w sposób bardziej precyzyjny i dogłębny niż kiedykolwiek wcześniej. Wykorzystanie zaawansowanych modeli językowych pozwala dostrzec subtelne różnice i cechy wspólne języków, co w tradycyjnych analizach lingwistycznych może być trudne do wychwycenia.

Niniejszy tekst ma na celu przedstawienie nowatorskiego podejścia do kontrastywnej analizy języka niemieckiego z innymi językami, z wykorzystaniem modeli sztucznej inteligencji, takich jak modele transformerowe (np. BERT, GPT). Przedmiotem analizy jest zestawienie języka niemieckiego z językiem polskim i japońskim.

Do analizy trzech tekstów wykorzystano model sztucznej inteligencji Generative Pre-Trained Transformer (GPT-4o), który dokonał opisowej charakterystyki podanych próbek tekstowych. Skorzystano również z programu opartego na architekturze Bidirectional Encoder Representations from Transformers (BERT) i pomocniczej sieci neuronowej biblioteki SpaCy. W dalszej części artykułu program ten nazywany będzie BERT-SpaCy. Do analizy semantycznej użyto modelu bert-base-multilingual-cased. Do oceny sentymentu tekstu w języku polskim wykorzystano model bert-base-polish-cased-v1, w języku niemieckim model gottbert-base, a w języku japońskim bert--base-japanese. Z kolei zadaniem biblioteki SpaCy była analiza struktury składniowej i podanie liczbowych danych na temat częstotliwości występowania poszczególnych części mowy, do czego użyto modeli de_core_news_lg dla języka niemieckiego, pl_core_news_lg dla polskiego i jp_core_news_lg dla japońskiego.

Języki można klasyfikować według różnych kryteriów, a jednym z nich jest typ morfologiczny, który określa sposób tworzenia i łączenia wyrazów. Język polski należy do języków fleksyjnych, co oznacza, że wykorzystuje bogaty system odmiany przez przypadki, liczby i rodzaje, a znaczenie wyrazów często zależy od końcówek fleksyjnych. Ze względu na specyfikę języka polskiego jako języka fleksyjnego, efektywne trenowanie modeli NLP wymaga dostosowania metod pretrenowania do bogatej

morfologii i licznych form odmiany. Jak zauważają Mroczkowski/Rybak/Wróblewska/Gawlik (2021: 1–2), model HerBERT został pretrenowany na polskojęzycznych korpusach, aby lepiej oddawać charakterystykę języka, co wskazuje na konieczność stosowania modeli dostosowanych do poszczególnych języków, a nie tylko korzystania z modeli wielojęzycznych.

Język niemiecki również jest językiem fleksyjnym, ale wykazuje cechy syntetyczno-analityczne, co oznacza, że obok odmiany fleksyjnej w dużym stopniu polega na stałym szyku wyrazów i użyciu przyimków do określania relacji gramatycznych. Język japoński, w przeciwieństwie do polskiego i niemieckiego, jest językiem aglutynacyjnym, co oznacza, że tworzy formy gramatyczne poprzez dodawanie sufiksów i partykuł, które mają jednoznaczne znaczenie i funkcję, bez większych zmian w rdzeniu wyrazu.

Analiza języków o różnej typologii oraz z różnych rodzin językowych może przyczynić się do lepszego zrozumienia tych aspektów języka, które są uniwersalne, jak i uwydatnienia takich struktur, które są dla danego typu lub grupy językowej unikalne. W tym kontekście szczególnie przydatne są nowoczesne narzędzia oparte na sztucznej inteligencji.

Rozwój technologii NLP, czyli Natural Language Processing (nie należy mylić z Neuro-Linguistic Programming, również związanym z językami, lecz bardziej w kontekście psychologii i komunikacji międzyludzkiej), oferuje nowe możliwości w dziedzinie analizy języków oraz badań kontrastywnych, pozwalając na szybsze i dokładniejsze eksplorowanie zawiłych struktur językowych oraz na porównywanie języków na większą skalę niż kiedykolwiek wcześniej. Dzięki zaawansowanym algorytmom NLP, badacze mogą analizować ogromne zasoby danych językowych, co sprzyja odkrywaniu ukrytych wzorców oraz zależności, które wcześniej były trudne do uchwycenia w tradycyjnych badaniach lingwistycznych. Rozwój NLP jest ściśle związany z ewolucją metod tłumaczenia maszynowego, począwszy od systemów opartych na regułach (RBMT), przez modele statystyczne (SMT), aż po nowoczesne sieci neuronowe (NMT). Obecnie modele transformerowe stanowią fundament tych technologii, oferując znacznie lepsze wyniki w przetwarzaniu tekstu dzięki ich zdolności do modelowania kontekstu na szeroką skalę (Bieda 2022: 3).

Transformerowe modele językowe (TLM) zrewolucjonizowały przetwarzanie języka naturalnego, a większość najnowocześniejszych modeli NLP opiera się na nich, wykorzystując kontekstualizację i indukcję wiedzy. Jak wskazują dotychczasowe badania, analiza tych modeli koncentrowała się głównie na języku angielskim, podczas gdy modele dla innych języków, takich jak niemiecki czy polski, pozostają mniej zbadane (Zaczyńska/Feldhus/Schwarzenberg/Gabryszak/Möller 2020: 1–2).

Modele transformerowe odgrywają szczególną rolę w przetwarzaniu języka naturalnego, a co za tym idzie, w zaawansowanych badaniach lingwistycznych, umożliwiając dokładniejsze eksplorowanie relacji między słowami i kontekstem w różnych językach. Jest to rodzaj architektury sieci neuronowej będący fundamentem wielu

zaawansowanych modeli sztucznej inteligencji, takich jak GPT czy BERT, działających na mechanizmie tzw. multi-head attention (‚uwagi wielogłowej'), która pozwala przetwarzać słowa i elementy tekstu jednocześnie, a nie, tak jak w tradycyjnych sieciach neuronowych, sekwencyjnie. Jego skuteczność ocenia się za pomocą różnych metryk, z których jedną z kluczowych jest perplexity (PPL) – miara określająca stopień niepewności modelu co do następnego słowa w sekwencji. Niższa wartość perplexity oznacza lepszą przewidywalność, co przekłada się na wyższą spójność generowanego tekstu oraz lepszą zdolność modelu do analizy struktur językowych i identyfikacji wzorców w tekście (Jagdishbhai/Thakkar 2023: 18–19).

Modele transformerowe mają szerokie zastosowanie w językoznawstwie, zwłaszcza językoznawstwie komputerowym, gdzie wykorzystuje się je w tłumaczeniu maszynowym, generowaniu tekstu, analizie sentymentu oraz klasyfikacji i wyszukiwaniu informacji w dużych zbiorach danych tekstowych. Wykorzystanie tych zaawansowanych technologii otwiera nowe perspektywy w badaniach nad złożonością i różnorodnością języków, co jest szczególnie przydatne w analizie kontrastywnej języków z tej samej grupy językowej, takich jak niemiecki i polski.

Jak zauważyli Dadas/Perełkiewicz/Poświata (2020), w przeciwieństwie do języka angielskiego, dla którego opracowano liczne modele oparte na architekturze transformerów, język polski pozostaje niedostatecznie reprezentowany w tym obszarze. Dostępne modele nie osiągają skali porównywalnej z największymi modelami anglojęzycznymi, zarówno pod względem rozmiaru korpusu, jak i liczby parametrów, co ogranicza ich skuteczność w różnorodnych zadaniach NLP.

Podobne wyzwania napotykano w przypadku języka niemieckiego, gdzie modele wielojęzyczne, takie jak mBERT, nie zawsze zapewniały optymalne rezultaty. W odpowiedzi na ten problem opracowano GottBERT – monolingwalny model transformerowy, który w testach przewyższył modele wielojęzyczne w zadaniach klasyfikacji tekstów i rozpoznawania nazw własnych (Scheible/Thomczyk/Tippmann/Jaravine/Boeker 2020: 2–3). To sugeruje, że dalszy rozwój modeli dostosowanych do konkretnych języków może być bardziej efektywny niż stosowanie uniwersalnych podejść.

Takie ograniczenie stanowi wyzwanie dla badań nad językiem polskim, zwłaszcza w kontekście porównań z lepiej opracowanymi modelami dla innych języków. W celu zilustrowania tych różnic oraz oceny skuteczności dostępnych narzędzi, przeprowadzono analizę fragmentów tekstów opisujących zamek Chambord we Francji jako atrakcyjną destynację turystyczną. Teksty te posłużyły do porównania języka niemieckiego[1] i polskiego[2], umożliwiając ocenę skuteczności modeli AI w analizie kontrastywnej.

Model GPT-4o uporządkował opis językowy, wyszczególniając kilka kategorii. Zaczął od stylu i tonu podanych tekstów. Określił niemiecki tekst jako precyzyjny

[1] Vgl. https://www.bloischambord.de/untersuchen/die-schloesser/schloss-chambord, dostęp: 4.11.2024.

[2] Cartwright, Mark. „Chateau de Chambord", 2020. https://www.worldhistory.org/trans/pl/1-19055/chateau-de-chambord/, dostęp: 3.11.2024.

i bezpośredni: Niemiecki opis zamku cechuje się formalnością i dokładnością, często używa terminów architektonicznych, np. *Schlossanlage* (‚kompleks zamkowy'), natomiast tekst polski ma bardziej narracyjny, opisowy charakter, z naciskiem na emocjonalne epitetowanie, np. *wspaniały zamek*. W strukturze gramatycznej model GPT-4o zwrócił uwagę na różnice w deklinacji (cztery przypadki w niemieckim, siedem w polskim), podkreślając większą elastyczność polskiego systemu przypadków. Niemiecki szyk zdania opiera się na sztywnych zasadach składniowych (np. szyk V2, czasownik na końcu w zdaniach podrzędnych), podczas gdy polski, dzięki fleksji, pozwala na większą swobodę w budowie zdań. W analizie leksykalnej model zaznaczył, że niemiecki często stosuje złożone rzeczowniki: „Długie słowa, takie jak Schlossparkverwaltung (zarząd parku zamkowego), są charakterystyczne dla niemieckiego i służą dokładności". W polskim dominują prostsze, jednoelementowe konstrukcje, a teksty są bardziej opisowe i nacechowane emocjonalnie. W budowie zdań niemiecki preferuje długie zdania podrzędnie złożone, co nadaje formalny i uporządkowany charakter, podczas gdy polski tekst wykazuje większą różnorodność długości i struktur zdań, co ułatwia przyswajanie treści.

Program BERT-SpaCy w swoim raporcie wskazał, że średnia długość zdań w tekście niemieckim to 19.7 słów, a w tekście polskim 19.3. Różnorodność leksykalną próbki w języku niemieckim określił na 0.6129, a w polskim na 0.6206 w skali 0-1, co wskazuje, że oba teksty cechują się porównywalnym stopniem zróżnicowania słownictwa. Nieznacznie wyższa różnorodność leksykalna w tekście polskim może wynikać z większej liczby synonimów i bardziej elastycznej składni, podczas gdy niemiecki często korzysta z rzeczowników złożonych, co ogranicza liczbę unikalnych leksemów. BERT-SpaCy ocenił również emocjonalność (sentyment) tekstu, stosując skalę od 1 do 5, niemiecki tekst plasując na 2.8107. Określił, że ton tekstu jest formalny, co jest sugerowane długimi i skomplikowanymi zdaniami. W przypadku polskiego wartość ta wyniosła 2.9582, co świadczy o nieco większej emocjonalności niż w niemieckim, wynikającej z większej różnorodności słownictwa. BERT-SpaCy ocenił podobieństwo semantyczne obu tekstów na 82%, co wskazuje, że prawidłowo określił, iż obydwie próbki traktują o tym samym.

Program korzysta z wielu liczbowych reprezentacji (embeddingów), aby uchwycić zależności semantyczne i syntaktyczne pomiędzy słowami, co umożliwia mu rozumienie kontekstu i subtelności językowych w przetwarzanych tekstach.

Podczas gdy analiza niemieckiego i polskiego pozwoliła na szczegółowe porównanie europejskich struktur językowych, zestawienie języka niemieckiego z japońskim otwiera nowe perspektywy na kontrast między językami z dwóch odmiennych kręgów kulturowych i językowych. W analizie również wykorzystano opisy zamku Chambord we Francji[3].

[3] Firiatoru, Pasukaru, „シャンボール城: 知っておくべき9つのポイント", 2018. https://www.france.fr/ja/article/9-choses-a-savoir-sur-le-chateau-de-chambord/, dostęp: 4.11.2024.

Model GPT-4o wykazał, że niemiecki opis ma charakter formalny i techniczny, podczas gdy japoński jest bardziej poetycki, oparty na ogólnych odczuciach i wrażeniach estetycznych. „Ton japońskiego opisu jest mniej bezpośredni i bardziej subtelny, z naciskiem na harmonijne wkomponowanie zamku w otaczający krajobraz. To podejście bardziej podkreśla atmosferę miejsca niż jego techniczne aspekty". W analizie gramatycznej sztuczna inteligencja zwróciła uwagę, że niemiecki posiada ścisłe zasady składniowe, takie jak szyk V2 i obowiązkowe rodzajniki, natomiast japoński jest bardziej elastyczny dzięki cząstkom gramatycznym (np. は, が, を), które określają relacje między elementami zdania. Jako język aglutynacyjny, japoński tworzy struktury gramatyczne poprzez dodawanie przyrostków zamiast odmiany przez przypadki. Badania Kuribayashiego, Ito, Suzukiego i Inuiego (2020) wykazały, że modele NLP dobrze odwzorowują preferowane kolejności składniowe w japońskim, ale napotykają trudności w analizie bardziej złożonych struktur. W warstwie leksykalnej model wskazał, że niemiecki tekst obfituje w techniczne terminy (*Schlossanlage*, *Fassaden*, *Architektur*), co podkreśla jego precyzyjny charakter, natomiast japoński operuje metaforami i ekspresyjnymi określeniami, jak 美しい (‚piękny'), odwołującymi się do estetyki wabi-sabi (uznanie dla niedoskonałości i piękna przemijania). Struktura zdań w języku niemieckim jest rozbudowana i często wielokrotnie złożona, natomiast w japońskim dominuje zwięzłość i rytmika, które sprawiają, że tekst staje się niemal poetycki. AI zauważyła, że aglutynacyjna natura japońskiego pozwala na elastyczne układanie zdań, podkreślając subiektywne wrażenia i kontekst kulturowy.

Ważnym aspektem kontrastowej analizy tych dwóch języków jest fakt, że język niemiecki stosuje alfabet łaciński, podczas gdy japoński zapisuje się głównie przy użyciu kanji (znaków logograficznych zapożyczonych z chińskiego) oraz dwóch sylabariuszy – hiragany i katakany. Każdy znak kanji reprezentuje nie tylko dźwięk, ale także pewne znaczenie. Dlatego jedno słowo może wizualnie przekazywać złożone idee lub odczucia, co odróżnia japoński od języków opartych na alfabecie. To zestawienie podkreśla wagę zrozumienia kontekstu kulturowego i strukturalnego, co pozwala na pełniejsze zrozumienie i interpretację tekstów. Wpływa także na skuteczność wykorzystywania modeli sztucznej inteligencji do analizy tekstów w językach, które korzystają z innych pism niż alfabet łaciński.

Program BERT-SpaCy przeprowadził analizę porównawczą tekstów w języku niemieckim i japońskim, oceniając m.in. ich strukturę, długość zdań oraz różnorodność leksykalną. Średnia długość zdań w niemieckim wyniosła 19.7 słów, natomiast w japońskim 36.9, co sugeruje, że tekst niemiecki charakteryzuje się większą zwięzłością. Różnorodność leksykalna niemieckiej próbki została określona na 0.6129, natomiast w japońskim osiągnęła wartość 0.4161. Niższy wynik w przypadku języka japońskiego może wynikać z jego specyfiki, obejmującej częste użycie partykuł, powtórzeń oraz stosunkowo mniejszą liczbę unikalnych leksemów w analizowanej próbce. Program dokonał również oceny emocjonalnego nacechowania tekstu, przyznając niemieckiemu opisowi sentyment na poziomie 2.8107 w skali od 1 do 5, a japońskiemu 3.0098,

co wskazywałoby, że tekst niemiecki jest bardziej rzeczowy i formalny, podczas gdy tekst japoński wykazuje tendencję do subtelniejszego wyrażania treści oraz większego nacisku na nastrój i atmosferę.

Podobnie jak przy porównaniu tekstów w języku niemieckim i polskim, także i w tym przypadku podobieństwo semantyczne dla próbki niemieckiej i japońskiej wyniosło 82%, co wskazuje, że program trafnie rozpoznał treść tekstów.

Podsumowując, analizy wykonane przez model GPT-4o oraz program BERT-SpaCy kierują do następujących wniosków:

Gramatyka: Polski i niemiecki to języki fleksyjne, ale niemiecki ma bardziej rozbudowaną deklinację i ustalony szyk zdania. Japoński, będący językiem aglutynacyjnym, znacząco różni się strukturalnie, co czyni go bardziej elastycznym w dodawaniu partykuł i określaniu znaczeń; Styl i ton: Niemiecki tekst jest bardziej formalny i precyzyjny, polski – neutralny i opisowy, natomiast japoński poetycki i nacechowany emocjonalnie, z elementami nawiązującymi do harmonii i estetyki. Słownictwo: Niemiecki używa długich słów złożonych, polski stosuje proste, opisowe słownictwo, a japoński preferuje wyrażenia metaforyczne. Różnica ta wskazuje, jak różne kultury traktują opisy miejsc – w Polsce i Niemczech bardziej dosłownie, a w Japonii – z poetyckim zabarwieniem. Japońska tendencja do metaforycznych opisów wynika z głęboko zakorzenionego naturalizmu w kulturze tego kraju. Język japoński często odwołuje się do przyrody jako kluczowego elementu estetyki, podkreślając piękno prostoty, niedoskonałości i przemijania.

Aby porównać wyniki liczbowe BERT-SpaCy z komentarzem modelu GPT-4o, poproszono go o interpretację liczbową (embeddingi) podanych tekstów. Uwagę zwraca fakt, że program i model przedstawiły podobne, ale nie tożsame wyniki. Różnice pomiędzy wynikami analizy języków polskiego, niemieckiego i japońskiego za pomocą modelu GPT-4o oraz programu BERT-SpaCy mogą wynikać z kilku czynników, w tym tokenizacji (czyli sposobu podziału tekstu na jednostki leksykalne – tokeny), klasyfikacji części mowy, metodologii liczenia zdań oraz traktowania interpunkcji. Jeśli chodzi o język polski, przykładowo, w przypadku słów z przyimkami, jak *na przykład* czy *z powodu*, program może traktować te wyrażenia jako oddzielne tokeny, podczas gdy GPT-4o może traktować je jako jedną jednostkę. Z kolei w przypadku złożonego słowa *Herrenhaus* w języku niemieckim, program podzieliłby je na dwie jednostki (np. *Herren* i *haus*), natomiast GPT-4o traktowałby to jako jedno słowo. W języku japońskim z kolei, struktura słów może być rozbita na jednostki leksykalne w sposób, który nie występuje w polskim czy niemieckim, co skutkuje innymi wynikami tokenizacji.

Różnice mogą również wynikać z klasyfikacji części mowy. Program BERT-SpaCy może uznawać imiesłowy, takie jak *będący* w polskim, za przymiotniki (ADJ), podczas gdy GPT-4o zaklasyfikuje je jako czasowniki (VERB). W niemieckim program może błędnie sklasyfikować formy czasownikowe w różnych przypadkach gramatycznych, jak na przykład w słowie *gesehen* (zamiast uznać go za czasownik w formie przeszłej, może traktować go jako przymiotnik). W japońskim, który jest bardziej złożony w kontekście

gramatyki i składni, program i model GPT-4o mogą inaczej klasyfikować partykuły (np. ∅ jako część mowy zamiast klasyfikować ją jako zaimek) w zależności od kontekstu.

Kolejnym czynnikiem są różnice w metodzie liczenia zdań. Na przykład w analizie języka polskiego program BERT-SpaCy naliczył 81 zdań, podczas gdy GPT-4o zaledwie 62, co może wynikać z różnic w traktowaniu zdań podrzędnych i skomplikowanych struktur gramatycznych. W niemieckim, GPT-4o może traktować niektóre zdania złożone jako jedno długie zdanie, podczas gdy program może liczyć każde zdanie osobno.

Również tokenizacja interpunkcji i spacji może przyczynić się do różnic. Program BERT-SpaCy, traktując spacje jako osobne tokeny, może mieć ich znaczną liczbę, co może wpływać na liczbę tokenów, podczas gdy GPT-4o może te tokeny traktować inaczej. W analizie polskiej program podał liczbę 279 znaków interpunkcyjnych, podczas gdy GPT-4o – 137. W przypadku języka japońskiego, interpunkcja jest bardziej uproszczona, a wyniki mogą różnić się w zależności od sposobu przetwarzania znaków i ich roli w strukturze gramatycznej.

Podsumowując, rozbieżności w wynikach analizy językowej można przypisać różnicom w sposobie segmentacji tekstu, przypisywaniu kategorii gramatycznych, liczenia zdań oraz podejściu do interpunkcji. Czynniki te sprawiają, że modele BERT-SpaCy i GPT-4o mogą generować odmienne wyniki, szczególnie w przypadku języków o odmiennych strukturach składniowych i morfologicznych.

Wyniki sugerują, że użycie modeli AI w analizach kontrastywnych może znacznie przyspieszyć identyfikację kluczowych różnic między językami oraz wspomóc określanie cech nawet bardzo długich tekstów. Dzięki wszechstronnym możliwościom AI, analiza kontrastywna nie tylko przyspiesza rozpoznawanie głównych różnic językowych, ale również uwidacznia wzorce stylistyczne i strukturalne w tekstach o znacznej długości. Pojawiają się przez to jednak kwestie etyki w wykorzystywaniu takich narzędzi w badaniach. Automatyzacja analizy niesie ryzyko marginalizacji wkładu badacza w proces naukowy. Jednakże, narzędzia sztucznej inteligencji, zwłaszcza na współczesnym etapie rozwoju, są dalekie od doskonałości i zawsze wymagane jest monitorowanie postępów przez osoby zaangażowane w badanie. AI pozostawiona samej sobie popełnia wiele błędów – m.in. literówki (np. *leksylokalne* zamiast *leksykalne*), zdarza się także, że nie rozumie polecenia. Przykładowo, model GPT miał za zadanie przeanalizować po kolei wysyłane pliki z tekstami zanim przejdzie do analizy kontrastywnej, jednakże, po dostarczeniu mu trzeciego pliku, stworzył jego streszczenie.

Warto więc zaznaczyć, że stosowanie AI w badaniach przyspiesza proces analizy i pozwala na większą dokładność w pomiarach, ale to ostatecznie badacze są odpowiedzialni za wybór tekstów, modeli, interpretację wyników i przedstawienie ich w kontekście badań kontrastywnych.

Precyzja analizy programu BERT-SpaCy zależała od załadowanych modeli językowych, takich jak de_core_news_lg czy gottbert-base dla języka niemieckiego.

Kluczowym elementem tego procesu jest kalibrowanie modeli w zakresie tokenizacji. Rolą badacza przy posługiwaniu się narzędziami sztucznej inteligencji nie jest jedynie pasywne korzystanie z gotowych algorytmów, ale także dostrajanie parametrów i dobór odpowiednich modeli do zadanego problemu badawczego lub tworzenie i trenowanie takiego modelu.

Wykorzystanie AI w wielu sferach nauki staje się normą, bo umożliwia badaczom skupienie się na interpretacjach, poszukiwaniach nowych pytań badawczych i tworzeniu nowych teorii, a nie tylko na ręcznym przetwarzaniu danych. W tym kontekście sztuczna inteligencja to przydatne, coraz powszechniej wykorzystywane narzędzie, lecz aktywność intelektualna ludzi w analizie tekstów, interpretacji wyników oraz wyciąganiu wniosków jest absolutnie niezbędna.

Choć narzędzia sztucznej inteligencji otwierają przed nauką nowe możliwości, to ich pełen potencjał ujawnia się dopiero wtedy, gdy są używane w sposób świadomy i odpowiedzialny, przyczyniając się do pogłębiania wiedzy o otaczającym świecie, do rozwiązywania interdyscyplinarnych problemów badawczych oraz inspirowania nowatorskich kierunków rozwoju w nauce i technologii.

Wykaz literatury

Bieda, Matthew. *Refining the state-of-the-art in Machine Translation, optimizing NMT for the JA <-> EN language pair by leveraging personal domain expertise*, 2022. https://arxiv.org/pdf/2202.11669. 10.1.2025.

Cartwright, Mark. *Chateau de Chambord*, 2020. https://www.worldhistory.org/trans/pl/1-19055/chateau-de-chambord/. 3.11.2024.

Dadas, Sławomir, Michał Perełkiewicz i Rafał Poświata. *Pre-Training Polish Transformer-Based Language Models at Scale*, 2020. https://arxiv.org/pdf/2006.04229. 10.1.2025.

Firiatoru, Pasukaru. シャンボール城：知っておくべき9つのポイント, 2018. https://www.france.fr/ja/article/9-choses-a-savoir-sur-le-chateau-de-chambord/. 4.11.2024.

Jagdishbhai, Nimit i Krishna Yatin Thakkar. *Exploring the Capabilities and Limitations of GPT and ChatGPT in Natural Language Processing*, 2023. https://pdf.ipinnovative.com/pdf/18673. 11.1.2025.

Kuribayashi, Tatsuki, Takumi Ito, Jun Suzuki i Kentaro Inui. *Language Models as an Alternative Evaluator of Word Order Hypotheses: A Case Study in Japanese*, 2020. https://arxiv.org/pdf/2005.00842. 16.1.2025.

Mroczkowski, Robert, Piotr Rybak, Alina Wróblewska i Ireneusz Gawlik. *HerBERT: Efficiently Pretrained Transformer-based Language Model for Polish*, 2021. https://arxiv.org/pdf/2105.01735. 10.1.2025.

Scheible, Raphael, Fabian Thomczyk, Patric Tippmann, Victor Jaravine i Martin Boeker. *GottBERT: a pure German Language Model*, 2020. https://arxiv.org/pdf/2012.02110. 11.1.2025.

Zaczyńska, Karolina, Nils Feldhus, Robert Schwarzenberg, Aleksandra Gabryszak i Sebastian Möller. *Evaluating German Transformer Language Models with Syntactic Agreement Tests*, 2020. https://arxiv.org/pdf/2007.03765. 10.1.2025.

Internetquellen

https://www.bloischambord.de/untersuchen/die-schloesser/schloss-chambord. 4.11.2024.

ZITIERNACHWEIS:

Grzybowska, Joanna. „Kontrastowa analiza językowa z wykorzystaniem modeli sztucznej inteligencji: porównanie języka niemieckiego i innych języków“, *Linguistische Treffen in Wrocław* 28, 2025 (II): 165–174. DOI: 10.23817/lingtreff.28-9.