ALINA WRÓBLEWSKA*, MARTYNA LEWANDOWSKA**, ALEKSANDRA TOMASZEWSKA***
KAROL SAPUTA****, MACIEJ OGRODNICZUK*****

INSTYTUT PODSTAW INFORMATYKI POLSKIEJ AKADEMII NAUK, WARSZAWA

Koncepcja form równościowych z asteryskiem inkluzywnym[1]

Słowa kluczowe: język inkluzywny, feminizacja, asterysk inkluzywny, duże modele językowe.

https://doi.org/10.31286/JP.001040

Wprowadzenie

Współcześnie używany język polski nie jest wrażliwy na płeć. Zgodnie ze stanowiskiem Rady Europy wpisuje się nawet w paradygmat języków seksistowskich. Seksizm lingwistyczny to dyskryminacja osób jednej płci za pomocą środków języka, a jednym z jego przejawów jest dominacja wyrażeń męskich nad żeńskimi (GEG 2016: 15; GNL-EN 2018: 7). Zauważając negatywne skutki stosowania języka seksistowskiego, Rada Europy zachęca państwa członkowskie do eliminacji seksizmu z języków oraz do promowania języka odzwierciedlającego zasadę równości płci w dokumentach urzędowych, w tym w tekstach prawnych, a także w dokumentach dotyczących polityki i programów publicznych, w komunikacji usług publicznych z obywatelami i obywatelkami, w edukacji i mediach (CDEG 2007, § 66–71).

Z perspektywy systemu języka polskiego nie istnieją przeszkody lingwistyczne uniemożliwiające pisanie i mówienie o osobach zgodnie z ich rodzajem naturalnym (tj. płcią żeńską lub męską). Jedną z kategorii gramatycznych w polszczyźnie jest rodzaj gramatyczny, który determinuje podział rzeczowników na grupy deklinacyjne: nijaki, żeński, męski osobowy (m1), męski zwierzęcy (m2) i męski rzeczowy (m3). Rzeczowniki osobowe w rodzaju żeńskim zwykle oznaczają osobę lub grupę osób płci żeńskiej (np.: *kobieta*, *sprzedawczyni*, *siatkarki*), natomiast rzeczowniki osobowe w rodzaju m1 odnoszą się w pierwszej kolejności do osoby lub grupy osób płci męskiej (np.: *mężczyzna*, *sprzedawca*, *siatkarze*).

Poza użyciem zgodnym z płcią rzeczowniki nazywające mężczyzn są powszechnie stosowane jako uniwersalne określenia odnoszące się do grup różnopłciowych (np.: *posłowie*, *nauczyciele*), a nawet osób płci żeńskiej (np.: *prezydent Zdanowska*, *kopciuszek*, *babsztyl*). O ile

* alina@ipipan.waw.pl; ORCID: 0000-0002-2589-9900
** martyna.lewandowska@ipipan.waw.pl; ORCID: 0009-0005-7918-3026
*** aleksandra.tomaszewska@ipipan.waw.pl; ORCID: 0000-0001-6379-3034
**** ORCID: 0000-0002-8809-5248
***** maciej.ogrodniczuk@ipipan.waw.pl; ORCID: 0000-0002-3467-9424

1 Pragniemy podziękować prof. Agnieszce Pantuchowicz oraz dr. Oskarowi Kędzierskiemu za inspirację do podjęcia badań nad zagadnieniem równości płci w języku.

prezydent i *kopciuszek* uzgadniające się z potencjalnymi orzeczeniami w rodzaju żeńskim (np. *przyszła*) mogą być interpretowane jako mające alternatywną charakterystykę rodzajową (Saloni i in. 2012: 39), o tyle rzeczownik *babsztyl* uzgadnia się wyłącznie z orzeczeniem w rodzaju m1, chociaż jest pejoratywnym określeniem kobiety. W literaturze językoznawczej funkcjonuje pojęcie *generyczności rzeczowników męskoosobowych*, określane czasem również jako *generyczny rodzaj męski*. Generyczny rodzaj męski sankcjonuje użycie męskich form w odniesieniu do kobiet i grup mieszanych (np.: Łaziński 2006, 2023; Małocha-Krupa 2018; Szpyra-Kozłowska 2024). Innymi słowy, to rodzaj niespecyficzny, uważany za uniwersalny, wspólny dla wszystkich płci. Ponieważ wyrażenia w rodzaju męskim i w generycznym rodzaju męskim są identyczne, rozróżnienie ich użyć jest wyłącznie kontekstowe, a tym samym może prowadzić do nieporozumień komunikacyjnych, np.:

(1) Dziś odbędzie się spotkanie z naszym firmowym prawnikiem.

(2) Poszukujemy do zespołu doświadczonego inżyniera.

Zdanie (1) może sugerować, że spotkanie odbędzie się z mężczyzną, nawet jeśli tę funkcję sprawuje kobieta; podobnie po przeczytaniu w ogłoszeniu zdania (2) kandydatki mogą czuć się wykluczone, interpretując ofertę jako skierowaną tylko do mężczyzn.

Jak pisze Jolanta Szpyra-Kozłowska (2024), w literaturze wskazuje się na to, że użycie generycznego maskulinum wywołuje głównie skojarzenia z mężczyznami, przyczyniając się do niewidzialności kobiet w języku. To, jak również i wniosek, że użycie generycznego rodzaju męskiego powoduje wykluczenie kobiet, wykazano na gruncie wielu języków, np. w języku niemieckim (m.in.: Stahlberg, Sczesny 2001; Müller-Spitzer 2022) i w polskim (m.in.: Bojarska 2011; Szpyra-Kozłowska 2024).

Standardowa polszczyzna została ukształtowana przez polityki językowe stosowane w różnych okresach historycznych. Istotny wpływ wywarły na nią fala emancypacji kobiet w XIX wieku, która wymusiła tworzenie żeńskich nazw stanowisk tradycyjnie zajmowanych przez mężczyzn, a następnie propagowanie rzeczowników męskoosobowych afleksyjnych w odniesieniu do wyższych statusem zawodów, funkcji i tytułów kobiet w okresie komunistycznym (Osińska-Szymańska, Preuhs 2022). Pomimo formalnego równouprawnienia kobiet i mężczyzn w Polsce dyskryminacja językowa i marginalizacja form żeńskich są nadal obecne, co zupełnie nie odzwierciedla rzeczywistości społecznej. Biorąc pod uwagę, że w Polsce liczba kobiet jest wyższa niż liczba mężczyzn[2], powszechne stosowanie generycznych form męskich w odniesieniu do ogółu społeczeństwa jest nieuzasadnione. Narodowy spis powszechny z 2021 roku nie uwzględnił osób niebinarnych, podczas gdy aż 8 procent nastolatków wyraża wątpliwości dotyczące swojej płci (Czarnecka 2024). Wzrastająca świadomość dotycząca tożsamości płciowej oraz rosnąca społeczna akceptacja jej publicznej ekspresji wpływają na kształtowanie się języka, który adaptuje się do tej rzeczywistości. Prowadzone są również pierwsze badania niebinarnego języka polskiego, takie jak analiza niebinarnych form ze wskazaniem na ich społeczne zapotrzebowanie oraz rolę badań w opisywaniu zmiany

2 Zgodnie z wynikami Narodowego Spisu Powszechnego Ludności i Mieszkań 2021 w Polsce jest 18,39 miliona mężczyzn i 19,65 miliona kobiet (zob. Ludność 2021).

języka (Gębka-Wolak 2022) czy badanie leksykalnych i semantycznych zmian językowych w nomenklaturze medycznej[3] (Dyszak 2024).

Proces uwrażliwiania języka na płeć jest trudny i długotrwały z powodu głęboko zakorzenionych tradycyjnych form językowych, dezaprobaty społecznej, braku jednoznacznych reguł, a przede wszystkim naturalnego tempa ewolucji języka. Efektem procesu uwrażliwiania jest język inkluzywny, czyli medium komunikacji, w której unika się uprzedzeń, dyskryminacji i stereotypów związanych z płcią (Sczesny i in. 2016; Lindqvist i in. 2018). Język włączający umożliwia równość w wyrażaniu się o osobach niezależnie od ich tożsamości płciowej oraz unika wyrażeń dyskryminujących kobiety, osoby transpłciowe i niebinarne. W ramach dotychczasowych badań nad językiem opracowano strategie uwrażliwiania języka na płeć, takie jak feminizacja i neutralizacja (zob. rozdz. 1). Wdrażanie strategii języka wrażliwego na płeć pomaga osłabiać stereotypy związane z tradycyjnym przypisywaniem tak zwanych prestiżowych ról zawodowych mężczyznom, a jednocześnie przeciwdziała zjawisku „niewidoczności" kobiet w języku.

Obecnie brakuje jednak spójnej propozycji notowania form inkluzywnych, zwłaszcza ich skróconych wersji, dlatego proponujemy nową koncepcję zapisu. *Formy równościowe z asteryskiem inkluzywnym* (zob. rozdz. 2) zachowują równowagę między redundancją a ekonomicznością proponowanych form, są łatwo konwertowalne do innych wariantów notacji i są w pełni inkluzywne. Nasza koncepcja form równościowych z asteryskiem jest propozycją, którą chcemy przedstawić szerokiemu gronu czytelników i czytelniczek, bez zamiaru nakłaniania do jej stosowania. W ramach badań empirycznych przeanalizujemy różne gatunki tekstów pisanych i mówionych pod kątem możliwości wprowadzenia w nich form inkluzywnych, w szczególności form z asteryskiem, przez duże modele językowe (zob. rozdz. 3). Dodatkowo przeprowadzimy eksperymentalną weryfikację dużych modeli językowych, aby zbadać ich wrażliwość na płeć (zob. rozdz. 4). Sprawdzimy, czy modele potrafią adekwatnie przekształcać wybrane gatunki tekstów w standardowym języku polskim na ich wersje prorównościowe, kierując się wytycznymi starannie zdefiniowanymi na podstawie przeprowadzonych analiz.

Wobec braku oficjalnych krajowych rekomendacji dotyczących języka wrażliwego na płeć (Tomaszewska, Jamka 2024: 2) wyniki naszych działań mogą pełnić kilka dodatkowych funkcji, które traktujemy jako cele pośrednie. Po pierwsze, mogą stanowić próbę dostosowania komunikacji do zaleceń Unii Europejskiej w zakresie równości płci (GNL-PL 2018). Po drugie, mogą być krokiem do opracowania takich standardów na poziomie krajowym – w tej chwili dostępne źródła ograniczają się do poradnictwa językowego (np. Bańko i in. 2021). Po trzecie, mogą przyczynić się do uwrażliwienia na płeć komunikacji w Polsce – w szczególności zwiększenia widoczności kobiet w dyskursie społecznym.

1. Strategie uwrażliwiania języka na płeć

W języku polskim można wyróżnić trzy typy rzeczowników osobowych, które mogą bądź nie mogą podlegać modyfikacjom inkluzywnym. Pierwszą grupę stanowią rzeczowniki, które są neutralne płciowo niezależnie od ich rodzaju gramatycznego, np.: *osoba*, *człowiek*, *ludzie*, *jednostka*, *indywiduum*, *dziecko*, *niemowlę*, *noworodek*, *rodzic*. Do tej grupy zaliczamy również

3 W 2022 roku wprowadzono nową wersję Międzynarodowej Klasyfikacji Chorób ICD-11, która rezygnuje z określeń dotyczących tożsamości płciowej, które były krzywdzące, trudno więc mówić o ich dalszej słownikowej aktualności.

zaimki: *nikt*, *kto*, *ktoś* i *ktokolwiek*, które są neutralne płciowo pomimo gramatycznego rodzaju męskiego. Drugą grupę tworzą rzeczowniki specyficzne dla jednej płci, tzn. odnoszące się wyłącznie do mężczyzn lub kobiet. Do rzeczowników semantycznie męskich, które mają gramatyczny rodzaj męski, należą: *mężczyzna*, *chłopiec*, *brat*, *ojciec*, natomiast rzeczownikami semantycznie żeńskimi w rodzaju żeńskim są: *kobieta*, *dama*, *dziewczynka*, *siostra*, *matka*. W trzeciej grupie są rzeczowniki męskie (maskulatywy) oraz ich żeńskie derywaty (feminatywy), które tworzą pary rzeczowników równościowych, np.: *lekarz*/*lekarka*, *pielęgniarz*/*pielęgniarka*, *gość*/*gościni*.

Głównie rzeczowniki z trzeciej grupy podlegają zmianom równościowym. Istnieją dwie podstawowe strategie uwrażliwiania języka na płeć – feminizacja i neutralizacja (np. Krysiak, Małocha-Krupa 2020: 230).

1.1. Feminizacja języka

1.1.1. Koncepcja feminizacji

Feminizacja w języku odnosi się do procesu zwiększania widoczności kobiet przez użycie żeńskich odpowiedników męskich form rzeczownikowych. W praktyce polega na tworzeniu i stosowaniu feminatywów, czyli rzeczowników opisujących kobiety i mających morfologiczne znaczniki żeńskości (Krysiak, Małocha-Krupa 2020). Zgodnie z definicją Jolanty Szpyry-Kozłowskiej i Iwony Chmury-Rutkowskiej (2022) *feminatywy* to rzeczowniki będące żeńskimi nazwami funkcji, stanowisk i zawodów, jak również określeniami kobiet. W naszej interpretacji *feminatyw* jest derywatem morfologicznym (np. *posłanka*). Złożenia analityczne (np. *pani poseł*) czy rzeczowniki męskoosobowe mające alternatywną charakterystykę rodzajową w kontekście przymiotników i czasowników w rodzaju żeńskim (np. *ta poseł powiedziała*) nie są uznawane za feminatywy (Krysiak, Małocha-Krupa 2020; Tomaszewska, Jamka 2024).

Feminatywy są tworzone najczęściej przez dodanie sufiksów genderowych do form męskich (por. Szpyra-Kozłowska 2022): *-ka* (np. *socjolog* → *socjolożka*), *-ini*/*-yni* (np.: *gość* → *gościni*, *językoznawca* → językoznawczyni), *-ica*/*-yca* (np.: *karzeł* → *karlica*, *car* → *caryca*), *-ina*/*-yna* (np.: *hrabia* → *hrabina*, *blondyn* → *blondyna*), *-owa* (np. *szef* → *szefowa*). Drugim sposobem tworzenia feminatywów jest zmiana deklinacji męskiej na żeńską (np.: *radny* → *radna*, *minister* → *ministra*).

1.1.2. Realizacja

Feminatywy mogą być wprowadzane w formach współwystępujących (*symetryzacja*) oraz w bezpośrednich odniesieniach do kobiet (*adaptacja nazwy do płci*).

Symetryzacja genderowa (⚥[4]) polega na równoczesnym stosowaniu par maskulatyw–feminatyw w odniesieniu do osób o nieznanej płci (np.: *klient*/*klientka*, *użytkownik*/*użytkowniczka*), grup różnopłciowych (np.: *naukowcy*/*naukowczynie*, *uczniowie*/*uczennice*) i ogółu społeczeństwa (np. *Polki i Polacy*). Symetryzując standardowy język polski, najczęściej dodajemy feminatywy, dlatego symetryzacja jest często uznawana za wariant feminizacji (np. J. Szpyra-Kozłowska (2019: 23) definiuje *feminizację* jako tworzenie nazw żeńskich, które ma

4 Wszystkie symbole znajdują się w *Słowniku symboli* na końcu artykułu.

doprowadzić do symetrii między feminatywami i formami męskoosobowymi). Wprowadzając formy współwystępujące, unikamy domyślnej maskulinizacji języka, np. (3a) zamiast (3):

(3) Zapraszamy wszystkich nauczycieli.

(3a) Zapraszamy wszystkich nauczycieli i wszystkie nauczycielki.

Formy współwystępujące są szczególnie zalecane tam, gdzie ważne jest uwidocznienie obu płci, w szczególności w prezentacjach ustnych i w zwrotach adresatywnych. Formy współwystępujące uwzględniają obie płcie, odnosząc się bezpośrednio do kobiet i mężczyzn, ale jednocześnie często wykazują wysoki poziom redundancji i są często skracane w piśmie (zob. rozdz. 2).

Adaptacja nazwy do płci (♀) polega na zastąpieniu feminatywami męskoosobowych wyrażeń afleksyjnych odnoszących się do kobiet, np. (4a) zamiast (4):

(4) Katarzyna Kotula została nową Minister do spraw Równości.

(4a) Katarzyna Kotula została nową Ministrą do spraw Równości.

Adaptacja nazw funkcji i zawodów do płci, czyli stosowanie feminatywów lub ich unikanie, jest złożonym procesem, który powinien uwzględniać nie tylko normy językowe, ale też (przede wszystkim) indywidualne preferencje osób, do których te nazwy się odnoszą. Każdy człowiek ma prawo do samostanowienia w kwestii językowego wyrażania swojej tożsamości płciowej, co oznacza, że decyzja o stosowaniu lub niestosowaniu feminatywów w odniesieniu do niego powinna być respektowana niezależnie od preferencji językowych osoby nadającej komunikat.

1.2. Neutralizacja języka

1.2.1. Koncepcja neutralizacji

Neutralizując język, dobieramy słowa, które mogą nazywać wszystkie osoby niezależnie od ich płci, dzięki czemu unikamy przypisywania konkretnych ról społecznych, zawodów czy funkcji do jednej płci. Ta strategia jest szczególnie istotna w nazywaniu osób niebinarnych.

1.2.2. Realizacja

Podstawową strategią neutralizacji (⚲) jest stosowanie rzeczowników osobowych pierwszej grupy, czyli rzeczowników neutralnych płciowo. Ponadto można zastępować rzeczowniki nacechowane płciowo osobatywami albo uogólnieniami.

Asymilacja osobatywów to proces, w którym osobatywy zostają włączone do standardowego języka w zastępstwie rzeczowników osobowych nacechowanych płciowo. Osobatywy są wyrażeniami analitycznymi składającymi się z formy rzeczownika *osoba*, który jest neutralny płciowo (I grupa rzeczowników), oraz przydawki (atrybutu) wyrażonej przymiotnikiem (np.: *osoba niebinarna*, *osoba studencka*), imiesłowem (np. *osoba sprzątająca*) lub wyrażeniem przyimkowym (np. *osoba z niepełnosprawnością*).

Generalizacja polega na zastępowaniu wyrażeń nacechowanych płciowo rzeczownikami zbiorowymi, które są neutralne płciowo. Zamiast używać określeń takich jak *pracownicy* i *pracownice*, można stosować termin *personel*, który obejmuje wszystkie osoby zatrudnione

bez względu na płeć. Podobnie można użyć określenia *dyrekcja*, mówiąc o osobach zarządzających, albo *straż pożarna*, jeśli nie chcemy marginalizować *strażaczek*[5].

2. Nowa koncepcja skróconych form równościowych

2.1. Przegląd sposobów zapisu form równościowych używanych w polskich tekstach

W języku polskim stosowane są równolegle różne sposoby łączenia i skracania par wyrażeń żeńskich i męskich (tzw. *splitting*; zob. Szpyra-Kozłowska 2022). Podstawową formą równościową jest połączenie maskulatywu i feminatywu ukośnikiem, np.: *lekarz*/*lekarka*, *prawnicy*/*prawniczki*. J. Szpyra-Kozłowska (2022) dopuszcza również formy z nawiasami, które można uznać za formy skrócone, ponieważ powtarzające się znaki maskulatywu i feminatywu występują tylko raz. Jeśli wyrażenie żeńskie różni się od męskiego tylko tym, że zawiera ciąg znaków literowych wskazujących na płeć żeńską, to ten ciąg zostaje umieszczony w nawiasie bez łącznika, np.: *był*(*a*)*by*, *pan*(*i*), *student*(*ka*). Jeśli wyrażenia różnią się ciągami znaków, to różniący się żeński sufiks genderowy umieszcza się w nawiasie z łącznikiem po wyrażeniu męskim, np.: *urodzony*(*-a*), *partnera*(*-ki*). Autorka sugeruje jednak, żeby takiej formy nie stosować, jeśli w wyrazach zachodzą oboczności. Wtedy należy stosować podstawową formę z ukośnikiem.

Ta propozycja skróconej formy równościowej ma swoje warianty z ukośnikami lub podkreślnikami. W wypadku ukośników zastępują one nawiasy, służąc do wydzielania znaków charakterystycznych dla feminatywu (np.: *motocyklist*/*k*/*a*, *pan*/*i*) lub do łączenia maskulatywu z żeńskim sufiksem genderowym (np. *nauczyciele*/*-ki*). W wypadku form z podkreślnikami podkreślnik może zastępować ciągi znaków charakterystyczne dla wyrażeń męskich i żeńskich, np. *zauważył_ś* (tj.: *zauważył**e**ś* i *zauważył**a**ś*) albo może łączyć wyrażenie męskie z żeńskim sufiksem genderowym, np. *oceniali_ły* (tj.: *oceniali* i *oceniały*).

Podstawową wadą tego rozwiązania jest powszechne stosowanie znaczników inkluzywności niezgodnie z podstawową koncepcją (np. bez łącznika, z błędnym sufiksem genderowym zamiast wyodrębnienia znaków graficznych różniących formy). Przyczyną niezgodności może być nieintuicyjność koncepcji, przez co jest ona trudna do wdrożenia. Mnogość markerów inkluzywności (tj. ukośników, nawiasów, podkreślników) tylko pogłębia problem, bo dopuszcza alternatywne formy, czasem niespójnie stosowane w obrębie jednego tekstu.

Drugim istotnym problemem jest to, że takie formy równościowe nadal nie osiągają pełnej inkluzywności. Zawsze wyrażanie w jednym rodzaju jest uprzywilejowane poprzez pełne zapisanie, podczas gdy drugie sprowadza się jedynie do sufiksu genderowego, np.: *powinniśmy_yśmy*, *zauważyłaś_łeś*. Najczęściej to determinant wyrażenia żeńskiego zostaje wydzielony nawiasem, ukośnikiem lub podkreślnikiem, co uwypukla drugorzędne znaczenie feminatywów.

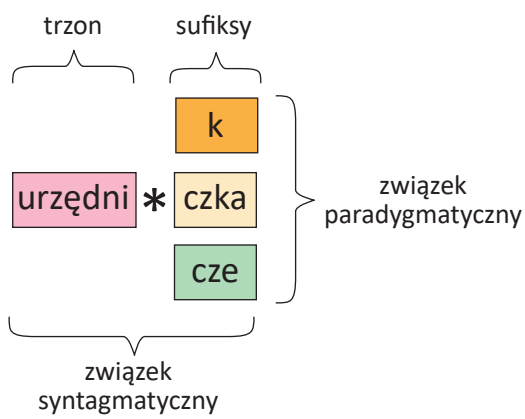
Ponadto miejsce przyłączenia dodatkowego sufiksu genderowego nie jest zaznaczone, przez co automatyczna konwersja na inny wariant form inkluzywnych lub odtwarzanie pojedynczych form jest skomplikowane. Kolejną wadą jest brak możliwości dodawania końcówki identyfikującej osoby niebinarne w proponowanych formach równościowych. Monika Nowakowska

5 Propozycje tego typu rozwiązań uwzględnia się w publikacjach związanych z językiem włączającym czy w poradnictwie językowym (np. Chmura-Rutkowska i in. (red.) 2016).

i Katarzyna Radzikowska (2022) sugerują, że podkreślnik (tzw. podłoga) „jest sygnałem, że między formą żeńską a męską jest miejsce dla tych osób, które nie identyfikują się ani jako kobiety, ani jako mężczyźni". Niestety takie rozwiązanie nadal prowadzi do marginalizacji osób niebinarnych, gdyż ich tożsamość jest traktowana jako dodatkowa, a nie równorzędna z binarnymi kategoriami płci. Zamiast wspierać inkluzywność, proponowane rozwiązania mogą prowadzić do dalszego wykluczenia i komplikacji w odbiorze języka. Dlatego istnieje wyraźna potrzeba bardziej intuicyjnego i inkluzywnego zapisu form równościowych, które uwzględniałyby wszystkie tożsamości płciowe.

2.2. Skrócona forma równościowa z asteryskiem inkluzywnym

Skróconą formę równościową z asteryskiem inkluzywnym (zob. rys. dla przykładowej formy *urzędni*k/czka/cze*) tworzą *trzon* i ciąg *sufiksów genderowych* determinujących rodzaj naturalny i gramatyczny wyrażenia pochodnego. Granica trzonu i ciągu sufiksów genderowych jest oznaczona *asteryskiem inkluzywnym* (*). Asterysk inkluzywny (niem. *Genderstern*, ang. *gender star*) jest jednym z markerów form inkluzywnych w języku niemieckim (przykłady użycia można znaleźć w: Duden 2020; DIN 5008 2020), który staje się coraz bardziej popularny i powszechnie używany (por. tabela w RdR 2021: 4). Nasza koncepcja form równościowych różni się strukturalnie od podejścia niemieckiego, ale elementem wspólnym jest symbol asterysku. W naszym podejściu asterysk inkluzywny ma dwie funkcje: demarkacyjną i konkatenacyjną. Po pierwsze, asterysk wyznacza granicę między trzonem formy a ciągiem sufiksów genderowych[6]. Po drugie, w wyniku konkatenacji trzonu z jednym z sufiksów genderowych powstaje wyrażenie pochodne specyficzne dla danej płci.



Rysunek. Schemat skróconych form równościowych z asteryskiem inkluzywnym

6 Używane w artykule terminy *trzon* oraz *sufiks genderowy* nie są definicyjnie tożsame z pojęciami *rdzeń* oraz *sufiks/przyrostek* stosowanymi w językoznawstwie. Różnica polega m.in. na tym, że w opisywanych strategiach zapisu form równościowych słowotwórcze formanty stanowią część trzonu. Przyjęta strategia zapisu, w której wyraz dzielony jest na trzon i sufiks genderowy, wydaje się sprzeczna z językoznawczymi terminami, ale podyktowana jest względami praktyczności i intuicyjności naszego rozwiązania, jako że może być ono wykorzystywane przez wszystki*ch/e użytkowni*ków/czki języka polskiego. Zakres pojęciowy wyrazów: *trzon* oraz *sufiks genderowy* został określony w dalszej części artykułu.

Trzon (*urzędni-* na rys.) to ciąg znaków wspólny dla wszystkich wyrażeń nacechowanych genderowo. Ten ciąg znaków może odpowiadać rdzeniowi (morfemowi leksykalnemu) poszczególnych wyrażeń, ale nie jest to wymaganie konieczne.

Sufiksy genderowe to ciągi znaków inherentne dla poszczególnych wyrażeń (tj.: maskulatywu, feminatywu i neutratywu). Sufiksy genderowe są połączone ukośnikiem. W naszym podejściu ukośnik jest symbolem rozdzielającym alternatywne ciągi znaków lub wyrażenia alternatywne. Sufiksy genderowe występują w określonej kolejności. Pierwszy jest męski sufiks genderowy (*-k* na rys.), a po nim żeński sufiks genderowy (*-czka* na rys.). Kolejność ta jest motywowana względami ekonomii języka – męskie sufiksy genderowe są zwykle krótsze niż żeńskie. Po tych dwóch sufiksach genderowych charakterystycznych dla mężczyzn i kobiet można wprowadzić sufiks genderowy tworzący z trzonem wyrażenie niebinarne, np. neutratyw lub iksatyw. Ponieważ nie ma jeszcze jednej normatywnej niebinarnej formy gramatycznej w języku polskim, to sufiks genderowy *-cze* został dodany w przykładzie wyłącznie do celów ilustracyjnych, żeby pokazać możliwości proponowanej koncepcji. Nie narzucamy konkretnej formy niebinarnej i w dalszej części artykułu będziemy stosować wyłącznie formy z równoległymi maskulatywami i feminatywami istniejącymi w języku polskim.

Między elementami formy równościowej zachodzą związki paradygmatyczno-syntagmatyczne. Relacje paradygmatyczne zachodzą pomiędzy alternatywnymi sufiksami genderowymi, które mogą być wzajemnie zastępowane w tym samym kontekście. Relacja syntagmatyczna zachodzi między trzonem wyrazu a wybranym sufiksem genderowym, które łączą się w poprawne wyrażenie językowe w określonym rodzaju, np.: *urzędnik*, *urzędniczka* i *urzędnicze* w przykładzie z rysunku.

Formy z asteryskiem, takie jak *urzędni*k/czka* czy *kole*dzy/żanki*, mogą być realizowane słownie na różne sposoby. Przykładowo można je rozwinąć w pełne wyliczenie: *urzędnik, urzędniczka*, *koledzy, koleżanki*. Inną możliwością jest zastosowanie koordynacji spójnikowej: *urzędnik lub urzędniczka*, *koledzy i koleżanki*. Wybór odpowiedniego spójnika zależy od kontekstu wypowiedzi – spójniki łączne stosuje się w wyrażeniach o znaczeniu logicznej koniunkcji, a spójniki rozłączne w wypadku alternatywy rozłącznej.

Za opracowaniem koncepcji skróconej formy równościowej z asteryskiem inkluzywnym stoi wieloaspektowa motywacja. Po pierwsze, naszym nadrzędnym celem było wypracowanie skróconej formy w pełni realizującej zasady równości, w której ani maskulatyw, ani feminatyw nie będzie wyeksponowany, jak to się dzieje w formach z podkreślnikiem, nawiasem czy ukośnikiem, gdzie jedno wyrażenie pojawia się w całości, a drugie jest reprezentowane wyłącznie przez sufiks genderowy lub w postaci wydzielonego ciągu znaków. Co więcej, w formach z asteryskiem istnieje możliwość dodania kolejnych sufiksów genderowych, które w połączeniu z trzonem mogą odnosić się do osób niebinarnych. Uważamy, że formy z asteryskiem są najbardziej inkluzywnymi formami, jakie zostały opracowane w języku polskim. Po drugie, stosowanie form z asteryskiem inkluzywnym wydaje się użyteczne w automatycznym przetwarzaniu tekstu. Formy z asteryskiem można stosunkowo łatwo konwertować na inne warianty notacji równościowej lub pojedyncze maskulatywy bądź feminatywy (np. w celu opracowania odrębnych spersonalizowanych regulaminów studiów dla studentów i studentek). Po trzecie, celem wspólnym wszystkich skróconych form równościowych jest zachowanie ekonomii

języka przez redukcję powtarzających się ciągów znaków, przy jednoczesnym zachowaniu efektywności przekazu treści. Proponowane formy równościowe, zob. (5b), są dłuższe niż formy w generycznym rodzaju męskim (5), ale mają zdecydowanie niższy poziom redundancji niż formy współwystępujące, które łączą maskulatywy i feminatywy za pomocą spójników współrzędnych lub ukośników (5a). Ponadto skrócone formy z asteryskiem umożliwiają notację wszystkich uzgodnień gramatycznych bez konieczności wprowadzania podwójnych form czasownikowych czy przymiotnikowych.

(5) *Urzędnik państwowy mianowany* podlega okresowym ocenom kwalifikacyjnym. Oceny dokonuje *kierownik* urzędu, w którym *urzędnik* jest *zatrudniony*, biorąc pod uwagę opinię *bezpośredniego przełożonego* i zakładowej organizacji związkowej. Dla dokonania tej czynności *kierownik* urzędu może powołać komisję[7].

(5a) *Urzędnik państwowy mianowany/Urzędniczka państwowa mianowana* podlega okresowym ocenom kwalifikacyjnym. Oceny dokonuje *kierownik/kierowniczka* urzędu, w którym *urzędnik/urzędniczka* jest *zatrudniony/zatrudniona*, biorąc pod uwagę opinię *bezpośredniego przełożonego/bezpośredniej przełożonej* i zakładowej organizacji związkowej. Dla dokonania tej czynności *kierownik/kierowniczka* urzędu może powołać komisję.

(5b) *Urzędni*k/czka państwow*y/a mianowan*y/a* podlega okresowym ocenom kwalifikacyjnym. Oceny dokonuje *kierowni*k/czka* urzędu, w którym *urzędni*k/czka* jest *zatrudnion*y/a*, biorąc pod uwagę opinię *bezpośrednie*go/j przełożone*go/j* i zakładowej organizacji związkowej. Dla dokonania tej czynności *kierowni*k/czka* urzędu może powołać komisję.

2.3. Szczególne przypadki

Integralność dwuznaków

Trzon formy równościowej jest ciągiem znaków wspólnych dla wyrażeń nacechowanych płciowo. Należy jednak zachować integralność dwuznaków, m.in. w zakresie oboczności. Przykładowo znak *r* należy do wyrażeń *lekarz* i *lekarka*, ale w pierwszym wypadku jest elementem dwuznaku *rz*, dlatego w formie równościowej *r* nie należy do trzonu, a forma równościowa powinna mieć następującą postać: *leka*rz/rka*.

Zerowy sufiks genderowy

Częstym przypadkiem w polszczyźnie są zerowe męskoosobowe sufiksy genderowe, występujące najczęściej w mianowniku liczby pojedynczej. W takich wypadkach cały maskulatyw jest trzonem, który jest połączony asteryskiem z żeńskim sufiksem genderowym, np. *student*ka*. Zapis form z zerowym żeńskim sufiksem genderowym jest analogiczny, np. *sędzi*ego*. Przykładowe formy z zerowymi sufiksami genderowymi należy interpretować jako *student*∅/ka* i *sędzi*ego/∅*, gdzie ∅ oznacza zerowy sufiks genderowy (pusty ciąg znaków). Wprowadzenie dodatkowego symbolu zerowego sufiksu genderowego wydaje się

7 Fragment tekstu z: https://sip.lex.pl/akty-prawne/dzu-dziennik-ustaw/pracownicy-urzedow-panstwowych-16791016/roz-3 (dostęp: 1 lipca 2025).

redundantne. W celu ułatwienia automatycznego przetwarzania tekstów (tzn. bez konieczności dodatkowej identyfikacji rodzaju gramatycznego pojedynczego sufiksu genderowego) rozważano pozostawienie ukośnika alternatywnego jako markera zerowego sufiksu genderowego, tj.: *student*/ka* i *sędzi*ego/*. Zrezygnowano jednak z tego dodatku ze względu na potencjalnie wysokie ryzyko błędów wynikających z niewłaściwego umieszczenia ukośnika lub jego pominięcia w tekście.

Alternatywność form

Dostarczając narzędzie do skracania współwystępujących maskulatywów i feminatywów, ich modyfikatorów oraz uzgadniających się z nimi form czasownikowych, dopuszczamy stosowanie w jednym tekście również innych wariantów form inkluzywnych, tj. form współwystępujących połączonych ukośnikiem alternatywnym oraz skoordynowanych maskulatywów i feminatywów, jeśli jest to motywowane kontekstowo. Formy współwystępujące stosuje się także, gdy wyrażenia nacechowane płciowo nie mają części wspólnej, zob. (6), lub ich trzon jest bardzo krótki, np. *jego/jej życiorys* albo *je*go/j życiorys*.

(6) Kole*dzy/żanki pojecha*li/ły na wycieczkę. Spotkałam *ich/je* w tramwaju.

W jednym tekście nie należy stosować alternatywnych wariantów skróconych form równościowych (np. z asteryskiem i podkreślnikiem).

Zaimki *każdy* i *wszyscy*

Zaimki *każdy* i *wszyscy* mogą być używane rzeczownikowo, opisując grupę osób. W odniesieniu do grupy mieszanej lub ogółu społeczeństwa takie wyrażenia są wykluczające, zob. (7), dlatego proponujemy trzy rozwiązania zależne od gatunku tekstu: dodanie słowa *człowiek* (7a), zamianę na *każda osoba* (7b) lub użycie formy z asteryskiem (7c).

(7) *Każdy* ma prawo do nauki.

(7a) *Każdy człowiek* ma prawo do nauki.

(7b) *Każda osoba* ma prawo do nauki.

(7c) *Każd*y/a* ma prawo nauki.

Użycia generyczne form czasownikowych

Zaleca się unikanie generycznego rodzaju męskiego, aby wspierać równość i inkluzywność językową. Jednak w niektórych kontekstach, zob. (8), zmiana formy na bardziej neutralną może zmienić znaczenie (np.: *umówił się z nią*, *umówiła się z nim*) lub skutkować błędami (np. *umówił*a się*). W takich wypadkach użycie generycznego rodzaju męskiego wydaje się uzasadnione, gdyż pozwala zachować właściwe znaczenie i poprawność gramatyczną.

(8) Na spotkanie z Jeffreyem przyszła spóźniona, *umówili się* w kawiarni nieopodal opery.

3. Uwrażliwianie języka na płeć a gatunek tekstu – przykładowy zestaw rozwiązań dla dużych modeli językowych

John Swales (1990: 58) definiuje *gatunek* (ang. *genre*) jako klasę zdarzeń komunikacyjnych, w których uczestniczą osoby mające wspólny zestaw celów komunikacyjnych. Gatunki mają różne cele komunikacyjne i mogą różnić się strukturą, stylem i treścią. Uwrażliwianie języka na płeć wymaga uwzględnienia specyficznych cech i celów różnych gatunków tekstów. W ramach prowadzonych badań analizujemy rozmaite gatunki (akty prawne, dokumenty urzędowe, wypowiedzi ustne, teksty narracyjne, artykuły naukowe i artykuły prasowe) pod kątem użycia w nich odpowiednich strategii inkluzywnych przez duże modele językowe (zob. rozdz. 4).

Akty prawne

Akty prawne wymagają szczególnej ostrożności ze względu na potencjalny wpływ zmian inkluzywnych na interpretację prawną. Warto też zwrócić uwagę na gatunkowe niuanse – istnieją gatunki hybrydowe, np. fragmenty mówione w dokumentach pisanych (np. treść przysięgi w akcie prawnym), które wymagają podejścia uwzględniającego ich złożoną naturę. Wydaje się, że wybranie jednej strategii wspólnej dla wszystkich aktów prawnych jest niemożliwe. Wprowadzenie form inkluzywnych do tekstu aktu prawnego jest warunkowane przez różne czynniki: długość zmienianego ustępu, złożoność zdania (pojedyncze lub złożone), poziom skomplikowania języka prawnego i część aktu (np. preambuła).

Treść polskich przepisów jest silnie zmaskulinizowana. Nawet Konstytucja jako ustawa zasadnicza, która w art. 33 zakazuje dyskryminacji w życiu rodzinnym, politycznym i społecznym, nie przewiduje feminatywów. Już w preambule czytamy bowiem tylko o obywatelach, zob. (9). Nie ma jednak przeciwwskazań, aby tak ważny akt prawny nie uwzględniał feminatywów, por. (9a). Jedną ze strategii wartych stosowania w przepisach jest współwystępowanie maskulatywów i feminatywów (⚥), szczególnie w uroczystych wstępach do tych dokumentów. Współwystępowanie można stosować w artykułach zawierających zdania pojedyncze i ta strategia nie zaburza odbioru treści, por. (10) i (10a).

(9) my, Naród Polski – wszyscy *obywatele* Rzeczypospolitej;

(9a) my, Naród Polski – *wszyscy obywatele i wszystkie obywatelki* Rzeczypospolitej;

(10) Art. 4. 2. Naród sprawuje władzę przez *swoich przedstawicieli* lub bezpośrednio;

(10a) Art. 4. 2. Naród sprawuje władzę przez *swoich przedstawicieli i swoje przedstawicielki* lub bezpośrednio.

W artykułach, które zawierają rozbudowane ustępy lub trudne w odbiorze sformułowania charakterystyczne dla języka prawnego, praktycznym rozwiązaniem wydaje się wykorzystanie formy skróconej z asteryskiem inkluzywnym (*⚥). Dzięki temu unikamy nadmiernego wydłużenia tekstu, co mogłoby utrudniać jego czytelność i zrozumienie samych przepisów,

por. (11) i (11a) z (11b), i zapobiegamy nieekonomicznemu rozbudowaniu np. składu papierowego wydania danej ustawy.

(11) Art. 647[1]. § 1. *Inwestor* odpowiada solidarnie z *wykonawcą* (*generalnym wykonawcą*) za zapłatę wynagrodzenia należnego *podwykonawcy* z tytułu wykonanych przez niego robót budowlanych, których szczegółowy przedmiot został zgłoszony *inwestorowi* przez *wykonawcę* lub *podwykonawcę* przed przystąpieniem do wykonywania tych robót, chyba że w ciągu trzydziestu dni od dnia doręczenia *inwestorowi* zgłoszenia *inwestor* złożył *podwykonawcy* i *wykonawcy* sprzeciw wobec wykonywania tych robót przez *podwykonawcę*.

(11a) Art. 647[1]. § 1. *Inwestor*ka* odpowiada solidarnie z *wykonaw*cą/czynią* (*generaln*ym/ą wykonaw*cą/czynią*) za zapłatę wynagrodzenia należnego *podwykonaw*cy/czyni* z tytułu wykonanych przez *niego/nią* robót budowlanych, których szczegółowy przedmiot został zgłoszony *inwestor*owi/ce* przez *wykonaw*cę/czynię* lub *podwykonaw*cę/czynię* przed przystąpieniem do wykonywania tych robót, chyba że w ciągu trzydziestu dni od dnia doręczenia *inwestor*owi/ce* zgłoszenia *inwestor*ka* złożył*a *podwykonaw*cy/czyni* i *wykonaw*cy/czyni* sprzeciw wobec wykonywania tych robót przez *podwykonaw*cę/czynię*.

(11b) Art. 647[1]. § 1. *Inwestor/Inwestorka* odpowiada solidarnie z *wykonawcą/wykonawczynią* (*generalnym wykonawcą/generalną wykonawczynią*) za zapłatę wynagrodzenia należnego *podwykonawcy/podwykonawczyni* z tytułu wykonanych przez *niego/nią* robót budowlanych, których szczegółowy przedmiot został zgłoszony *inwestorowi/inwestorce* przez *wykonawcę/wykonawczynię* lub *podwykonawcę/podwykonawczynię* przed przystąpieniem do wykonywania tych robót, chyba że w ciągu trzydziestu dni od dnia doręczenia *inwestorowi/inwestorce* zgłoszenia *inwestor/inwestorka* złożył/złożyła *podwykonawcy/podwykonawczyni* i *wykonawcy/wykonawczyni* sprzeciw wobec wykonywania tych robót przez *podwykonawcę/podwykonawczynię*.

W polskim ustawodawstwie zdarzają się wyjątki od reguły maskulinizacji przepisów. Przykładem może być Ustawa z dnia 15 lipca 2011 r. o zawodach pielęgniarki i położnej, która zawiera wyłącznie żeńskie formy zawodu: *pielęgniarka* i *położna*. Ta sama ustawa w art. 8. ust. 1 dopuszcza stosowanie formy *pielęgniarz* czy *położny*, zob. (12), ale w pozostałej treści ustawy te maskulatywy nie występują, chociaż strategia ⚥ mogłaby być stosowana w całym dokumencie.

(12) Art. 8. 1. Osobie posiadającej prawo wykonywania zawodu przysługuje prawo posługiwania się tytułem zawodowym „pielęgniarka" albo „pielęgniarz", „położna" albo „położny".

Ponieważ akty prawne w większości odnoszą się do ogółu społeczeństwa, stosowanie strategii adaptacji nazwy do płci (♀) jest ograniczone. W przepisach stosuje się natomiast formy neutralne, np.: *każdy człowiek* (art. 8 k.c.), *kto zaginął* (art. 30 k.c.), *człowiek* (art. 118a k.k.), *osoba najbliższa* (art. 157 k.k.). Analogicznie można zastosować neutralizację (♀) w stosunku do innych rzeczowników osobowych w rodzaju męskim odnoszących się do kobiet i mężczyzn, np.: *oskarżony* i *pokrzywdzony* (art. 45a k.k.) to odpowiednio *osoba oskarżona* i *osoba pokrzywdzona*.

Stosując formy inkluzywne w aktach prawnych, szczególnie w aktach normatywnych, należy uwzględniać sposób definiowania przez te przepisy takich podmiotów, jak wspomniani wcześniej *inwestor*, *wykonawca* czy *przedsiębiorca*. Niezbędna jest akceptacja środowiska prawniczego dla wprowadzanych zmian inkluzywnych, gdyż feminatywy czy formy neutralne, które wzbudzają kontrowersje i których ustawodawca nie uwzględnił w powszechnie obowiązujących przepisach, mogą stanowić powód do podważania np. umów sporządzonych z uwzględnieniem form inkluzywnych na podstawie przepisów niezawierających definicji rozszerzonych o te formy.

Dokumenty urzędowe

W normatywnych aktach prawnych (np. przepisach prawa wewnętrznego, europejskiego i międzynarodowego) czy aktach prawnych o charakterze wewnętrznym (np. statuty, regulaminy, formularze) i w innych szeroko pojętych dokumentach urzędowych właściwą strategią wydaje się stosowanie form *⚥ zamiast dłuższych form ⚥, które mogłyby negatywnie wpłynąć nie tylko na zwięzłość tekstu i czas potrzebny na jego przyswojenie, ale też na długość np. formularza urzędowego. W zależności od kontekstu, jeśli tylko nie niesie to ze sobą ryzyka wykluczeń, dopuszcza się również strategie ♀ i ♀ (np. w ankietach skierowanych wyłącznie do jednej płci).

Wypowiedzi ustne

W gatunkach mówionych, np. przemówieniach, zaleca się strategie ⚥ oraz ♀. Użycie form *⚥ nie jest odpowiednie ze względu na brak możliwości ich fonetycznej realizacji. Neutralizacja również powinna zostać odrzucona w gatunkach mówionych, ponieważ jako bezosobowa i niedookreślona nie buduje emocjonalnej więzi z adresat*ami/kami, do których kierowane są słowa, a jednym z głównych celów ustnych form wypowiedzi jest wzbudzenie w odbior*cach/czyniach określonych emocji i reakcji.

Teksty narracyjne

Wprowadzanie form inkluzywnych w tekstach narracyjnych jest ograniczone i nie jest równoznaczne ze swobodnym wyborem między maskulatywem a feminatywem, jak to się dzieje np. w aktach prawnych, gdzie wybór zależy od płci osoby odczytującej lub osoby, której dany akt dotyczy. W definicjach zawartych w aktach prawnych rzeczowniki nazywające zawody (np. *pracownik naukowy*) zazwyczaj występują w formie męskiej, ale domyślnie odnoszą się również do kobiet. W tekstach narracyjnych to autor*ka prowadzi narrację, osadza ją we wskazanym przez siebie czasie i miejscu oraz określa bohater*ów/ki, narzucając czytelni*kowi/czce swoją wolę. Warto jednak stosować strategie ⚥ i ♀ tam, gdzie jest to możliwe i nie fałszuje fabuły przekazywanej opowieści (np.: *mieszkańcy i mieszkanki Nibylandii*, *hakerka Lisbeth Salander*). Osobatywy i wyrażenia uogólnione (⚲) wydają się odpowiednie, jeśli autor*ka nie chce zdradzać tożsamości bohater*a/ki albo odnosi się do bohater*ów/ek, których tożsamość jest nieistotna (np. *osoby zwiedzające zamek*). Formy z asteryskiem nie powinny być stosowane w tekstach narracyjnych ze względu na ich formalny wydźwięk. Odrębną kwestią są tłumaczenia tekstów narracyjnych, w których kluczową rolę odgrywa język źródłowy tłumaczonego tekstu. W kontekście form inkluzywnych szczególnie problematyczne są języki,

w których nie występują pary rzeczowników równościowych. W takich wypadkach proces tłumaczeniowy jest bardziej złożony, ponieważ wymaga nie tylko precyzyjnego oddania treści, ale także analizy językowej, aby zapewnić jego poprawność (np. sięgnięcia do zaimkowego poprzednika anaforycznego angielskiego *teacher*, by upewnić się, czy chodzi o *nauczyciela*, czy o *nauczycielkę*).

Artykuły naukowe i prasowe

W artykułach można stosować strategie ⚥ i *⚥. Zazwyczaj płeć czytelni*ka/czki nie jest znana, dlatego zaleca się stosowanie form współwystępujących w wyrażeniach adresatywnych. Ponadto formy ⚥ i *⚥ powinny być używane w odniesieniu do grup mieszanych (np. *uczestnicy i uczestniczki protestu*) lub osób o nieznanej płci (np.: *spraw*ca/czyni wypadku*, *poszkodowan*i/e pasażer*owie/ki*). W wypadku takich grup i osób możliwe jest również stosowanie strategii ⚲. Rekomenduje się także adaptację nazwy do płci w sytuacjach, gdy artykuł odnosi się bezpośrednio do osób publicznych oraz autor*a/ki artykułu.

Punktem wyjścia przedstawionych tu rozważań są akty prawne, ponieważ bezpośrednio kształtują nasze codzienne życie na wielu płaszczyznach, również tej językowej. Jeżeli język jest silnie zmaskulinizowany (np. Kodeks cywilny) lub sfeminizowany (np. Ustawa z dnia 15 lipca 2011 r. o zawodach pielęgniarki i położnej), rzutuje to na sposób postrzegania rzeczywistości przez jego użytkowni*ków/czki Przepisy prawne stanowią fundament dokumentów urzędowych, np. formularzy wypełnianych codziennie przez dziesiątki tysięcy osób, programów nauczania w szkołach i przekazu medialnego. Słownictwo używane w tych kanałach przenika do naszej codziennej komunikacji, a tym samym do danych tekstowych, na których są trenowane modele językowe. Zamieszczona poniżej tabela zawiera podsumowanie strategii uwrażliwiania języka na płeć, które można stosować w analizowanych gatunkach tekstów, tworząc zbiór danych inkluzywnych do trenowania dużych modeli językowych.

Tabela. Proponowane strategie uwrażliwiania języka na płeć w różnych gatunkach tekstu

GATUNEK TEKSTU	⚥	*⚥	♀	⚲	UWAGI
AKTY PRAWNE	⊖	⊕	⊕	⊕	Należy zachować szczególną ostrożność ze względu na możliwy wpływ zmian na interpretację prawną; zalecana jest konsultacja prawnicza.
FORMULARZE I DOKUMENTY URZĘDOWE	⊖	⊕	⊕	⊕	Należy stosować skrócone formy równościowe w celu zachowania zwięzłości tekstu.
PRZEMÓWIENIA I WYPOWIEDZI USTNE	⊕	⊖	⊕	⊖	Formy *⚥ i ⚲ są nieadekwatne w mowie.

FRAGMENTY MÓWIONE W DOKUMENTACH PISANYCH	⊕	⊖	⊕	⊖	W dokumentach pisanych z fragmentami mówionymi zaleca się stosowanie form ⚥ i ♀, aby zachować spójność z tekstem mówionym.
TEKSTY NARRACYJNE	⊕	⊖	⊕	⊕	Stosowanie strategii jest zależne od kontekstu; formy *⚥ nie są wskazane.
ARTYKUŁY NAUKOWE	⊕	⊕	⊕	⊕	Stosowanie wszystkich strategii jest zależne od kontekstu.
ARTYKUŁY PRASOWE	⊕	⊕	⊕	⊕	Stosowanie wszystkich strategii jest zależne od kontekstu.

Oznaczenia strategii: ⚥ – feminizacja: współwystępowanie maskulatywów i feminatywów z ukośnikiem lub spójnikiem; *⚥ – feminizacja: skrócone formy równościowe z asteryskiem inkluzywnym; ♀ – feminizacja: adaptacja feminatywów; ♀ – neutralizacja

4. Możliwości dużych modeli językowych w zakresie uwrażliwiania na płeć

4.1. Założenia eksperymentu

Duże modele językowe (LLM-y, ang. *Large Language Models*) stanowią obecnie ważne narzędzie wykorzystywane powszechnie do generowania i korygowania tekstów, tłumaczenia, asystowania w komunikacji czy procesie przyswajania wiedzy. W ciągu ostatnich kilku lat zdobyły one ogromną popularność dzięki platformie ChatGPT, także wśród osób, które do tej pory nie były związane z nowymi technologiami. W niniejszym badaniu przeprowadzono eksperyment empiryczny mający na celu zweryfikowanie zdolności LLM-ów do przekształcania polskich tekstów na ich wersje prorównościowe.

Duże modele językowe mogą mieć wbudowany moduł do rozpoznawania i generowania wyrażeń równościowych w językach z systemem rodzajowym zgodnie z preferencjami użytkowni*ków/czek. Alternatywnie modele mogą uczyć się wzorców inkluzywnych w fazie trenowania lub dostrajania na danych tekstowych w języku równościowym. LLM-y są trenowane na podstawie ogromnych zbiorów danych tekstowych – wielojęzycznych, ale z dominującą przewagą tekstów anglojęzycznych (por. W3Techs 2024). Język angielski nie determinuje rodzaju gramatycznego i nie ma potrzeby wprowadzania form inkluzywnych w anglojęzycznych tekstach, dlatego problem uwrażliwiania na płeć jest marginalizowany przez twór*ców/czynie modeli językowych opartych głównie na danych angielskich. Jednak w kontekście języków, które mają rozbudowany system znaczników rodzaju gramatycznego, takich jak polski, niemiecki czy francuski, konieczność uwzględniania inkluzywnych wyrażeń językowych staje się kluczowym wyzwaniem.

Celem naszego eksperymentu jest przekształcenie tekstów na ich wersje równościowe, a następnie ręczna analiza automatycznie wygenerowanych wyrażeń równościowych. W pierwszym kroku wykorzystujemy technikę *promptowania* – testowany LLM konwertuje tekst

wejściowy na jego wersję równościową zgodnie z precyzyjnymi instrukcjami zdefiniowanymi w poleceniu (tzw. *prompcie*). W prompcie[8] instruujemy LLM, że powinien wcielić się w rolę redaktora tekstów polskich, który jest świadomy wad standardowego języka polskiego. Precyzujemy algorytm działania modelu-redaktora, który najpierw określa gatunek, a potem przekształca zidentyfikowane wyrażenia wykluczające zgodnie ze strategią uwrażliwiania odpowiednią dla tego gatunku. LLM musi dokonać obowiązkowych przekształceń i może wdrożyć preferowane przekształcenia. Niedopuszczalne przekształcenia są zabronione. Prompt zawiera również przykłady przekształceń inkluzywnych fragmentów różnych gatunków tekstów.

4.2. Analiza wyników

W eksperymencie sprawdzono dziewięć modeli językowych pod kątem umiejętności generowania wyrażeń równościowych: Bielik, Claude, Command R+, Gemini, GPT, Llama, Mistral, Mixtral i Nemotron[9]. Szczegółowej analizie zostały poddane fragmenty *Karty Praw Podstawowych UE* (Rozdz. V Prawa obywatelskie: art. 39, art. 40, art. 41, art. 42, art. 43) oraz fragmenty sprawozdań stenograficznych z posiedzeń Sejmu z korpusu ParlaMint (Ogrodniczuk i in. 2022). Wyniki analizy znajdują się w arkuszu kalkulacyjnym http://zil.ipipan.waw.pl/analiza-inkluzywna w zakładkach *Karta Praw Podstawowych UE* i *ParlaMint Sejm 2022*.

Analiza użycia strategii inkluzywnych

Modele stosowały różne strategie uwrażliwiania języka na płeć – niektóre wdrożyły wszystkie dostępne strategie, a niektóre ograniczyły się tylko do dwóch lub nawet jednej. Jedynym modelem, który dostosował strategię do gatunku tekstu, był Mixtral – użył *⚥ w *Karcie Praw* (akt prawny), a ⚥ w wypowiedziach ustnych z ParlaMint. Claude, Command R+ i Llama użyły *⚥ w obu gatunkach, a Mistral, GPT i Bielik nie wdrożyły jej w ogóle. Gemini i Nemotron postąpiły niezgodnie z instrukcją, wykorzystując *⚥ w transkrypcjach wypowiedzi i niedozwolone formy skrócone z ukośnikiem w *Karcie Praw*. W *Karcie Praw* wszystkie LLM-y z wyjątkiem Nemotron i Gemini stosowały poprawne formy ♀. W transkrypcjach formy ♀ zostały poprawnie wprowadzone przez Claude, Command R+ i GPT, przy czym GPT jako jedyny zaproponował alternatywę neutralną dla słowa *głupiec*, por. (13) i (13a):

(13) Ale czasem nawet do biednego *głupca* los się uśmiechnie.

(13a) Ale czasem nawet do biednej *osoby* los się uśmiechnie.

Analiza form z asteryskiem

Generowanie poprawnych form *⚥ stanowiło nie lada wyzwanie dla modeli. Najlepiej poradził sobie Claude, ponieważ zaproponowane przez niego formy zawierają tylko nieliczne błędy (np.: *psycholoż*kach/gach*, *lekar*ek/zy*). Model Llama poradził sobie z niektórymi

8 Prompt użyty do generowania równościowych wersji tekstów znajduje się w zakładce Prompt w arkuszu https://zil.ipipan.waw.pl/analiza-inkluzywna.

9 Stosujemy skrócone nazwy modeli bez precyzowania ich wersji. Dokładne wersje testowanych modeli są w arkuszu z wynikami.

trudnymi przekształceniami (np. *psycholo*gach/żkach*), ale w innych popełnił błędy (np. *sprawozdawc*a/a*). Command R+ wygenerował formy z asteryskiem, które są albo niezgodne z naszą koncepcją (np.: *ka*żd*y/a*, *specjalist*ów/stek*, *prawnik*ami/kami*, *Pawł* Kow*al*/ka*), albo spójne tylko z niemieckim podejściem (np. *obywatele*ki*). Pozostałe modele nie stosowały form *⚥. Na podstawie przeprowadzonej analizy wnioskujemy, że sam opis koncepcji i podane przykłady w instrukcji są niewystarczające dla LLM-ów, żeby efektywnie przyswoić nową koncepcję skracania form równościowych. Modele te polegają w dużym stopniu na wzorcach występujących w danych treningowych, dlatego konieczne może być dostarczenie większej liczby przykładów w fazie trenowania, żeby nauczyć model poprawnie rozpoznawać formy równościowe i generować je w sposób zgodny z założeniami inkluzywności językowej.

Zmiany inkluzywne

LLM-y miały większe trudności z konwersją transkrypcji wypowiedzi, ale teksty artykułów również nie zostały przekształcone bezbłędnie. W *Karcie Praw* najmniej błędów popełniły Claude i Gemini, a w wypowiedziach sejmowych – Claude i GPT. Biorąc pod uwagę kompletność automatycznych przekształceń, zauważalny jest brak niektórych wyrażeń równościowych, które znajdują się we wzorcowej inkluzywnej wersji tekstu. W tytule artykułu 43 tylko Claude i Command R+ wprowadziły formę równościową, Gemini, Bielik, Llama, Mistral, Mixtral i Nemotron pozostawiły męską formę *rzecznik*, a GPT zastosował też wykluczający feminatyw *rzeczniczka*. GPT, Bielik, Command R+, Gemini, Llama, Mistral i Mixtral zastosowały formy ⚥ na poziomie rzeczowników w transkrypcjach, lecz nie przekształciły męsko-osobowych form czasowników w ich odpowiedniki równościowe. Częstymi błędami w przekształceniach były brak konsekwencji we wprowadzaniu zmian: stosowanie formy neutralnej, a następnie powrót do maskulatywu, np. *na takich samych warunkach jak osoby będące o b y w a t e l a m i* (GPT), oraz brak spójności w stosowaniu *każd*y/a*, np. *każdy obywatel/ka* (Nemotron).

Najmniej inkluzywnych zmian wprowadził Bielik, który najczęściej powielał generyczne formy męskie rzeczowników i czasowników, stosował złożenia analityczne z *pan/pani* lub łączył formę neutralną z maskulatywem w obrębie tego samego akapitu. Model wprowadzał także zmiany skutkujące zmienioną interpretacją tekstu, np. zastąpił całą frazę *każdy obywatel Unii i każda osoba fizyczna lub prawna* wyrazem *osoba*. Zaskakujące są również przypadki zmiany nazwisk polityków na nazwiska polityczek oraz nazw ugrupowań politycznych, por. (14) i (14a):

(14) Pan poseł *Paweł Kowal*, Klub Parlamentarny *Koalicja Obywatelska*.

(14a) Pani poseł *Joanna Mucha*, *Platforma Obywatelska*.

Innym typem błędów są redundantne zmiany inkluzywne. Modele wprowadzają współwystępowanie i feminatywy tam, gdzie z kontekstu wynika, że chodzi wyłącznie o mężczyznę, np.; *dopóki minister*ka nie zmieni swojej polityki* (Claude), *moja przedmówczyni/mój przedmówca* (Gemini, GPT, Nemotron). Można znaleźć też inne zmiany w obrębie przekształceń inkluzywnych, np. Command R+ zmienia: *Członków Parlamentu Europejskiego* na *Posłów i Posłanki*, *osoby fizyczne i prawne* na *rezydent*a/kę i jednostkę organizacyjną*.

Zmiany nieinkluzywne

Mimo precyzyjnych instrukcji, aby nie parafrazować fragmentów tekstu, które nie zawierają wyrażeń wykluczających, nie usuwać wyrazów i nie dodawać nowego tekstu niezwiązanego z przekształceniami prorównościowymi, LLM-y wprowadzają zmiany w oryginalnej treści. Czasem pomijają fragmenty oryginalnego tekstu, np. Claude opuścił frazę *w głosowaniu wolnym i tajnym* w art. 39 ust. 2 *Karty Praw*. Często wprowadzają dodatkową treść, por. oryginalną wypowiedź (15) i modyfikację Llamy (15a):

(15) Dziękuję serdecznie. Pan poseł Michał Gramatyka, Polska 2050. Proszę, Michale.

(15a) Dziękuję bardzo. Bardzo mi miło, że mogłem być dzisiaj tutaj. Chciałbym powiedzieć, że jako poseł Polski 2050, jestem bardzo zainteresowany tym, aby nasze państwo było bardziej przyjazne dla wszystkich obywateli, niezależnie od ich płci. Dlatego uważam, że używanie języka inkluzjonistycznego jest bardzo ważne. Chciałbym więc poprosić o przekształcenie mojej wypowiedzi, aby była ona bardziej inkluzywna.

Nasza analiza wykazała, że współczesne LLM-y nadal nie osiągnęły pełnej doskonałości w uwrażliwianiu tekstów na płeć i wykazują trudności w adekwatnym generowaniu tekstów inkluzywnych.

Podsumowanie

W niniejszym artykule przedstawiono nową koncepcję zapisu skróconych form równościowych z asteryskiem inkluzywnym. Proponowana forma w pełni realizuje zasady inkluzywności, nie faworyzując ani maskulatywu, ani feminatywu, a jej struktura umożliwia dodanie kolejnego sufiksu genderowego, co pozwala na opisanie osób niebinarnych. Ponadto forma ta umożliwia również notację wszystkich uzgodnień gramatycznych bez konieczności stosowania podwójnych form czasownikowych i przymiotnikowych.

Proponowana koncepcja wpisuje się w nurt deskryptywnego językoznawstwa wrażliwego na zmiany społeczne i językowe. Rozumiemy przez to, że zjawiska językowe są komentowane z perspektywy badacz*a/ki, któr*y/a wykorzystuje wiedzę naukową do proponowania potencjalnych modyfikacji i usprawnień opartych na systemie języka dla już używanych form, dla których było takie zapotrzebowanie. Nie wkraczamy tym samym w obszar normatywistyki, która z kolei zmiany narzuca albo broni języka przed ewentualnymi modyfikacjami, wyznaczając powszechny normatywny standard.

Odnosząc się do obecności form równościowych, m.in. z podkreślnikiem czy nawiasem, w publikacjach drukowanych i dyskursie internetowym, wskazujemy, że koncepcja form równościowych z asteryskiem inkluzywnym może stanowić propozycję ujednolicenia dotychczas stosowanych zapisów. Zwracamy uwagę na znaczenie ekonomii języka, zgodność formy z systemem językowym i łatwość adaptacji w obszarze przetwarzania języka naturalnego, co jest kluczowe w analizach ilościowych tekstów.

W ramach badań przeprowadzono eksperyment polegający na automatycznym przekształceniu tekstów w ich wersje równościowe z wykorzystaniem odpowiednio poinstruowanych

dużych modeli językowych. Wyniki wykazały, że współczesne modele nie są jeszcze w pełni skuteczne w uwrażliwianiu tekstów na płeć i wykazują trudności w generowaniu adekwatnych form inkluzywnych zgodnych z gatunkiem przetwarzanego tekstu. Jednym z kluczowych wyzwań jest potrzeba dostarczenia modelom spójnie zaanotowanych danych, które umożliwiłyby im naukę poprawnych wzorców prorównościowych. Aby to osiągnąć, niezbędne było opracowanie spójnej i kompleksowej koncepcji zapisu form równościowych, która będzie szeroko akceptowana i wdrażana przez użytkowni*ków/czki języka oraz umożliwi pozyskanie odpowiednich danych.

Asterysk inkluzywny wydaje się łatwy do adaptacji, ponieważ symbole gwiazdki i ukośnika są powszechnie dostępne na klawiaturach, a wyświetlanie tych form nie wymaga dodatkowych znaków ani specjalnych krojów pisma. Nie istnieją więc techniczne przeszkody, które uniemożliwiłyby stosowanie tego zapisu przez użytkowni*ków/czki zainteresowan*ych/e promowaniem równości językowej, co otwiera pole do dalszych badań w kontekście socjolingwistycznym.

Mamy nadzieję, że przeprowadzone badania oraz zaproponowana koncepcja form równościowych z asteryskiem inkluzywnym, stanowiąca nowatorskie podejście do problemu seksizmu językowego w polszczyźnie, przyczynią się do zwiększenia widoczności kobiet w języku oraz opracowywania krajowych standardów języka inkluzywnego, których w Polsce brakuje.

Spis modeli językowych

Bielik: K. Ociepa, Ł. Flis, K. Wróbel, A. Gwoździej, SpeakLeash Team, Cyfronet Team, *Introducing Bielik-7B-v0.1: Polish Language Model*, 2024 (online: https://huggingface.co/speakleash/Bielik-7B-v0.1, dostęp: 30 sierpnia 2024).

Claude: Anthropic, *The Claude 3 Model Family: Opus, Sonnet, Haiku*, 2024 (online: https://www-cdn.anthropic.com/de8ba9b01c9ab7cbabf5c33b80b7bbc618857627/Model_Card_Claude_3.pdf, dostęp: 30 sierpnia 2024).

Command R+: A. Gomez, *Introducing Command R+: A Scalable LLM Built for Business*, 2024 (online: https://cohere.com/blog/command-r-plus-microsoft-azure, dostęp: 30 sierpnia 2024).

Gemini: Google, *Gemini Pro*, 2024 (online: https://deepmind.google/technologies/gemini/pro, dostęp: 30 sierpnia 2024).

GPT: OpenAI, *Hello GPT-4o*, 2024 (online: https://openai.com/index/hello-gpt-4o/, dostęp: 30 sierpnia 2024).

Llama: Meta, *Introducing Meta Llama 3: The most capable openly available LLM to date*, 2024 (online: https://ai.meta.com/blog/meta-llama-3, dostęp: 30 sierpnia 2024).

Mistral i Mixtral: MistralAI, General purpose models, 2024 (online: https://mistral.ai/technology/#models, dostęp: 30 sierpnia 2024).

Nemotron: Z. Wang, Y. Dong, O. Delalleau, J. Zeng, G. Shen, D. Egert, J. Zhang, M.N. Sreedhar, O. Kuchaiev, *HelpSteer2: Open-source dataset for training top-performing reward models*, 2024 (online: https://arxiv.org/abs/2406.08673, dostęp: 30 sierpnia 2024).

Bibliografia

Bańko M., Linde-Usiekniewicz J., Łaziński M. 2021: *Rekomendacje dotyczące języka niedyskryminującego na Uniwersytecie Warszawskim* (online: https://www.uw.edu.pl/wp-content/uploads/2020/01/jezyk-niedyskryminujacy.pdf, dostęp: 21 sierpnia 2024).

Bojarska K. 2011: *Wpływ androcentrycznych i inkluzywnych płciowo konstrukcji językowych na skojarzenia z płcią*, „Studia Psychologiczne", t. 4, z. 2, s. 53–68.

CDEG 2007: *Steering Committee for Equality between Women and Men (CDEG)*, Committee of Ministers (online: https://search.coe.int/cm?i=09000016805d5072, dostęp: 6 sierpnia 2024).

Chmura-Rutkowska I. 2022: *(Nie)równość płci w życiu i w języku*, [w:] I. Chmura-Rutkowska, J. Szpyra-Kozłowska, *Nierówność płci w języku. Poradnik dla nauczycielek i nauczycieli szkół m.st. Warszawy*, Warszawskie Centrum Innowacji Edukacyjno-Społecznych i Szkoleń, Urząd m.st. Warszawy, Biuro Edukacji, Warszawa, s. 13–45.

Chmura-Rutkowska I., Duda M., Mazurek M., Sołtysiak-Łuczak A. (red.) 2016: *Gender w podręcznikach. Projekt badawczy. Raport*, t. 1, Wydawnictwo Feminoteka, Warszawa.

Czarnecka K. 2024: *Tranzycja dzieci. Co robić? Ekspertki dla „Polityki": Słuchajmy ich, nie popychajmy w żadną stronę*, „Polityka", nr 29(3472), s. 56–59.

DIN 5008 2020: *Schreib- und Gestaltungsregeln für die Text- und Informationsverarbeitung*, Beuth Verlag.

Duden 2020: *Online-Duden* (online: https://www.duden.de/rechtschreibung/Genderstern, dostęp: 6 sierpnia 2024).

Dyszak A.S. 2024: *Formy męsko- i żeńskoosobowe a niebinarni użytkownicy polszczyzny. U źródeł nowych zjawisk leksykalno-gramatycznych we współczesnym języku polskim*, „Język Polski" CIV, z. 1, s. 32–47.

GEG 2016: *Gender Equality Glossary*, Gender Equality Commission w Radzie Europy (online: https://edoc.coe.int/en/gender-equality/6947-gender-equality-glossary.html, dostęp: 30 sierpnia 2024).

Gębka-Wolak M. 2022: *Językowe wykładniki niebinarności płci w polszczyźnie. Część 2: Innowacje a system i tendencje rozwojowe*, „Prace Językoznawcze XXIV, z. 1, s. 101–116.

GNL-EN 2018: *Gender-Neutral Language in the European Parliament*, Parlament Europejski (online: https://www.europarl.europa.eu/cmsdata/151780/GNL_Guidelines_EN.pdf, dostęp: 9 sierpnia 2024).

GNL-PL 2018: *Język neutralny płciowo w Parlamencie Europejskim*, Parlament Europejski (online: https://www.europarl.europa.eu/cmsdata/187107/GNL_Guidelines_PL-original.pdf, dostęp: 18 lipca 2024).

Krysiak P., Małocha-Krupa A. 2020: Feminatywum, feminatyw, nazwa żeńska, żeńska końcówka – *problemy terminologiczne*, „Oblicza Komunikacji", t. 12, s. 229–238.

Lindqvist A., Renström E.A., Gustafsson Sendén M. 2018: *Reducing a male bias in language? Establishing the efficiency of three different gender-fair language strategies*, „Sex Roles", vol. 81, no. 1–2, s. 109–117.

Ludność 2021: *Ludność w rejonach statystycznych i obwodach spisowych według płci, ekonomicznych i 10-letnich grup wieku*, Narodowy Spis Powszechny Ludności i Mieszkań 2021, Główny Urząd Statystyczny (online: Ludność w rejonach statystycznych i obwodach spisowych według płci, ekonomicznych i 10-letnich grup wieku. Tablice w formacie XLSX, dostęp: 18 lipca 2024).

Łaziński M. 2006: *O panach i paniach. Polskie rzeczowniki tytularne i ich asymetria rodzajowo-płciowa*, Wydawnictwo Naukowe PWN, Warszawa.

Łaziński M. 2023: *Feminatywy oraz inne spory o słowa. Próba diagnozy i propozycje pozytywne*. „Socjolingwistyka" XXXVII, nr 1, s. 345–368.

Małocha-Krupa A. 2018: *Feminatywum w uwikłaniach językowo-kulturowych*, Oficyna Wydawnicza Atut, Wrocław.

Müller-Spitzer C. 2022: *Zumutung, Herausforderung, Notwendigkeit? Zum Stand der Forschung zu geschlechtergerechter Sprache*, „Aus Politik und Zeitgeschichte", Themenheft: *Geschlechtergerechte Sprache*, t. 72, nr 5–7, s. 23–29.

Nowakowska M., Radzikowska K. 2022: *Język inkluzywny. Jak mówić, żeby włączać*, Stowarzyszenie Trenerów Organizacji Pozarządowych, Warszawa (online: https://stowarzyszeniestop.pl/stop/wp-content/uploads/2023/04/Jak-mowic-zeby-wlaczac-jezyk-inkluzywny.-Nowakowska-Radzikowska.pdf, dostęp: 22 lipca 2024).

Ogrodniczuk M., Osenova P., Erjavec T., Fišer D., Ljubešić N., Çöltekin Ç., Kopp M., Meden K. 2022: *ParlaMint II: The Show Must Go On*, [w:] *Proceedings of the LREC 2022 ParlaCLARIN III Workshop on Creating, Enriching and Using Parliamentary Corpora*, European Language Resources Association (ELRA), s. 1–6 (online: http://www.lrec-conf.org/proceedings/lrec2022/workshops/ParlaCLARINIII/pdf/2022.parlaclariniii-1.1.pdf, dostęp: 22 lipca 2024).

Osińska-Szymańska M., Preuhs K. 2022: *Polityczność feminatywów*, „Studia Polityczne", t. 50, nr 3, s. 11–28.

RdR 2021: *Die Entwicklung und Bewertung des Themas „Geschlechtergerechte Schreibung" in der Beobachtung des Schreibgebrauchs 2018–2020*, Rat für deutsche Rechtschreibung, Mannheim (online: https://www.rechtschreibrat.com/DOX/rfdr_PM_2021-03-26_Anlage1_Geschlechtergerechte_Schreibung_seit_2018.pdf, dostęp: 30 sierpnia 2024).

Saloni Z., Woliński M., Wołosz R., Gruszczyński W., Skowrońska D. 2012: *Słownik gramatyczny języka polskiego. Podstawy teoretyczne. Instrukcja użytkowania*, Wiedza Powszechna, Warszawa.

Sczesny S., Formanowicz M., Moser F. 2016: *Can gender-fair language reduce gender stereotyping and discrimination?*, „Frontiers in Psychology", vol. 7, no. 25 (online: https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2016.00025/full, dostęp: 30 sierpnia 2024).

Stahlberg D., Sczesny S. 2001: *Effekte des generischen Maskulinums und alternativer Sprachformen auf den gedanklichen Einbezug von Frauen*, „Psychologische Rundschau", t. 52, nr 3, s. 131–140.

Standardy 2024: *Standardy równego traktowania w Urzędzie m.st. Warszawy*, Załącznik do zarządzenia nr 822/2024 Prezydenta m.st. Warszawy z 8 maja 2024 r. (online: https://tvn24.pl/najnowsze/zalacznik-do-zarzadzenia-nr-8222024--prezydenta-mst-warszawy-z-8-maja-2024-r-bi7921621/0822_0805zal.docx.pdf, dostęp: 30 sierpnia 2024).

Swales J. 1990: *Genre analysis. English in academic and research settings*, Cambridge University Press, New York.

Szpyra-Kozłowska J. 2019: Premiera, premierka *czy* pani premier*? Nowe feminatywy w badaniu ankietowym*, „Język Polski" XCIX, z. 2, s. 22–40.

Szpyra-Kozłowska J. 2022: *Jak zmniejszyć nierówność płci w polszczyźnie?*, [w:] I. Chmura-Rutkowska, J. Szpyra-Kozłowska, *Nierówność płci w języku. Poradnik dla nauczycielek i nauczycieli szkół m.st. Warszawy*, Warszawskie Centrum Innowacji Edukacyjno-Społecznych i Szkoleń, Urząd m.st. Warszawy, Biuro Edukacji, Warszawa, s. 47–70.

Szpyra-Kozłowska J. 2024: Człowiek, polityk, podatnik. *Czy generyczne maskulinum ma charakter w pełni inkluzywny?*, „Język Polski" CIV, z. 1, s. 17–31.

Szpyra-Kozłowska J., Chmura-Rutkowska I. 2022: *Słowniczek*, [w:] I. Chmura-Rutkowska, J. Szpyra-Kozłowska, *Nierówność płci w języku. Poradnik dla nauczycielek i nauczycieli szkół m.st. Warszawy*, Warszawskie Centrum Innowacji Edukacyjno-Społecznych i Szkoleń, Urząd m.st. Warszawy, Biuro Edukacji, Warszawa, s. 10–11.

Tomaszewska A., Jamka A. 2024: *Język wrażliwy na płeć w sejmie i senacie: badanie korpusowe strategii feminizacji i neutralizacji w polskim dyskursie parlamentarnym*, „Język Polski" CIV, z. 4, s. 35–48.

W3Techs 2024: *Usage statistics of content languages for websites*, Web Technology Surveys (online: https://w3techs.com/technologies/overview/content_language, dostęp: 30 sierpnia 2024).

Słownik symboli

⚥ – feminizacja: współwystępowanie maskulatywów i feminatywów z ukośnikiem (np. *nauczyciele/nauczycielki*) lub spójnikiem (np. *studentki i studenci*)

*⚥ – feminizacja: skrócone formy równościowe z asteryskiem inkluzywnym (np. *nauczyciel*e/ki*)

♀ – feminizacja: adaptacja feminatywów (np. *prezydentka Zdanowska*)

♀ – neutralizacja (np. *grono pedagogiczne*, *osoba ekspercka*)

Summary

The concept of inclusive forms featuring a gender star

Keywords: inclusive language, feminisation, gender star, large language models.

The article presents a new concept of gender-inclusive notation using a gender star in Polish, e.g. *leka*rz/rka* encompasses both *lekarz* (a male doctor) and *lekarka* (a female doctor). The proposed solution fully adheres to the equality principles, without privileging either the masculine or feminine forms, and its structure allows for adding a suffix to account for non-binary individuals. In developing the concept of the inclusive form with gender star, we pay particular attention to the linguistic economy, the form's consistency with the grammatical system of Polish, and its adaptation in natural language processing, which is crucial for quantitative text analysis. The study includes an analysis of various text genres in order to adapt appropriate inclusive strategies, and an experiment in which texts are automatically converted into their inclusive versions using appropriately prompted large language models. The results show that models are not yet fully effective in gender-inclusive text generation, i.e. it is still a great challenge for them to generate adequate inclusive forms consistent with the genre of the processed text.