



Hélène Włodarczyk

 <https://orcid.org/0000-0003-0277-0490>

Sorbonne-Université
Paris, France

The Polish Language Is Not a Male-Chauvinistic Language

(A Computer-Aided Research on the Polish Gender
Category Presented in a Nutshell)

Język polski nie jest męskoszowinistyczny
(krótka informacja o wspomaganym komputerowo badaniu
nad rodzajem w języku polskim)

Abstract:

I took up the topic of grammatical gender in Polish due to the fact that the theories of 5, 7 and even 9 genders (including the so-called personal-masculine gender in plural) accepted in Poland since the second half of the 20th century give rise to unnecessary difficulties for foreigners attempting at correctly building Polish sentences such as *Matka i dziecko nie mogli się sobą nacieszyć*. ('Mother and child could not stop enjoying each other'). The research I conducted as part of interactive linguistics (using the Prolog artificial intelligence language and the SEMANA platform) showed that the theory distinguishing the category of gender from the category of animacy (as in comparative Slavic linguistics) is more economical and gives better results (than the theory of 9 genders) both in the analysis and production of utterances.

Additionally, due to the "personal-masculine" gender introduced into Polish grammar, there arose a debate on the possible "male-chauvinistic" character of the Polish language, and consequently of the entire Polish culture. In response to this, I argue that one should not too hastily identify grammatical categories with cognitive categories, especially since in all European languages I know, it is the masculine gender which is the unmarked gender (-feminine). It follows from this that the masculine gender has two meanings, one broad and more abstract (e.g., *man* as a human being in general) and one narrow indicating only men (e.g., *man* as a male person). Recently, however, the non-marked meaning of the masculine gender has been undermined by the so-called inclusive speech. In Polish, in the plural, the so-called masculine-personal form indicates either a group of men or a group of entities among which is at least one human being regardless of gender (in the utterance quoted above, in which the verb *nie mogli* ('could not') exhibits the so-called personal-masculine form, the feminine noun *matka* ('mother') and the neuter noun *dziecko* ('child') constitute a collective subject that occurs in grammatical agreement with the verb form in the masculine gender, acting as an exponent of the category of humanity (not "masculinity"). Thus, the so-called personal-masculine form is used with the unmarked gender meaning of a personal (or human) form (+*human*, -*feminine*). The research I conducted shows that in contemporary Polish the category of gender has the following

meanings: masculine, feminine, and neuter, and the category of animacy: inanimate (-animate), human animate (+animate, +human), and non-human or animal animate (+animate, -human).

Keywords: gender in Polish, interactive linguistics, linguistic view of the world

Abstrakt: W niniejszym artykule podejmuję temat rodzaju w języku polskim ze względu na fakt, że upowszechniane w Polsce od drugiej połowy XX wieku teorie pięciu, siedmiu, a nawet dziewięciu rodzajów (w tym tzw. rodzaju męskoosobowego) sprawiają cudzoziemcom niepotrzebne trudności w poprawnym budowaniu polskich zdań typu: *Matka i dziecko nie mogli się sobą nacieszyć*. Badania, jakie prowadziłam w ramach lingwistyki interaktywnej (za pomocą języka sztucznej inteligencji Prolog oraz platformy SEMANA), wykazały, że teoria odróżniająca kategorię rodzaju od kategorii żywotności (jak w porównawczej lingwistyce słowiańskiej) jest bardziej ekonomiczna i daje lepsze wyniki (niż teoria dziewięciu rodzajów) tak w analizie, jak i w produkcji wypowiedzi.

Ponadto wprowadzenie do gramatyki opisowej języka polskiego rodzaju męskoosobowego zapoczątkowało debatę nad rzekomo „męskozwinińskim” charakterem polszczyzny, a co za tym idzie – także całej kultury polskiej. W odpowiedzi na to twierzę, że nie można zbyt pochopnie utożsamiać kategorii gramatycznych z kategoriami kognitywnymi, zwłaszcza że w znanych mi językach europejskich odróżniających rodzaj męski od żeńskiego rodzajem nienacechowanym jest rodzaj męski (-żeński). Z tego wynika, że rodzaj męski ma dwa znaczenia – szerokie i bardziej abstrakcyjne (np. człowiek) albo wąskie, wskazujące tylko na mężczyzn (np. mężczyzna). Ostatnio jednak charakter nienacechowany rodzaju męskiego został podważony przez tzw. mowę inkluzywną. W języku polskim w liczbie mnogiej tzw. forma męskoosobowa wskazuje albo na grupę mężczyzn, albo na grupę istot, wśród których znajduje się przynajmniej jedna istota ludzka niezależnie od płci – jak w cytowanym przykładzie wypowiedzi, w której czasownik *nie mogli* ma tzw. formę męskoosobową, a *matka* (rzeczownik rodzaju żeńskiego) i *dziecko* (rzeczownik rodzaju nijakiego) stanowią zbiorowy podmiot, który występuje w zgodzie z formą czasownika w rodzaju męskim jako wykładnik kategorii ludzkości (a nie „męskoosobowości”). Zatem tzw. forma męskoosobowa jest użyta z nienacechowanym znaczeniem rodzaju (-feminine) oraz z nacechowanym znaczeniem formy osobowej, czy inaczej ludzkiej (+human). Z przedstawionych badań wynika, że w języku polskim kategoria rodzaju ma obecnie następujące znaczenia: męski, żeński i nijaki, a kategoria żywotności: nieżywotny (-animate), żywotny ludzki (+animate, +human) i żywotny nie-ludzki (+animate, -human).

Słowa kluczowe: rodzaj w języku polskim, lingwistyka interaktywna, językowy obraz świata

For more than 50 years, both native and foreign students have been taught that Polish is an exotic language having from 5 (Mańczak, 1956) up to 9 genders (Saloni, 1976; Przepiórkowski, 2003). As an amazing consequence, in recent years, the development of so-called *cognitive linguistics*¹ has given rise to the idea that Polish is supposed to be a male-chauvinistic language due to grammatical descriptions using the term “masculine-human” gender in plural number (Pol. *rodzaj męskoosobowy*), which apparently leads to the reduction of the class of human beings in Polish grammar to male only.

¹ I mean here by “cognitive” not the interdisciplinary research between linguistics and neuroscientists, but the approach proposed in linguistics by Ronald Langacker.

The computer-aided research carried out at CELTA (Centre de Linguistique Théorique et Appliquée) research centre with the SEMANA platform² led us to call into question the contemporary theory of gender in Polish grammar (Sauvet & Włodarczyk A. & H., 2008; Włodarczyk H., 2008, 2009, 2012, 2018). We carried out the attempt at describing Polish gender morphology using the often accepted theory of 5 (up to 9) gender values. In view of the inconsistent results we obtained³, we had to put forward another theory that sets apart the *gender* category (with its three values: *masculine*, *feminine*, and *neuter*) from the *animacy* category (*inanimate*, *animate-non-human*, and *animate-human*). Therefore, we ruled out the so-called male-personal gender value (*rodzaj męskoosobowy*). Although Polish morphology uses amalgamated forms, it seems justified to separate the said categories just as we separate the *case*, *gender*, and *number* which are expressed by a single morpheme, for instance, the *-a* with the three meanings: nominative case, feminine gender, and singular number in the noun: *dziewczyn-a* ('girl').

Moreover, the syntax of subject-verb agreement in utterances having more than one subject noun phrase (with a multiple subject, cf. Karolak, 1984) brings evidence of the relevance of the \pm human semantic feature in Polish morphology and syntax. At the same time, it can be demonstrated that the characterization of the form *-li* of verbs (3rd person plural in the past tense) as "masculine-human" is misleading, since multiple subjects which do not contain any "masculine-human" noun, nonetheless agree with this form.

² The SEMANA software was designed for CELTA by André Włodarczyk (2007) and Georges Sauvet (2008). It has been presented in A. Włodarczyk and H. Włodarczyk (2019), and on the website <http://celta.paris-sorbonne.fr/anasem/Semana-Corner/index.html>

³ Using the SEMANA platform database builder we could gather and analyze as many samples of expressions as necessary. Indeed, SEMANA makes it possible to easily gather in the database hundreds of examples randomly found in corpora. The platform database builder does not only help to add or remove attributes but its statistical algorithms reveal when the database is exhaustive enough to account for all the possible language uses and attributes configurations.

- (1) *Matka i dziecko patrzyli na siebie.*
 Mother and child were looking at each other
 Nom Fem +Hum Sing Nom Neu +Hum Sing Past 3rd Pers Plur +Hum
 ‘The mother and her child were looking at each other.’
- (2) *Dziecko i pies patrzyli na siebie.*
 Child and dog were looking at each other
 Nom Neu +Hum Sing Nom Masc -Hum Sing Past 3rd Pers Plur +Hum
 ‘The child and the dog were looking at each other.’

From the observation of such utterances as (1) and (2) above I proposed to consider the form *-li* as +Hum without any gender specification and the opposite *-ły* as -Hum as well without any gender specification. When the subject phrase is composed of more than one noun, it is sufficient that one of these nouns refers to a human being (of any sex) to trigger the use of *-li*. As a matter of fact, verbs in past plural have two possible endings: *-li* or *-ły*. But the difficulty for linguists to describe these forms results from the fact that each of them has two meanings depending on the gender and animacy category of the subject noun with which the verb is used. Thanks to our platform SEMANA designed for interactive research in linguistics (Włodarczyk H., 2009), we built a database collating examples of uses of the forms *-li* and *-ły*. As a result of the analysis of the database with data-mining tools, we could represent the oppositions of these forms by the following grid (ENT means *entity*):

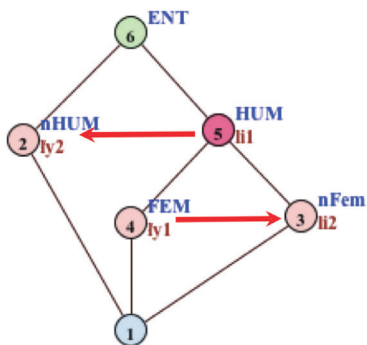


Figure 1

We put forward the hypothesis that two privative oppositions (4 different meanings) are expressed by only two forms:

- (a) on the one hand, we have the opposition between: $-li_1$ (+hum) and $-ly_2$ (-hum);
- (b) on the other hand, the opposition between $-ly_1$ (+fem) and $-li_2$ (-fem).

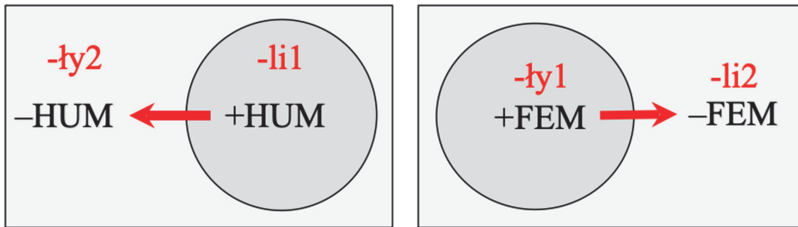


Figure 2

Thus, one should recognize that in contemporary Polish, nouns designing women are not treated as nouns designing non-human entities (which is currently interpreted as the result of a “male-chauvinistic” worldview) but rather as a special “marked” or distinguished class among human entities.

In conclusion, it appears that as far as the classification of nouns is concerned, Polish replaced in the plural the *animate/inanimate* opposition – common to other Slavic languages (e.g., Russian) where it is present both in singular and in plural forms of masculine nouns – with a *human/non-human* opposition as in contemporary English. That would rather reflect an anthropocentric *Weltanschauung* also very common in most European languages. However, one should not consider that the grammatical categories of a given language directly reflect the image of the world in the mind of people who speak that language. Although in prehistory, such close relations between language and worldview might have been possible, however, in historical times, grammatical categories are purely morphological devices playing an important role in syntax. Thus, although it is widespread and fashionable in present times, one should not confuse *gender* in language and *sex* in society. The fact that a *table* is feminine in French and masculine in Polish has

no philosophical consequences. It would be equally grotesque to pretend that languages like French and Polish – in which nouns designating inanimate objects have gender forms – are “animist” languages.

Bibliography

- Karolak S., 1984, *Składnia wyrażen predykatywnych*, in: *Gramatyka współczesnego języka polskiego*, ed. S. Urbańczyk, vol. 1, *Składnia*, ed. Z. Topolińska, Wydawnictwo Naukowe PWN, Warszawa 1984, pp. 11–212.
- Mańczak W., 1956, *Ile jest rodzajów w języku polskim?*, „*Język Polski*”, R. XXXIV, pp. 116–121.
- Przepiórkowski A., 2003, *A Hierarchy of Polish Genders*, in: *Generative Linguistics in Poland: Morphosyntactic Investigations*, eds. P. Bański, A. Przepiórkowski, Instytut Podstaw Informatyki Polskiej Akademii Nauk, Warszawa, pp. 109–122.
- Saloni Z., 1976, *Kategoria rodzaju we współczesnym języku polskim*, in: *Kategorie gramatyczne grup imiennych we współczesnym języku polskim*, eds. S. Urbańczyk, J. Basara, K. Rymut, Zakład Narodowy im. Ossolińskich, Wrocław, pp. 41–75.
- Sauvet G., Włodarczyk A., Włodarczyk H., 2008, *Morphological Data Exploration – Using the Semana Platform (Feature Granularity Problem in the Definition of Polish Gender)*, CASK Sorbonne 2008, Language Data Mining International Conference, June 13th–14th, Université Paris-Sorbonne (Paris 4), <http://celta.paris-sorbonne.fr/anasem/papers/miscelanea/PolishGender.pps>.
- Włodarczyk A., Włodarczyk H., 2019, *The Interactive Method for Language Science and Some Salient Results*, „*Zagadnienia Naukoznawstwa*”, vol. 3 (221), Komitet Naukoznawstwa PAN, Warszawa, pp. 73–92. <http://celta.paris-sorbonne.fr/anasem/papers/miscelanea/PolishGender.pps>
- Włodarczyk H., 2009, *Lingwistyka na polonistyce krajowej i zagranicznej w dobie filozofii informatyczno-logicznej*, „*LingVaria*”, no. 1 (7), pp. 65–79.
- Włodarczyk H., 2012, *Les catégories du genre et de l'animation dans la grammaire polonaise*, in: *L'Enseignement du polonais en France*, eds. K. Siatkowska-Callebat, A. Synoradzka-Demadre, IIIe Assises, Centre Scientifique de l'Académie Polonaise des Sciences, Varsovie–Paris, pp. 25–38.

Włodarczyk H., 2018, *O Potrzebie wiedzy o języku polskim kompatybilnej z wiedzą o innych językach narodowych*, in: *Polonistyka na początku XXI wieku. Diagnozy, koncepcje, perspektywy*, vol. 5, *W kręgu (glotto)dydaktyki*, eds. A. Achtełik, K. Graboń, Wydawnictwo Uniwersytetu Śląskiego, Katowice, pp. 205–218.

HÉLÈNE WŁODARCZYK – emeritus professor at the Department of Slavic Studies, Paris-Sorbonne University (currently Sorbonne-Université), Paris, France / emerytowany profesor doktor habilitowany, Wydział Sławistyki Uniwersytetu Paris-Sorbonne (obecnie Sorbonne-Université), Paryż, Francja.

Hélène Włodarczyk was the founder and director of the Centre de Linguistique Théorique et Appliquée (CELTA), Université Paris-Sorbonne – Paris IV (2000–2014). At CELTA, linguistic research was conducted on the three families of modern Indo-European languages (Romance, Germanic, and Slavic). Her main scientific interests include: Polish, Slavic, general theoretical and interactive (computational) linguistics, and Polish literature of the 20th century. Among her major publications are: *The Interactive Method for Language Science and Some Salient Results*, „Zagadnienia Naukoznawstwa (Problems of Epistemology)”, Komitet Naukoznawstwa PAN, Warszawa 2022, pp. 73–92; *Qu'est-ce au juste que la prédication ?*, „Bulletin de la Société de Linguistique de Paris”, t. CXIV (2019), fasc. 1, pp. 1–54 (both co-authored with André Włodarczyk).

Założycielka i dyrektorka Centre de Linguistique Théorique et Appliquée (CELTA), Université Paris-Sorbonne – Paris IV (2000–2014). Prowadzono tam badania językowe nad trzema rodzinami współczesnych języków indoeuropejskich (romańskimi, germańskimi i słowiańskimi). Zainteresowania naukowe badaczki to: polonistyka, sławistyka i językoznawstwo ogólne. Najważniejsze publikacje: *Lingwistyka na polonistyce krajowej i zagranicznej w dobie filozofii informatyczno-logicznej* (Kraków 2010), *O potrzebie wiedzy o języku polskim kompatybilnej z wiedzą o innych językach narodowych* (Katowice 2016), *Mnie czy mi? O użyciu zaimka pierwszej osoby w celowniku* („Poradnik Językowy” 9/2018).

E-mail: helene.wlodarczyk@gmail.com