

## Opis morfo-syntaktyczny liczebników zbiorowych oraz zaimków osobowych na potrzeby konwersji danych słownika Morfologik do tagsetu KIPI

26.01.2011

Słownik Morfologika [Miłkowski 2010] udostępniany jest na otwartej licencji: GNU LGPL bądź Creative Commons ShareAlike (użytkownik może wybrać jedną z nich). Daje to mu istotną przewagę nad Morfeuszem. Z drugiej strony, tagset Korpusu IPI PAN [Przepiórkowski 2004] stał się standardem *de facto* w językoznawstwie komputerowym. Podjęliśmy się odtworzyć tagi Korpusu IPI PAN (odtąd: KIPI) na podstawie tagów Morfologika oraz prac językoznawczych. O ile konwersję wielu klas gramatycznych przeprowadziliśmy pół-automatycznie, kilka klas wymagało ręcznej analizy. Szczególną uwagę poświęciliśmy zaimkom osobowym i liczebnikom — ponieważ słownik Morfologika zawiera w obrębie tych klas znaczną liczbę błędów gramatycznych. Dane otrzymane w wyniku konwersji zostały włączone do systemu analizy morfologicznej Maca [Radziszewski, Śniatowski 2011] — stąd też w dalszej części tego dokumentu przekonwertowane dane Morfologika skrótowo określamy mianem „Maca”.

### 1. RODZAJE

Tagset KIPI posługuje się zestawem rodzajów ustalonym dla języka polskiego przez Salonię i Świdzińskiego.

#### 1.1. Rodzaje n1 i n2

Podkategorie rodzaju nijakiego n1 oraz n2 są atrybutami rzeczowników, które zależą od łączliwości rzeczownika z formami liczebników (głównych bądź zbiorowych):

Klasa n1, typu *dzieci*, łączy się wyłącznie z liczebnikami zbiorowymi (*siedmioro, ośmioro* etc.), podczas gdy klasa n2, typu *kół*, łączy się z liczebnikami głównymi (*pięć, osiem*, por. w google np. formę *\*pięcioro kół*, 0 wystąpień, *\*czworo kół* tylko 3 wystąpienia wobec 150 tys. użyć wyrażenia *cztery koła*) [por. Saloni, Świdziński, 1998: 197].

#### 1.2. Rodzaje p1, p2 i p3

Rzeczowniki plurale tantum łączą się w składni zgody wyłącznie z (nijakimi) przymiotnikami oraz czasownikami w liczbie mnogiej [Saloni, Świdziński, 1998: 100, 170, 196, 198], niektóre z nich mogą się łączyć z liczebnikami zbiorowymi [ibidem: 196-7]. Wyróżnia się wśród nich plurale tantum p1, p2 oraz p3. Rzeczowniki z atrybutem p1, typu *państwo, wujostwo, rodzice*, łączą się z przymiotnikami oraz czasownikami w rodzaju m1 (rodzaj męskoosobowy), rzeczowniki te łączą się wyłącznie z liczebnikami zbiorowymi [Saloni, Świdziński 1998: 197]. Pozostałe rzeczowniki zbiorowe (p2 oraz p3) nie łączą się w składni zgody z przymiotnikami i czasownikami w rodzaju m1. Klasa p2, typu *skrzypce*, dopuszcza współwystępowanie z liczebnikami zbiorowymi (i tylko takimi), zaś klasa p3, typu *spodnie, okulary, perfumy*, nie dopuszcza łączliwości z żadnymi liczebnikami, „jeśli zachodzi konieczność komunikowania wielości nazywanych nimi przedmiotów – są na ogół używane z pomocniczymi wyrazami, przede wszystkim *para*, np. *pięć par spodni, pięć par okularów*” [Saloni, Świdziński 1998: 198].

Podsumowujemy to w **tabelce**:

	atrybut				
Łączliwość z:	n1	n2	p1	p2	p3
Czasownikiem	(n)*		plur (m1)*	plur (m2 m3 f n)*	
Przymiotnikiem	n		plur : m1	plur : m2 m3 f n	
Liczebnikiem	zbiorowym	głównym	zbiorowym	zbiorowym	-----

\*formy rodzajowe czasownika

rodzaje n1, n2, p1, p2, p3 są atrybutami wyłącznie rzeczowników, czasowniki i przymiotniki nie znają tych rozróżnień.

Rozróżnienie między n1 i n2 oraz p1, p2 i p3 nie występuje w rzeczywistym korpusie IPI PAN (korpus.pl). Wygląda na to, że na etapie planowania istniały dwa warianty tagsetu — np. w pracy [Woliński 2004] rozróżnienie to udokumentowano, natomiast w kilku innych źródłach — w tym w książce [Przepiórkowski 2004] — rozróżnienia tego nie ma. My zdecydowaliśmy się na użycie tagsetu rzeczywistego korpusu, zastępując rodzaje w słowniku Morfologika w następujący sposób:

- n1, n2 → n,
- p1 → m1,
- p2, p3 → n.

## 2. ZAIMEK OSOBOWY

### 2.1. ZAIMKI 1 i 2 OS.

Zaimek *ja* i *ty* są uznawane przez Saloniego za rzeczowniki, są to singularia tantum [Saloni, Świdziński 1998: 175]. Oba zaimki wykazują łączliwość ze wszystkimi rodzajami (tj. m1, m2, m3 oraz n i f) [Saloni, Świdziński 1998: 176, 182]. „W podobny sposób rozwiązany został problem uzgodnienia rodzaju dla zaimków osobowych *ja*, *ty*, *my* i *wy*. Nie mają one w słowniku zdefiniowanej dokładnej wartości rodzaju. Odpowiedni wybór dokonywany jest w momencie uzgadniania wymagań formy czasownikowej.” [FOJP 2001: 215]. Zaimki MY i WY łączą się z następującymi rodzajami: p1 oraz p-1 [Saloni, Świdziński 1998: 177]. Ponieważ zrezygnowaliśmy z kategorii p1 (= p1) oraz p-1 (= p2 i p3), dlatego — podobnie jak w przypadku zaimków JA i TY — połączymy zaimki MY i WY ze wszystkimi rodzajami (m1, m2, m3, f i n), w myśl zasady, że p1 odpowiada m1, a p2 oraz p3 odpowiadają m2, m3, f i n (w Macy zamieniamy to na n).

Oto zestaw kategorii zaimkowych:

- liczba: pojedyncza (*ja*, *ty*), mnoga (*my*, *wy*)
- przypadek
- rodzaj (dla wszystkich m1, m2, m3, f i n)
- osoba (*ja*, *my* – pri, *ty*, *wy* – sec)
- akcentowość (relewantna tylko dla tych par, w przypadku których występuje oboczność formy akcentowanej i nieakcentowanej)
- poprzyimkowość (charakterystyczna tylko dla zaimka trzeciej osoby)

### Akcentowość

„Kategoria akcentowości została wprowadzona dla odróżnienia form zaimka osobowego występujących w zdaniu na pozycji akcentowanej (np. *jego*) od form występujących na pozycji nieakcentowanej (np. *go*). Poprzyimkowość różnicuje te formy zaimka *on*, które mogą wystąpić wyłącznie po przyimku (*przez niego*, *do-ń*), i te, które występują w innych kontekstach (*jego*, *go*). (Kategorie te po raz pierwszy zostały wprowadzone przez Saloniego (1981)).” [Woliński 2004: 46].

### 2.1.1. JA

Oto paradygmat odmiany zaimka *ja* [za: GWJP-M1 1998: 337; *Słownik języka polskiego*, pod red. M. Szymczaka, Warszawa 1982, t. I: 815]

M. ja

D. mnie, mię

C. mnie, mi

B. mnie, mię

N. mną

Ms. mnie

„Formy D.-C.-B. *mnie* i D.-B. *mię*, C. *mi* oraz D.-B. *ciebie* i *cię*, C. *tobie* i *ci* pozostają w stosunku do siebie w dystrybucji uzupełniającej jako formy ortotoniczne (akcentowane) i enklityczne (nieakcentowane). Traktujemy tu enklityki *mię*, *mi*, *cię*, *ci* jako warianty pozycyjne odpowiednich form ortotonicznych. Por. Inaczej Saloni 1980, wprowadzający dla zaimków osobowych kategorię fleksyjną opartą na przeciwstawieniu pozycji mocnej (formy ortotoniczne) i pozycji słabej (enklityki) form fleksyjnych tych leksemów” [GWJP-M1 1998: 337].

JA	Morfologik – zestaw tagów
ja	ppron12:sg:nom.voc:f.m.n.n1.n2:pri
mnie	ppron12:sg:acc.dat.gen:f.m.n.n1.n2:pri:akc ppron12:sg:loc:f.m.n.n1.n2:pri
mi	ppron12:sg:dat:f.m.n.n1.n2:pri:nakc
mną	ppron12:sg:inst:f.m.n.n1.n2:pri
mię	---

Uwagi:

(1) Morfologik nie podaje kategorii akc/nakc w przypadkach, gdy są one nerelevantne (tj. w M, N, Ms), jest to podejście słuszne.

(2) Morfologik nie ma formy *mię*, ale akcentowość, która odróżnia *mnie* od *mię*, jest przez Morfologika utrzymana w tagach formy *mnie*.

(3) W B. i D. akcentowanej formie *mnie* odpowiada nieakcentowane, przestarzałe *mię*, w C. *mnie* odpowiada nieakcentowane *mi*. Tę formę *mię*, niewystępującą w Morfologiku, przez analogię do formy *mi*, zaopatrzylibyśmy w następujące tagi:

- pron12 (część mowy)
- sg (singulare tantum [Saloni, Świdziński 1998: 175]),
- m1, m2, m3, f, n (kookurencja ze wszystkimi rodzajami),
- pri (zaimek pierwszej osoby, współwystępuje z czasownikami w 1 os. l.poj.),
- acc.gen (forma biernikowa i dopełniaczowa [WSPP I; Morfologia I]),
- nakc (forma nieakcentowana [WSPP I: *ja*]).

Pełny zestaw tagów dla formy *mię* jest następujący:

mię	ppron12:sg:acc.gen:m1.m2.m3.f.n:pri:nakc
-----	--

(4) Błędny jest Morfologikowy tag voc w przypadku formy *ja* (zaimki osobowe nie mają wołacza)

	<a href="#">MACA</a>
ja	<a href="#">ppron12:sg:nom:f.m1.m2.m3.n:pri</a>
mnie	<a href="#">ppron12:sg:acc.dat.gen:f.m1.m2.m3.n:pri:akc</a> <a href="#">ppron12:sg:loc:f.m1.m2.m3.n:pri</a>
mną	<a href="#">ppron12:sg:dat:f.m1.m2.m3.n:pri:nakc</a>
mi	<a href="#">ppron12:sg:inst:f.m1.m2.m3.n:pri</a>
mię	<a href="#">ppron12:sg:acc.gen:f.m1.m2.m3.n:pri:nakc</a>

### 2.1.2. TY

Paradygmat odmiany [GWJP-M1 1998: 337; WSPP II: 1210]:

M. ty

D. ciebie, cię

C. tobie, ci

B. ciebie, cię

N. tobą

Ms. tobie

„Formy pełne (ciebie, tobie) mają samodzielny akcent; mogą być używane na początku i na końcu zdania (akcent wyrazowy pokrywa się wtedy z akcentem zdaniowym) oraz po przyimkach, natomiast nieakcentowane formy skrócone (cię, ci) w takich właśnie pozycjach nie występują; najczęściej używa się ich po czasowniku (...).” [WSPP II: 1210]

	Morfologik
ty	ppron12:sg:nom:f.m.n.n1.n2:sec
ciebie	ppron12:sg:acc.gen:f.m.n.n1.n2:sec:akc
tobie	ppron12:sg:dat:f.m.n.n1.n2:sec:akc ppron12:sg:loc:f.m.n.n1.n2:sec
tobą	ppron12:sg:inst:f.m.n.n1.n2:sec
ci	ppron12:sg:dat:f.m.n.n1.n2:sec:nakc
cię	---

Uwagi:

(1) Morfologik nie ma formy *cię*. Nadamy jej następujące tagi:

- pron12 (część mowy)
- sg (singulare tantum)
- gen.acc (WSPP, USJP, GWJP-M1)
- m1, m2, m3, f, n (wszystkie rodzaje)
- sec (współwystępowanie z czasownikami w 2 osobie)
- nakc (GWJP-M1)

Ostatecznie uzyskujemy taki zestaw tagów dla formy *cię*:

cię	ppron12:sg:gen.acc:m1.m2.m3.f.n:sec:nakc
-----	--

(2) Zwróćmy uwagę, że loc *tobie* nie ma tagu akc, jest to całkowicie poprawne, w miejscowniku formie *tobie* nie przeciwstawia się żadna forma nieakcentowana.

Zestaw tagów Macy przybiera następującą postać:

TY	MACA – zestaw tagów
ty	ppron12:sg:nom:m1.m2.m3.f.n:sec
ciebie	ppron12:sg:acc.gen:m1.m2.m3.f.n:sec:akc
tobie	ppron12:sg:dat:m1.m2.m3.f.n:sec:akc ppron12:sg:loc:m1.m2.m3.f.n:sec
tobą	ppron12:sg:inst:m1.m2.m3.f.n:sec
ci	ppron12:sg:dat:m1.m2.m3.f.n:sec:nakc
cię	ppron12:sg:gen.acc:m1.m2.m3.f.n:sec:nakc

### 2.1.3. MY

Zestaw form [GWJP-M1 1998: 338]:

M. my  
D. nas  
C. nam  
B. nas  
N. nami  
Ms. nas

W paradygmacie odmiany zaimka 1 os. l.mn. nie ma specjalnych form nieakcentowanych, dlatego kategorie akc/nakc są nerelevantne dla zaimka MY (analogicznie dla WY): „Enklityczne formy D., C. I B. Zaimków MY i WY są równokształtne z odpowiednimi formami ortotonicznymi; por. np. *Wie to od nas, Wam to powiedział* (formy ortotoniczne) i *Nie słyszy nas, Chce nam to dać* (formy enklityczne)” [GWJP-M1 1998: 338].

	Morfologik
my	ppron12:pl:nom:pri
nas	ppron12:pl:acc.gen.loc:pri
nam	ppron12:pl:dat:pri
nami	ppron12:pl:inst:pri

Po zamianie rodzaju p1 na m1.m2.m3.f.n otrzymujemy właściwe tagi MAC-y:

MY	MACA
my	ppron12:pl:nom:m1.m2.m3.f.n:pri
nas	ppron12:pl:acc.gen.loc:m1.m2.m3.f.n:pri
nam	ppron12:pl:dat:m1.m2.m3.f.n:pri
nami	ppron12:pl:inst:m1.m2.m3.f.n:pri

Uwagi:

(1) Zaimkom osobowym pierwszej i drugiej osoby liczby mnogiej powinny być nadane wszystkie tagi rodzajowe m1.m2.m3.f.n – przez analogię do zaimków osobowych JA i TY.

(2) „Tak zwane zaimki osobowe zostały opisane za pomocą dwóch klas leksemów: zaimków niertzecioosobowych i trzecioosobowych. Pierwsza z klas zawiera cztery leksemy: *ja, ty, my* i *wy*. Przyjęto, że leksemy tej klasy są odmienne przez przypadek, rodzaj (mimo że zachodzi pełna neutralizacja tej kategorii) i akcentowość.” [Woliński 2004]. - Zwróćmy uwagę, że w przypadku zaimków MY oraz WY kategorię akcentowości/nieakcentowości – jako nerelevantne – zostały pominięte.

### 2.1.4. WY

Paradygmat fleksyjny zaimka WY [GWJP-M1 1998: 338]:

M. wy  
D. was  
C. wam  
B. was  
N. wami  
Ms. was

	Morfologik
wy	ppron12:pl:nom:sec
was	ppron12:pl:acc.gen.loc:sec
wam	ppron12:pl:dat:sec
wami	ppron12:pl:inst:sec

Oto tagi MAC-y

WY	MACA
wy	ppron12:pl:nom:m1.m2.m3.f.n:sec
was	ppron12:pl:acc.gen.loc:m1.m2.m3.f.n:sec
wam	ppron12:pl:dat:m1.m2.m3.f.n:sec
wami	ppron12:pl:inst:m1.m2.m3.f.n:sec

## 2.2. ON

W przypadku zaimka trzeciej osoby dochodzi istotna kwestia rodzaju: zaimek ON odmienia się przez rodzaje (zaimki Ppron12 były obojętne na kategorię fleksyjną rodzaju).

Dla zaimka 3. osoby istotne są następujące kategorie gramatyczne [GWJP-M1 1998: 336-337]

- liczba (sg.pl),
- rodzaj (m1.m2.m3.f.n),
- przypadek,
- osoba (ter),

fakultatywne są następujące kategorie (por. [Woliński 2004: 49]):

- akcentowość,
- poprzyimkowość.

„Klasa zaimków trzecioosobowych zawiera tylko jeden leksem o postaci hasłowej ON, odmienny przez przypadek, rodzaj, akcentowość i poprzyimkowość. Leksemy obu tych klas mają przypisaną słownikową wartość liczby i osoby.” [Woliński 2004: 50-1].

PARADYGMAT FLEKSYJNY ZAIMKA ON wygląda następująco [GWJP-M I: 338-9]

l.poj.	m1.m2.m3	n	f
M.	on	ono	ona
D.	jego, niego, go		jej, niej
C.	jemu, niemu, mu		jej, niej
B.	jego, niego, go	je, nie	ją, nią
N.	nim		nią
Ms.	nim		niej
l.mn.	m1	m2.m3.n.f	
M.	oni	one	
D.	ich, nich		
C.	im, nim		
B.	ich, nich	je, nie	
N.	nimi		
Ms.	nich		

GWJP-M I zauważa opozycyjną relację pomiędzy formami akcentowanymi (ortotonicznymi) i nieakcentowanymi (enklitycznymi):

„Podobnie jak u zaimków JA, TY i SIĘ, w D., C. i B. zaimka ON występuje formalna opozycja między ortotonicznym (jego || niego, jemu || niemu) i enklitycznym (go, mu) wariantem formy fleksyjnej. Ograniczona jest ona do D. i C. lp. rodzaju nieżeńskiego oraz B. lp. rodzaju męskiego. Enklityczne i ortotoniczne formy rodzaju żeńskiego oraz ortotoniczne i enklityczne formy liczby

mnoziej są równokształtne; np. Jej jest dobrze – Powiedz jej to, Ich troje – Spotkałem ich.” [GWSJ-M I: 339].

Zauważa również *Żółta gramatyka* opozycję poprzyimkowości – niepoprzyimkowości:

„W dopełniaczu, celowniku i bierniku wszystkich rodzajów i liczb zaimka ON występuje formalna opozycja między wariantem poprzyimkowym i niepoprzyimkowym danej formy fleksyjnej. Wykładnikiem pozycji syntaktycznej zaimka jest w tych formach oboczność (supletywizm) tematu: alternant *j-/ø-* sygnalizuje pozycję niepoprzyimkową (np. *Jemu możesz się poskarżyć, Pokaż im drogę, Słyszę ją*); alternant *ń-*, będący nienacechowanym członem tej opozycji, w formach D., C. i B. sygnalizuje pozycję poprzyimkową, natomiast w formie N. jest neutralnym wykładnikiem znaczenia leksykalnego zaimka ON (por. *przez niego, dla niej, od nich*, ale: *cieszy się nim i idzie z nim*).

Formy: *jego, jej, ich* występujące w wyrażeniach typu *do jego domu, z jej płaszczka, ich matka* itp. nie należą do paradygmatu fleksyjnego zaimka ON, lecz są odrębnymi leksemami – przymiotnikami dzierżawczymi (typu *mój, wasz*) różnią się one swoją nieodmiennością; por. *jego dom, jego domu, jego domowi, ...* Stąd w parach typu *od jego domu – od niego* mamy do czynienia z opozycją dzierżawczego przymiotnika *jego* i poprzyimkowej formy D. lp. rodzaju nieżeńskiego zaimka ON), nie zaś z dwoma wariantami formy poprzyimkowej ON.” [WGJP-M1: 339].

Później odwołują się autorzy WGJP-M1 do prac Saloniego:

- Saloni wprowadza cztery zaimki ON: on-1 (m1), on-2 (m2.m3), ono (n), ona (f)
- przypisuje im cztery kategorie fleksyjne: - przypadku, liczby, akcentowości, przyimkowości
- rodzaj i osoba są dla tych leksemów kategorią klasyfikującą, nie – fleksyjną [ibidem].

Przerabiam powyższą tabelę na tabelę z fakultatywnymi tagami (akc – nakc, praep – npraep):

l.poj.	m1.m2.m3	n	f
M.	on	ono	ona
D.	jego [akc.npraep], niego [akc.praep], go [nakc.npraep]		jej [npraep], niej [praep]
C.	jemu [akc.npraep], niemu [akc.praep], mu [nakc.npraep]		jej [npraep], niej [praep]
B.	jego [akc.npraep], niego [akc.praep], go [nakc.npraep]	je [npraep], nie [praep]	ją [npraep], nią [praep]
N.		nim	nią
Ms.		nim	niej
l.mn.	m1	m2.m3.n.f	
M.	oni	one	
D.	ich [npraep], nich [praep]		
C.	im [npraep], nim [praep]		
B.	ich [npraep], nich [praep]	je [npraep], nie [praep]	
N.	nimi		
Ms.	nich		

Na podstawie tej tabelki oraz dopisanych przed chwilą fakultatywnych tagów tworzymy tabelkę tagów MACY i porównujemy z tagami Morfologika. Zachowujemy przy tym kolejność tagów:

- klasa (część mowy) : liczba : przypadek : rodzaj : osoba : akcentowość : poprzyimkowość.

ON	MACA	Morfologik
on	ppron3:sg:nom:m1.m2.m3:ter	ppron3:sg:nom:m:ter
ono	ppron3:sg:nom:n:ter	ppron3:sg:nom:n.n1.n2:ter
ona	ppron3:sg:nom:f:ter	ppron3:sg:nom:f:ter
jego	ppron3:sg:gen:m1.m2.m3.n:ter:akc:npraep ppron3:sg:acc:m1.m2.m3:ter:akc:npraep	ppron3:sg:gen:m.n.n1.n2:ter:akc:npraep ppron3:sg:acc:m:ter:akc:npraep
niego	ppron3:sg:gen:m1.m2.m3.n:ter:akc:praep ppron3:sg:acc:m1.m2.m3:ter:akc:praep	ppron3:sg:gen:m.n:ter:akc:praep ppron3:sg:acc:m:ter:akc:praep
go	ppron3:sg:gen:m1.m2.m3.n:ter:nakc:npraep ppron3:sg:acc:m1.m2.m3:ter:nakc:npraep	ppron3:sg:gen:m.n.n1.n2:ter:nakc:npraep ppron3:sg:acc:m:ter:nakc:npraep
jemu	ppron3:sg:dat:m1.m2.m3.n:ter:akc:npraep	ppron3:sg:dat:m.n.n1.n2:ter:akc:npraep
niemu	ppron3:sg:dat:m1.m2.m3.n:ter:akc:praep	ppron3:sg:dat:m.n:ter:akc:praep
mu	ppron3:sg:dat:m1.m2.m3.n:ter:nakc:npraep	ppron3:sg:dat:m.n.n1.n2:ter:nakc:npraep
jej	ppron3:sg:gen.dat.f:ter:npraep	ppron3:sg:dat.gen.f:ter:npraep
niej	ppron3:sg:gen.dat.f:ter:praep ppron3:sg:loc:f:ter	ppron3:sg:dat.gen.f:ter:praep ppron3:sg:loc:f:ter
ją	ppron3:sg:acc:f:ter:npraep	ppron3:sg:acc:f:ter:npraep
nią	<b>ppron3:sg:acc:f:ter:praep</b> <b>ppron3:sg:inst:f:ter</b>	<b>ppron3:sg:acc.inst.f:ter:praep</b>
nim	ppron3:sg:inst.loc:m1.m2.m3.n:ter ppron3:pl:dat:m1.m2.m3.n.f:ter:praep	ppron3:sg:inst.loc:m.n.n1.n2:ter ppron3:pl:dat:ter:praep
oni	ppron3:pl:nom:m1:ter	ppron3:pl:nom:m1.p1:ter
one	ppron3:pl:nom:m2.m3.n.f:ter	ppron3:pl:nom:f.m2.m3.n.n1.n2.p2.p3:ter
ich	ppron3:pl:gen:m1.m2.m3.n.f:ter:npraep ppron3:pl:acc:m1:ter:npraep	ppron3:pl:gen:ter:npraep ppron3:pl:acc:m1.p1:ter:npraep
nich	<b>ppron3:pl:gen:m1.m2.m3.n.f:ter:praep</b> <b>ppron3:pl:acc:m1:ter:praep</b> <b>ppron3:pl:loc:m1.m2.m3.n.f:ter</b>	<b>ppron3:pl:gen.loc:ter:praep</b> <b>ppron3:pl:acc:m1.p1:ter:praep</b>
im	ppron3:pl:dat:m1.m2.m3.n.f:ter:npraep	ppron3:pl:dat:ter:npraep
je	ppron3:sg:acc:n:ter:npraep ppron3:pl:acc:m2.m3.n.f:ter:npraep	ppron3:sg:acc:n.n1.n2:ter:npraep ppron3:pl:acc:f.m2.m3.n.n1.n2.p2.p3:ter:npraep
nie	ppron3:sg:acc:n:ter:praep ppron3:pl:acc:m2.m3.n.f:ter:praep	ppron3:sg:acc:n:ter:praep ppron3:pl:acc:f.m2.m3.n.p2.p3:ter:praep
nimi	ppron3:pl:inst:m1.m2.m3.n.f:ter	ppron3:pl:inst:ter

Uwagi:

- (1) Błędnie zanalizowana przez Morfologika forma *nią* (inst sg), ma przypisany tag praep. Jest to błąd! Forma *nią* N. nie ma partnera niepoprzyimkowego.
- (2) Identyczny błąd popełniają autorzy Morfologika w przypadku formy *nich* loc pl.

### 3. LICZEBNIKI ZBIOROWE

W Morfologiku formy typu *siedmioro*, *siedemnaścioro* mają własne lematy. Ponieważ Saloni uznaje formy typu *siedmioro*, *ośmioro* (liczebniki zbiorowe) za reprezentanty lematów *siedem*, *osiem* (liczebniki główne) [Saloni, Świdziński 1998], powinniśmy sprowadzić lematy Morfologika do lematów KIPI (*siedmioro* > *siedem*, *ośmioro* > *osiem* itd.).

PRZYKŁAD:

Morfologik:		
siedmiorga	siedmioro	num:pl:gen:p1.p2
siedmiorgiem	siedmioro	num:pl:inst:p1.p2
siedmiorgu	siedmioro	num:pl:dat.loc:p1.p2
siedmioro	siedmioro	num:pl:acc.nom:p1.p2



## LISTA LEMATÓW LICZEBNIKÓW ZBIOROWYCH

STARY LEMAT MORFOLOGIKA	LEMAT MACY
czterdzieścioro	czterdzieści
czworo	cztery
dwadzieścioro	dwadzieścia
dwanaścioro	dwanaście
dwoje	dwa
dziesięcioro	dziesięć
dziewięćdziesięcioro	dziewięćdziesiąt
jedenascioro	jedenaste
kilkoro	kilka
oboje	oba
osiemdziesięcioro	osiemdziesiąt
ośmioro	osiem
pięcioro	pięć
pięćdziesięcioro	pięćdziesiąt
siedemdziesięcioro	siedemdziesiąt
siedmioro	siedem
sześcioro	sześć
sześćdziesięcioro	sześćdziesiąt
troje	trzy
trzydzieścioro	trzydzieści
czterdzieścioro	czterdzieści

Ponieważ Morfologik nie podaje w obrębie liczebników zbiorowych wołacza, tag voc dopisujemy do tagu nom

nom → nom.voc
---------------

### AKOMODACYJNOŚĆ liczebników zbiorowych

„Składnia liczebników zbiorowych występujących z rzeczownikami jest prawie taka sama jak liczebników głównych od pięciu wzwyż. Różnica między nimi polega tylko na tym, że liczebniki zbiorowe łączą się ze swoimi określeniami w związku zgody w celowniku i miejscowniku, podczas gdy liczebniki główne łączą się tak jeszcze w narzędniku, np.

C. *Podarował pamiątki pięciu uczniom – pięciorgu uczniom;*

Ms. *Myszę o pięciu uczniach – pięciorgu uczniach,*

ale N. *Kieruję pięcioma uczniami – pięciorgiem uczniów.*”

[WSPP II 2004, tu: *liczebnik*]”

Z rozważań tych wynika, że powinniśmy odpowiednim przypadkom przypisać następujące akomodacyjności (rec = składnia rządu, congr = składnia zgody):

- nom.gen.acc.voc.inst → rec
- dat.loc → congr

Z rozdziału książki Saloniego, Świdzińskiego [Saloni, Świdziński 1998: 208] poświęconego akomodacji liczebnikowej w zakresie przypadku wynika, że rząd dopełniaczowy (rec) mamy we wszystkich przypadkach:

- nom.gen.dat.loc.acc.inst.voc  $\Rightarrow$  rec

W dat i loc szerzy się rzeczywiście składnia rządu, uznawana jednak przez WSPP za niepoprawną (por [WSPP I: *dwoje, czworo, troje*]).

Ostatecznie ustalamy regułę, która łączy stanowisko normatywne z koncepcją Saloniego:

nom.gen.acc.inst.voc $\Rightarrow$ rec dat.loc $\Rightarrow$ congr.rec
---

Formy te — co prawda — są niepoprawne, ale — powszechne.

#### Bibliografia:

GWJP-M1 = *Gramatyka współczesnego języka polskiego. Morfologia*, pod red. R. Grzegorzczkovej, R. Laskowskiego, H. Wróbla, wyd. 2. zm., Warszawa 1998

Przepiórkowski 2004 = A. Przepiórkowski, *Korpus IPI PAN. Wersja wstępna*. Warszawa 2004

Radziszewski, Śniatowski 2011 = A. Radziszewski, T. Śniatowski, *Maca: a configurable tool to integrate Polish morphological data*, 2011 [materiały z konferencji: FreeRBMT, Barcelona, 2011]

Miłkowski 2010 = M. Miłkowski, *Developing an open-source, rule-based proofreading tool*. W: *Software: Practice and Experience*, 40, ss. 543–566. <http://morfologik.blogspot.com>

Woliński 2004 = M. Woliński, *System znaczników morfosyntaktycznych w korpusie IPI PAN*. W: *Polonica*, 2004, XII, s. 39–55.

WSPP I i II = *Wielki słownik poprawnej polszczyzny* PWN, pod red. A. Markowskiego, Warszawa 2006, t. I-II.

Saloni, Świdziński 1998 = Z. Saloni, M. Świdziński, *Składnia współczesnego języka polskiego*, wyd. 4. zm., Warszawa 1998.