

PIOTR WOJDAK\* | UNIwersytet SZCZECIŃSKI

# Oboczność *-a/-u* w D.lp. a dwurodzajowość M2/M3 w świetle badań korpusowych sterowanych słownikami

Słowa kluczowe: rzeczownik, rodzaj męski, dwurodzajowość, końcówka fleksyjna, dopełniacz liczby pojedynczej, wariantywność, korpus.

doi: <http://dx.doi.org/10.31286/JP.100.2.3>

W dobie językoznawstwa korpusowego aktualizują się wyzwania badawcze wynikające z ułatwionego dostępu do danych tekstowych. Po zarysowaniu problematyki związanej z krzyżowaniem się wariantywności końcówek D.lp. masculinów z ich dwurodzajowością (Wojdak 2020) wypada zatem wykorzystać korpus (NKJP) także do zweryfikowania listy dwurodzajowych jednostek rzeczownikowych<sup>1</sup> typu M2/M3 z obocznością *-a/-u* (sporządzonej uprzednio na podstawie współczesnych polskich słowników) oraz do ustalenia względnych frekwencji obu końcówek i obu alternantów rodzajowych. W tym wypadku nad danymi z korpusu trzeba się jednak napracować, aby je wydobyć i zinterpretować<sup>2</sup>. Zanim przedstawione zostaną wyniki badań korpusowych, omówimy więc zasady pozyskiwania danych z NKJP oraz wyjaśnimy, skąd się wzięła wstępna lista dwurodzajowych jednostek rzeczownikowych typu M2/M3 z obocznością *-a/-u* w D.lp. (siłą rzeczy odwołamy się do wcześniejszego opracowania – Wojdak 2013).

W artykule nie zostaną wskazane czynniki (semantyczne, morfologiczne czy jeszcze inne) współdecydujące o tym, które to jednostki rzeczownikowe odznaczają się zarówno dwurodzajowością M2/M3, jak i wariantywnością końcówek *-a/-u* w D.lp. Autor, owszem, zadawał sobie pytanie o takie prawidłowości, ale nie znalazł na nie odpowiedzi.

## 1. Pozyskiwanie danych ze słowników i korpusu

Badania korpusowe przeprowadzono na zbiorze jednostek, które w świetle niektórych przynajmniej współczesnych słowników języka polskiego odznaczają się dwurodzajowością M2/M3,

\* piotr.wojdak@usz.edu.pl; ORCID: 0000-0002-8636-7107

1 Celowo unika się tutaj w odniesieniu do takich jednostek terminu *lekssem (rzeczownikowy)* czy nawet *rzeczownik*, gdyż w świetle kanonicznych definicji rzeczowniki rozumiane jako klasa leksemów (część mowy) nie powinny mieć w swoim paradygmacie opozycji rodzajowych.

2 O „nadmiarze” danych do analiz językoznawczych pisze ostatnio Mirosław Bańko (2018).

a zarazem neutralną znaczeniowo obocznością *-a/-u* w D.lp.<sup>3</sup> Zbiór ten sporządzono na podstawie słownika apelatywnych jednostek wielorodzajowych współczesnej polszczyzny (SAJW, w ramach monografii: Wojdak 2013), uwzględniającego dane z siedmiu źródeł: EPLar, ISJP, SGJP<sup>4</sup>, USJP, WSJPDun, WSPP, WSWO (dziś do tego zestawu dołączony by został WSJP PAN). Jako źródłowe siłą rzeczy wykorzystane mogły być tylko te słowniki, które rozróżniają – eksplcytnie lub chociaż implcytnie – trzy rodzaje męskie (wewnątrzżeńskie opozycje biernikowe)<sup>5</sup>.

Ze względu na oboczność D.lp. z SAJW wybrano 62 z 468 (ponad 13 procent) haseł typu M2/M3:

BAKŁAŻAN, BATAT/PATAT, BATON, BEFSZTYK, BEZAN, BISZKOPI, BROKUŁ, BUS, CALVADOS, CHABER, CHIP/CZUPI, CIENKUSZ, DWÓJNIAK, DYNGUS, ESEMES, FANT, FART, FILET, FILODENDRON, FLOKS, GICZOŁ, GIMBUS, GLOBUS, GNIOT, GOKART/GO-CART, GYROS, HALOBIONT, HIACYNT, KART, KAWON, KLEKOT, KLOPS, KOKOS, KONDOM, ŁYK, MELEKS/MELEX, MIGDAŁ, MIKST, MLECZ, MLECZAJ, NENUFAR, NIEFART, OMLECICK, OMLET, PAINTBALL, PARCH, PIRUET, PLATFUS, POST, POLAROID, PULPET, RAUSZ, SMS, STĘP, SWEET, ŚWIERK, TOST, TRÓJNIAK/TROJNIAK, WALKOWER, WILCZOMLECZ, WIS/WIS/VIS/VIS, ZIMOWIT<sup>6</sup>.

Dwa hasła z wyżej przytoczonej listy – ESEMES i MIGDAŁ – w SAJW nie są notowane z końcówką *-u* w D.lp., a więc zostały uwzględnione na zasadzie wyjątku. Pierwsze z nich sprawdzono w korpusie ze względu na bliskość do SMS, a drugie pod wpływem orzeczenia M. Bańki (2001: 229; 2002: 149); w obu wypadkach korpus potwierdził oboczność *-a/-u*.

Spoza zbioru haseł SAJW – czyli tym bardziej w ramach wyjątku – uwzględniono w badaniach tylko jedną jednostkę: KONDOM (ze względu na walor porównania z jednostką KONDOM); w tym wypadku korpus potwierdził orzeczony wcześniej (Wojdak 2013: 209) brak końcówki *-u*

3 W językoznawstwie korpusowym rozróżnia się badania: (1) ilustrowane przez korpus (*corpus illustrated*), (2) oparte na korpusie (*corpus based*) i (3) sterowane korpusem (*corpus driven*) (Górski 2012: 292–293). Realizowany jest tutaj drugi z wymienionych typów: oblicza się bowiem frekwencje jednostek ze zbioru ustalonego niezależnie od korpusu, nie wykluczając falsyfikacji czy modyfikacji przyjętych hipotez. Chodzi zatem o coś więcej niż tylko poszukiwanie przykładów bez zwracania uwagi na frekwencje jednostek (typ pierwszy), ale nie ma mowy o całkowitym zdaniu się na korpus, bez „uprzedzeń” co do zakresu badania (typ trzeci, który ma jednak charakter raczej tylko idealny).

4 W trakcie opracowywania SAJW (w latach 2011–2013) wykorzystywano niepublikowaną, ale udostępnioną przez autorów wersję SGJP oznaczoną jako 1.0.27 (późniejszą niż SGJP 2007, a wcześniejszą niż SGJP 2012).

5 Nie wszystkie współczesne słowniki, nawet spośród tych największych, por. PSWP, spełniają to kryterium. Spośród słowników uwzględnionych w SAJW, jeśli chodzi o informacje o bierniku, zastrzeżenia budzi USJP. Przykładowo, przywoływane w poprzednim artykule (Wojdak 2020: 47) masculina CERATYT, EMBRION i MSZYWIOL są notowane z wyjątkową jak na rzeczowniki żywotne (oboczna bądź jedyną) końcówką *-u* w D.lp., ale bez wskazania biernika. Nie wiadomo, czy tej formy, a tym samym (pod)rodzaju męskiego, należy się domyślać wedle znaczenia (M2), czy wedle końcówki (M3). Częściej jeszcze w USJP daje o sobie znać biernikowe nieookreślenie masculinów nieżywotnych (zob. tam np. BATON, BROKUŁ, CHABER, CHIP, DYNGUS); w takich wypadkach w SAJW kierowano się zasadą domyślnego (wedle znaczenia) nieobocznego rodzaju M3 (oznaczenie z trójką w nawiasie: M(3)). Ich liczba nie była jednak na tyle duża, aby całkowicie zrezygnować z tak ważnego i popularnego słownika, jakim jest USJP.

6 Faktycznie w SAJW zapis haseł i ich wariantów jest inny, por. np. bata|t lub patat|, mele|ks/x, mlecz. Pionowa kreska rozdziela niezmienną część hasłowanej jednostki (temat graficzny) od swoistej części końcowej hasła (jeśli taka w nim występuje), a na szeregowanie alfabetyczne znak ten działa tak jak koniec wyrazu graficznego (dlatego np. najpierw jest polst, a potem polaroid). Tutaj zastosowano zapis zgodny z konwencją przyjętą w całym artykule (a także w: Wojdak 2020; Wojdak, w druku).

w D.lp., ale nie potwierdził (tylko) jednorodzaowości (M<sub>3</sub>). Dodanie tej jednostki zrównoważono połączeniem haseł GOKART/GO-CART i KART (uzasadnienie dalej). Ostatecznie liczba przebadanych jednostek pozostała taka sama jak liczba haseł SAJW z prezentowanej wyżej listy (62). Dwa hasła przejęte z SAJW z dodaniem wariantów (w zapisie jednostek na drugim miejscu): SMS/sms i DYNGUS/ŚMIGUS-DYNGUS. Ten ostatni – zestawieniowy – wariant (pełna nazwa wielkanocnego zwyczaju) nie mógł się w SAJW pojawić ze względu na przyjęte tam ograniczenie co do jednostek graficznie złożonych (także tych z łącznikiem).

Przypomnijmy, że warunkiem włączenia jednostki rzeczownikowej do SAJW było wskazanie jej dwurodzajowości (wielorodzajowości) w co najmniej dwóch źródłach. Przykładowo tego warunku nie spełniła komentowana w poprzednim artykule (Wojdak 2020: 52) jednostka FAUL – dwurodzajowo notowana tylko w jednym słowniku (WSJPDun).

Przejęta z SAJW – z nieznacznymi modyfikacjami – lista jednostek rzeczownikowych M<sub>2</sub>/M<sub>3</sub> z obocznością końcówek *-a/-u* jest więc wybiórcza, ale też i dość reprezentatywna dla współczesnego stanu leksykograficznego.

Zarówno w SAJW, jak i w niniejszym opracowaniu jednostki wyodrębniano, dopuszczając z jednej strony ich wariantywność formalną (jakby chcąc przesądzić, że chodzi tu o jednostki rzeczownikowe „większe” niż leksemy, por. np.: TRÓJNIAK/TROJNIAK, CHIP/CZIP), z drugiej strony polisemię – jeśli tylko nie osłabiało to ich tożsamości gramatycznej: fleksyjnej (także co do oboczności D.lp.) ani składniowej (rodzajowej, choćby właśnie tej „podwójnej”). To ze względu na różnice gramatyczne nie scalono równoważnych znaczeniowo jednostek KONDOM i KONDON (różniących się także pragmatycznie – nacechowaniem drugiej jednostki). Rozdzielenie w SAJW synonimicznych jednostek GOKART/GO-CART i KART wynikało nie z różnicy ilościowej w temacie (wariant pełny i skrócony), lecz z drobnej – w istocie arbitralnie przyjętej – różnicy fleksyjnej (co do D.lp. alternantu M<sub>3</sub> – GOKART/GO-CART: *-a/-u*; KART: *-u*). Ponieważ korpus jej nie potwierdził, scalono obie jednostki.

W wypadku kilku dalszych jednostek, zob. BISZKOPT, FANT, GLOBUS, GNIOT, MIGDAŁ w zamieszczonej dalej tabeli wraz z przypisami, badania korpusowe doprowadziły do innej niż rozszerzenie wariantywności haseł (GOKART/GO-CART/KART, a także DYNGUS/ŚMIGUS-DYNGUS i SMS/sms) rewizji wyjściowego – takiego jak w SAJW – ujęcia. Trudno się temu dziwić, skoro NKJP był dla SAJW źródłem co najwyżej pomocniczym, wobec rozbieżności między źródłowymi słownikami siłą rzeczy znalazło się więc w tamtym opracowaniu dużo rozstrzygnięć arbitralnych (dyktowanych również duchem normatywizmu<sup>7</sup>).

Frekwencje form dopełniaczowych na *-a* i na *-u* oraz męskozwierzęcych i męskonieżywotnych form biernikowych zliczano (w latach 2015–2018) z NKJP na podstawie wyszukiwarki

7 Ambicje normatywne można zilustrować różnicą w opisie jednostki ESEMES w stosunku do SMS. Otóż uznano, że jeśli ktoś „oswaja” obcy z pochodzenia skrótowiec pisownią (zgodną z polską – literową – wymową), to wybiera w D.lp. „swojską” końcówkę *-a* (czyli że *-u* do tej pisowni nie pasuje). Trochę analogicznie można by się spodziewać, że jeżeli rozdzielnie – ze względu na D.lp. – ujęte zostały synonimiczne jednostki GOKART i KART, to końcówka *-a* powinna być dowartościowana raczej w ramach jednostki skróconej (bardziej oswojonej) niż pełnej. Niestety, jak już wskazano wyżej, w SAJW przyjęto rozstrzygnięcie odwrotne.

Trzeba przyznać, że zarówno sposób (mały stopień) wykorzystania korpusu, jak i „normatywizm” w monografii Wojdak 2013 były przedmiotem krytyki lub przynajmniej wątpliwości recenzentów (Zbróg 2014; Żmigrodzki 2014).

PELCRA<sup>8</sup>: z podkorpusu zrównoważonego (ok. 300 mln segmentów) – gdy reprezentowane w nim były wszystkie cztery typy form danej jednostki, z pełnego korpusu – gdy w podkorpusie zrównoważonym choćby jeden z czterech typów form danej jednostki nie był reprezentowany (miał zerową frekwencję).

Przy jednostkach o wysokiej frekwencji podkorpus zrównoważony pozwala uniknąć zliczania tysięcy wystąpień (choć powyżej 10 tys. wyszukiwarka i tak ich nie wykazuje), a jednocześnie jest bardziej reprezentatywny (więcej na temat reprezentatywności: Górski, Łaziński 2012). Nie faworyzuje języka potocznego, który w NKJP wyłania się przede wszystkim z nieoficjalnych tekstów internetowych (teksty mówione są mniej liczne i trudniej interpretowalne). W wypadku jednostek dwurodzajowych M2/M3 o obocznych końcówkach *-a* i *-u* w D.lp. nadreprezentacja języka potocznego może skutkować zawyżonym udziałem alternantu M2 oraz końcówki *-a* w D.lp. Danych szukano jednak głównie z nastawieniem na uzyskanie niezerowych wartości dla obu końcówek D.lp. i obu rodzajów w B.lp., dlatego także w wypadku jednostek o ogólnie (stosunkowo) wysokiej frekwencji nie poprzestawano na podkorpusie zrównoważonym.

Czasem trzeba się było naszukać odpowiednich przykładów w dużej – z racji szeroko pojętej homonimii – puli wystąpień danej formy, a i tak nie zostały one znalezione, na przykład *sweet* ma aż 6816 wystąpień w korpusie pełnym (m.in. jako angielski przymiotnik ‘słodki’<sup>9</sup>, poufały apostrof/wołacz odnoszący się do kobiet, składnik adresu mailowego, nazw własnych), ale ani jednej formy B.lp. w znaczeniu ‘odmiana jazzu’ (!). Podobnie było z formami *kart* i *karta* (odpowiednio aż 7039 i 5107 wystąpień już w samym podkorpusie zrównoważonym), bezowocnie przeszukiwanymi w znaczeniu nazwy pojazdu (KART = GOKART)<sup>10</sup>.

Największym wyzwaniem przy pozyskiwaniu danych korpusowych była – zakładana na podstawie słowników – wieloznaczność niektórych jednostek. Trudności w rozróżnianiu znaczeń (w dezambiguacji użyć) oraz użycia nieprzewidywalne w świetle wyjściowych opisów słownikowych skłaniały do poszukiwań alternatywnego ujęcia monosemicznego („unikowa” dezambiguacja – przez rozszerzenie znaczenia, por. FANT) albo przynajmniej do rozszerzania znaczeń i zmniejszania ich liczby (por. PIRUET) – żeby ograniczyć liczbę użyć „odrzuć”, nie dających się przyporządkować do żadnego ze znaczeń, także spoza horyzontu zakreślonego przez wiadome skrzyżowanie fenomenów gramatycznych.

Monosemantyzacja bywa wprawdzie trudna czy nawet niemożliwa do zrealizowania (albo może mieć nieprzekonującą realizację), ale jako ogólny kierunek metodologiczny jest słuszna (wszak nauka dąży do formułowania sądów ogólnych). A korpus, ujawniając realną różnorodność odniesień, bogatszą, niż wynikałoby to z definicji słownikowych, paradoksalnie wcale nie osłabia dążenia do ujęć monosemicznych, lecz jeszcze je potęguje<sup>11</sup>.

8 Druga z wyszukiwarek NKJP – POLIQARP – daje wprawdzie większe możliwości filtrowania pod względem kryteriów gramatycznych, ale i tak (poza niewielkim – milionowym – podkorpusem, dla którego anotacja gramatyczna została opracowana „ręcznie”) nie rozróżnia form synkretycznych rzeczowników, w tym M. i B. lp. rzeczowników męskozwrotnych czy D. i B. lp. rzeczowników męskozwrotnych. W stosunku do PELCRY istotną słabością POLIQARPA jest to, że zwraca maksymalnie 1000 wyników na zapytanie (PELCRA – 10 tys.).

9 Istnieje problem niewystarczającej selekcji źródeł NKJP, skoro trafiają się tam niestety także dłuższe teksty angielskie.

10 Znalaziono tylko jeden wątpliwy (niejasny) przykład z języka mówionego: *czekaj czekaj czekaj trzymaj ten tyłek zobacz jaki ten ma włochaty taki k a r t nie? ten komar przestań* (bez szerszego kontekstu w NKJP).

11 Poszukiwanie tego, co semantycznie wspólne, pod jakimś względem tożsame, oczywiście otwiera drogę także do bardziej rozbudowanych opisów znaczenia, ukazujących wewnętrzne zróżnicowanie pojęcia.

Bardziej dyskusyjne jest niejednolite podejście do jednostek o zachowanej polisemiczności, ich frekwencje zliczano bowiem aż na trzy alternatywne sposoby: (1) rozdzielnie (według znaczeń) bez sumowania – GLOBUS, GNIOT, MIGDAŁ (znaczenia 1 i 2, por. punkt następny), MIKST, PARCH; (2) rozdzielnie, ale z sumowaniem – BEZAN, BISZKOPT, MIGDAŁ (znaczenia 1a i 1b); (3) łącznie, tzn. bez rozróżniania (pod)znanień – MLECZ, KOKOS, OMLECIK, PIRUET. W wypadku BISZKOPTA wyjątkowo ograniczono się do podkorpusu zrównoważonego mimo tamtejszych deficytów jednego z – osobno liczonych – podznanień (zob. przypis 16).

W sumie rozwarstwienie znaczeniowe uwzględniono – na różne sposoby i z różnym skutkiem – w wypadku 11 jednostek (w tym przy zliczaniu frekwencji – 7 jednostek, por. w akapicie wyżej sposoby (1) i (2)).

Sieci korpusowych zliczeń nie zarzucano jednak na tyle szeroko, by objąć nią wszystkie znaczenia w ramach homonimii czy niepewnej (gramatycznie rozchwianej) polisemii. Przykładowo nie odnotowano GLOBUSA w znaczeniu ‘model kuli ziemskiej’, bo nie było podejrzeń, że może to być jednostka dwurodzajowa M2/M3 z obocznością -a/-u. Dodatkowo wyszczególniano składowe frekwencje niektórych jednostek o wariantowości formalnej: ilościowej (związanej z redukcją, DYNGUS/ŚMIGUS-DYNGUS), fonetyczno-graficznej (TRÓJNIAK/TROJNIAK) lub tylko graficznej (CHIP/CZIP, SMS/SMS). Pojedynczo były liczone konteksty identyczne, wynikające na przykład z racji zamieszczania tych samych ogłoszeń w różnych numerach gazety/czasopisma albo przytaczania czyjejs wypowiedzi w komentarzach internetowych<sup>12</sup> (szczególnego przykładu dostarcza wyrażenie *Temat postu*, które w PELCRZE zrównoważonej pojawia się 827 razy (!)). Nie uwzględniano nazw własnych (*Chip* – czasopismo; *Zimowit* – klub sportowy, także osiedle mieszkaniowe; *Melex*, *Polaroid* – firmy). Zliczane były jednak formy zapisane (błędnie) wielką literą, które odnosiły się do produktu firmy.

Pomijano przykłady, które nie poddawały się jednoznacznej identyfikacji przypadku: ze względu na nierozróżnialną wariantowość dopełniacza i biernika (*mamo! chcę tosta, chcesz tosta?, tosta chciałeś, prawda?, mogą tosta?, może (jakiegoś) tosta?*<sup>13</sup>), nierozróżnialne schematy składniowe (mianownik lub biernik w konstrukcjach z *się*: *klops to czym dłużej się piecze*) czy też użycie nawiasowe (mianownik lub biernik: *uprawiać rośliny odstraszające gryzonie (czosnek, nostrzyk, wilczomlecz)*).

Pomijano również przykłady o niepewnym znaczeniu (niedające się ujednoznaczyć):

Mnie film też nie zachwyił. Początek i rozwinięcie akcji było świetne, ale finał marny, jakby komuś zabrakło pomysłów (urwanie się *globusa*, przypadkowe pojawienie się policji).

Bez znajomości filmu nie można zdecydować, czy w powyższym użyciu chodzi o znaczenie podstawowe (geograficzne) *globusa*, czy przenośne (‘głowa’). Pierwsze jednak i tak nie byłoby specyfikowane (rodzaj tylko M3).

12 Ogólnie należy wspomnieć o wcale nierzadkim zjawisku powtarzania się (wielokrotnego występowania) tekstów w NKJP.

13 O rozchwianiu przypadku form rzeczownika współwystępujących z czasownikiem modalnym CHCIEĆ oraz modalizatorem MOŻE świadczą frekwencje połączeń z formami *kawy* i *kawę* (w korpusie pełnym PELCRY): *chcesz kawy* – 16, *chcesz kawę* – 17; *może kawy* – 26, *może kawę* – 10.

Pominięto też użycie przytoczone jako niepoprawne (przynajmniej w opinii autora wypowiedzi)<sup>14</sup>:

Nie można na przykład nakłaniać ludzi, by wysyłał *esemesa*, tylko *esemes*, bo wiadomo, że zawiązuje się nie buta tylko but.

## 2. Wstępna lista jednostek dwurodzajowych M2/M3 z obocznością -a/-u w D.lp. w świetle danych frekwencyjnych

Wyniki badań korpusowych zestawiono w poniższej tabeli.

Tabela. Weryfikacja korpusowa dwurodzajowości M2/M3 i oboczności -a/-u w D.lp. wybranych jednostek rzeczownikowych (na podstawie wersji zrównoważonej (z) lub pełnej (p) NKJP)

JEDNOSTKA RZECZOWNIKOWA I JEJ ZNACZENIE	DANE FREKWENCYJNE Z NKJP					WNIOSKI <sup>15</sup>	
	typ korpusu	formy D.lp.		wykładniki rodzaju w B.lp.		koń- cówki D.lp.	rodzaj grama- tyczny
		na -a	na -u	M2	M3		
BAKŁAŻAN 'roślina warzywna i sam jej jadalny owoc'	z	22	2	9	6	-a/-u	M2/M3
BATAT/PATAT 'tropikalna roślina upraw- na i sama jej jadalna mą- czysta bulwa'	p	0	1	0	2	-u	M3
BATON 'wyrób cukierniczy o podłużnym kształcie'	z	15	3	29	14	-a/-u	M2/M3
BEFSZTYK 'porcja smażonego mięsa wołowego'	z	28	5	1	74	-a/-u	M3/M2
BEZAN ŻEGLUGA a) 'tylny maszt, = bezanmaszt'; b) 'rozpinany na nim żagiel, = bezanżagiel'	p	5 (4 + 1)	1 (0 + 1)	2 (1 + 1)	2 (1 + 1)	-a/-u	M3 = M2

14 Pod wpływem uwagi recenzenta autor zdał sobie jednak sprawę, że jest to arbitralne rozwiązanie, bo dowodzenie niepoprawności danej formy pośrednio świadczy o jej używalności.

15 Na pierwszym miejscu (przed ukośnikiem) notowano końcówki i alternanty rodzajowe o wyższej frekwencji. W wypadku jednakowych frekwencji stosowano znak „=”.

BISZKOPI <sup>16</sup> a) 'ciasto'; b) 'ciastko z tego ciasta'	z	3 (2 + 1)	12 (12 + 0)	4 (2 + 2)	25 (16 + 9)	-u/-a	M <sub>3</sub> /M <sub>2</sub>
BROKUŁ 'jednoroczna roślina warzywna i jej jadalne nierozwinięte kwiatostany i części łodyg'	z	7	5	3	3	-a/-u	M <sub>3</sub> = M <sub>2</sub>
BUS <i>pot.</i> 'mikrobus'	z	223	15	102	30	-a/-u	M <sub>2</sub> /M <sub>3</sub>
CALVADOS 'francuska wódka z jabłek'	z	1	17	1	18	-u/-a	M <sub>3</sub> /M <sub>2</sub>
CHABER 'roślina polna o niebieskich kwiatach'	p	15	25	1	11	-u/-a	M <sub>3</sub> /M <sub>2</sub>
CHIP/CZIP <sup>17</sup> 'miniaturowy układ scalony'	z	51 (43/8)	58 (52/6)	12 (8/4)	74 (67/7)	-u/-a	M <sub>3</sub> /M <sub>2</sub>
CIENKUSZ <i>przest.</i> 'słaby lub zły jakości alkohol (wino, piwo, miód)'	p	18	0	10	8	-a	M <sub>2</sub> /M <sub>3</sub>
DWÓJNIAK 'miód pitny o określonych proporcjach (w połowie rozcieńczony wodą lub sokiem)'	p	9	2	0	13	-a/-u	M <sub>3</sub>
DYNGUS/ŚMIGUS-DYNGUS <sup>18</sup> <i>pot.</i> 'zwyczaj polewania się wodą w Poniedziałek Wielkanocny'	z	27 (14/13)	1 (1/0)	6 (3/3)	21 (2/19)	-a/-u	M <sub>3</sub> /M <sub>2</sub>
ESEMES 'wiadomość wysyłana przez telefon komórkowy'	z	37	3	218	28	-a/-u	M <sub>2</sub> /M <sub>3</sub>

16 W nawiasie odrębnie wykazano frekwencje znaczeń 'ciasto' i 'ciastko' (odpowiednio – przed znakiem „+” i po). Ze względu na brak wystąpień formy D.lp. *biszkopty* w znaczeniu 'ciastko' podjęto próbę zliczania frekwencji jednostki BISZKOPI także w korpusie pełnym, ale spięrzyły się tam trudności w rozróżnianiu obu znaczeń. W SAJW uznano tę jednostkę za dwurodzajową tylko w (pod)znaczeniu 'ciastko' (jako 'ciasto' – tylko M<sub>2</sub>).

17 W nawiasach osobno (odpowiednio: przed ukośnikiem i po ukośniku) wskazano liczbę wystąpień form wariantu CHIP i CZIP.

18 W nawiasach osobno (odpowiednio: przed ukośnikiem i po ukośniku) wskazano liczbę wystąpień form wariantu prostego DYNGUS i (nieujętego w SAJW) zestawienia ŚMIGUS-DYNGUS (z uwzględnieniem także pisowni rozdzielnej (bez łącznika) oraz zmodyfikowanego członu: *śmingusa*).

FANT <sup>19</sup> 'rzecz, o której myślimy, że jest cenna'	z	1	9	1	22	-u/-a	M3/M2	
FART <i>pot.</i> 'sprzyjająca okoliczność, = powodzenie, szczęście'	z	11	56	74	73	-u/-a	M2/M3	
FILET 'płat ryby bez ości lub mięsa bez kości'	z	15	12	2	45	-a/-u	M3/M2	
FILODENDRON 'drzewo, krzew lub pnącze z rodziny obrazkowatych, także roślina doniczkowa'	p	13	3	3	8	-a/-u	M3/M2	
FLOKS 'roślina ozdobna'	p	3	0	1	1	-a	M2 = M3	
GICZOŁ <i>pot. reg.</i> '[zwykle w lm.] noga'	p	1	0	0	1	-a	M3	
GIMBUS 'autobus szkolny'	z	63	2	35	12	-a/-u	M2/M3	
GLOBUS <sup>20</sup>	1. 'głowa'	p	1	0	0	7	-a	M3
	2. 'wrażenie zatykania się gardła (pod wpływem negatywnych emocji lub na tle alergicznym)'	p	8	3	9 <sup>21</sup>	6 <sup>22</sup>	-a/-u	M2/M3

19 Wobec referencjalnej różnorodności użyć jednostki FANT w korpusie zdecydowano się na jedną szeroką formułę definicyjną, która jest jeszcze pojemniejsza niż jej trzy najczęściej rozróżniane w słownikach znaczenia razem wzięte: 1. 'przedmiot do wygrania na loterii'; 2. 'w niektórych grach: rzecz oddawana w zastaw'; 3. 'rzecz pochodząca z kradzieży'. W SAJW wyspecyfikowano dwa pierwsze z tych znaczeń, tutaj uwzględniono także trzecie, uznając, że wprawdzie genetycznie ma ono charakter środowiskowy, ale jego współczesne użycie zdecydowanie wykracza poza gwarę złodziejską (upowszechniło się w mediach).

Tylko użycie w znaczeniu zinterpretowanym jako 'problem, szkopuł' (będące pochodną frazeologizmu *Co z tym fantem zrobić?*) – *W szkołach czort wie, co się dzieje. Weźcie chociażby taki fant* – nie dawało się podciągnąć pod proponowane szerokie znaczenie i nie było liczone.

20 W SAJW GLOBUS figuruje tylko w znaczeniu 'głowa', ale korpus wykazał, że jednostka o takiej postaci w innym znaczeniu ('wrażenie zatykania się gardła') jest dwurodzajowa, a zarazem odznacza się obocznością -a/-u.

21 W niektórych wypadkach nie można też było wykluczyć innego – niepotwierdzonego jednak w słownikach – s frazeologizowanego użycia zwrotu *dostać czy mieć globusa*, które znaczeniowo bliskie by było zwrotowi *dostać/mieć kręčka* ('zacząć wariować/'wariować').

22 Uwzględniono także użycie nawiązujące do przypadku bohaterki powieści Elizy Orzeszkowej, ale utożsamiające globus z bólem głowy: *Emilia Korczyńska z „Nad Niemnem” na tzw. gl o b u s, czyli ból głowy dostawała proszki bromowe.*



GNIOT <i>pot.</i> <i>pejorat.</i>	1. 'nieudany wypiek'	z (p)	0 (0)	0 (0)	0 (1)	3 <sup>23</sup>	–	M <sub>3</sub> (M <sub>3</sub> /M <sub>2</sub> )
	2. 'wytwór pracy intelektualnej o niskiej wartości (utwór literacki, film, wypracowanie szkolne)'	z	7	2	4 <sup>24</sup>	7	-a/-u	M <sub>3</sub> /M <sub>2</sub>
GOKART/GO-CART/KART <sup>25</sup> 'mały jednosobowy samochód wyścigowy bez nadwozia'		z	8	1	8	3	-a/-u	M <sub>2</sub> /M <sub>3</sub>
GYROS 'grecka potrawa z mięsa'		p	20	2	11	27	-a/-u	M <sub>3</sub> /M <sub>2</sub>
HALOBIONT <i>BIOLOGIA</i> 'organizm, który wymaga środowiska o dużym zasoleniu' <sup>26</sup>		p	0	0	0	0	–	–
HIACYNT 'roślina ozdobna'		p	18	20	8	6	-u/-a	M <sub>2</sub> /M <sub>3</sub>
KAWON <i>przest., reg. lwowsko-krakowski</i> 'arbuz'		p	5	2	0	1	-a/-u	M <sub>3</sub>
KLEKOT <i>pot. pejorat.</i> 'stary zużyty przedmiot (zwłaszcza pojazd)'		p	3	0	0	1	-a	M <sub>3</sub>
KLOPS 'okrągły kotlet'		p	13	1	4	8	-a/-u	M <sub>3</sub> /M <sub>2</sub>
KOKOS a) 'owoc palmy kokosowej'; b) <i>pot.</i> 'palma kokosowa'		p	77	61	16	20	-a/-u	M <sub>3</sub> /M <sub>2</sub>

Pominięto natomiast użycie, w którym *globus* pojawia się w ramach dwuczłonowej łacińskiej nazwy objawu hysterii (polski odpowiednik to *kula histeryczna*), gdyż taki kontekst determinuje biernik równy mianownikowi (nieodmienność): *Byliśmy już u lekarzy, którzy stwierdzili globus histerikus i zalecili jeść hydroksyzynę w takich sytuacjach*.

23 W korpusie pełnym nie zliczono B.lp. M<sub>3</sub> (z puli 961 przykładów), gdy się okazało, że i tak nie występują tam formy D.lp. (w tym samym znaczeniu 'nieudany wypiek') – ani na -a, ani na -u.

24 Pominięto jedno użycie w odniesieniu do samochodu (negatywnie ocenianego) – ze względu na rozszerzenie znaczenia na przedmiot konkretny, ale bardziej jeszcze ze względu na to, że wśród męskich nazw (marek) samochodów narzuca się rodzaj M<sub>2</sub>.

25 W grę wchodziłyby także postaci hasłowe GO-KART i GOCART, ale ich formy nie były reprezentowane w uwzględnionym podkorpusie zrównoważonym.

26 W SAJW przyjęto, że jednostka ta jest dwurodzajowa tylko w odniesieniu do roślin (w odniesieniu do zwierząt nie ma rodzaju M<sub>3</sub>). Ze względu na brak wystąpień to systemowo słuszne ograniczenie nie miało jednak znaczenia.

KONDOM '= prezerwatywa'		z	14	5	5	15	-a/-u	M3/M2
KONDON <i>pot.</i> 'kondom'		p (z) <sup>27</sup>	19 (2) <sup>28</sup>	0 (0)	22 (1)	3 (0)	-a	M2/M3
ŁYK <i>pot.</i> 'ilość płynu lub powietrza przefiltrowana jednorazowo'		z	26	25	25	742	-a/-u	M3/M2
MELEKS/MELEX 'mały odkryty samochód z napędem elektrycznym'		p	29	5	12	19	-a/-u	M3/M2
MIGDAŁ	1. 'roślina i jej jadalne nasienie' <sup>29</sup>	p (z)	44 (12)	16 (6)	3 (0) <sup>30</sup>	27 (8)	-a/-u	M3/M2
	2. 'narząd ciała' <sup>31</sup>	p (z)	106 (1)	5 (1)	5 (0)	161 (8) <sup>32</sup>	-a/-u	M3/M2
MIKST SPORT	a) 'gra między parami mieszanymi'	p (z)	193 (10)	1 (0)	40 (7)	5 (0)	-a/-u	M2/M3
	b) 'para biorąca udział w tej grze'	p (z)	10 (0)	4 (1)	2 (0)	13 (1)	-a/-u	M3/M2
MLECZ 'roślina': a) <i>pot.</i> '= mniszek lekarski'; b) 'podobna do mniszka lekarskiego'		p	83	45	2	27	-a/-u	M3/M2
MLECZAJ 'grzyb kapeluszowy wydzielający przy uszkodzeniu białawy sok'		p	8	0	11	0	-a	M2
NENUFAR 'roślina wodna o dużych pływających liściach i żółtych kwiatach'		p	4	9	0	4	-u/-a	M3
NIEFART <i>pot.</i> '= pech'		p	4	14	15	127	-u/-a	M3/M2

27 Wartości z korpusu zrównoważonego podano (w nawiasie) dla porównania z jednostką KONDON.

28 Jeden z przykładów mniej przekonujący z racji rymu: *Ja ostatnio zmieniałam redute ordona na redute kondona* (pisownia oryginalna).

29 Dominowały wystąpienia w znaczeniu 'nasienie' (z uwzględnieniem użyc metaforycznych: 'rzecz o podobnym kształcie, kolorze, wielkości...'; np. *oczy w migdał wycięte, oczy z migdału*) – odpowiednio (według kolejności kolumn): 36, 12, 3, 25; wyróżniały się jeszcze użycia w znaczeniu 'roślina' (6, 3, 0, 1); pozostałe wystąpienia w znaczeniu nieokreślonym (2, 1, 0, 1).

30 Jeden z przykładów mniej przekonujący z racji rymu: *Która dostała migdała, dostanie Michała*.

31 Tego znaczenia nie uwzględniał SAJW (przyjmowano tylko rodzaj m2).

32 Sześć użyc (cztery w podkorpusie zrównoważonym) w zwrotach typu *dać w migdał* ('uderzyć').

OMLECIEK 1. <i>pot.</i> ' = omlet'; 2. 'ciastko w kształcie omleta'		p	1	2	0	10	-u/-a	M3
OMLET 'smażona potrawa z mąki, mleka i ubitych jaj, o postaci placka'		z	2	25	3	41	-u/-a	M3/M2
PAINTBALL 'gra, w której uczestnicy strzelają do siebie nabojami wypełnionymi farbą'		z	24	23	1	40	-a/-u	M3/M2
PARCZ <sup>33</sup>	1. a) 'grzybicza lub bakteryjna choroba roślin'; b) 'objaw tej choroby'	p	36	0	28	6	-a	M2/M3
	2. <i>pot.</i> a) 'grzybicza skóry'; b) '[zwykle w lm.] zmiana na skórze wywołana tą chorobą'	p	0	0	14	4	-	M2/M3
PIRUET 1. <i>SPORT</i> 'figura w tańcu i łyżwiarstwie oraz w jeździe konnej'; 2. <i>pot. przen.</i> 'inny ruch obrotowy (ciała, pojazdu), zwłaszcza szybki i zręczny' <sup>34</sup>		p	3	47	23	97	-u/-a	M3/M2
PLATFUS <i>pot.</i> ' = płaskostopie'		p	17	0	25	2	-a	M2/M3
POLAROID 'aparat fotograficzny do szybkiego uzyskiwania zdjęć'		z	5	6	1	3	-u/-a	M3/M2
POST 'list elektroniczny zamieszczony na forum'		z	95	196 <sup>35</sup>	157	241	-u/-a	M3/M2

33 Nie uwzględniono pejoratywnych znaczeń przenośnych: 'złośliwe oprogramowanie', 'słaby film', '„znamię” żydowskości lub wprost 'Żyd' (w tym ostatnim wypadku i tak nie zgadzałby się rodzaj, odpowiednio do znaczenia – męskoosobowy).

34 Uwzględniono także użycia z dalej posuniętą metaforyzacją (w znaczeniu „akrobacji” intelektualnej czy logicznej).

35 Gdyby liczyć osobno wszystkie użycia frazy *temat postu* (w ramach tych samych portali), liczba wystąpień formy *postu* skoczyłaby do 1022 (!).

PULPET 'mielone mięso w formie kulki, = klops, klopsik'	p	1	0	5	4	-a	M2/M3
RAUSZ 'lekkie odurzenie lub podniecenie (m.in. pod wpływem alkoholu) <sup>36</sup>	p	3	31	3	21	-u/-a	M3/M2
SMS/sms 'wiadomość wysyłana przez telefon komórkowy'	z	452 <sup>37</sup>	9	558 <sup>38</sup>	567	-a/-u	M3/M2
STĘP 'najwolniejszy, miarowy chód konia'	p	12	1	12	4	-a/-u	M2/M3
SWEET 'odmiana jazzu wykonywana bez improwizacji, charakteryzująca się sentymentalnym brzmieniem'	p	0	1	1	0	-u	M2
ŚWIERK 'drzewo iglaste'	z	210	95	11	66	-a/-u	M3/M2
TOST 'opiekana kromka chleba'	z	1	8	15	9	-u/-a	M2/M3
TRÓJNIAK/TROJNIAK <sup>39</sup> 'miód pitny o określonych proporcjach (rozcieńczony wodą lub sokiem w stosunku 1:2)'	p	16 (15/1)	4 (2/2)	0	10 (4/6)	-a/-u	M3
WIS/WiS/viS/ViS 'dawny pistolet samopowtarzalny'	z (p)	8	0 (0 <sup>40</sup> )	7	0	-a	M2

36 Z NKP wyłania się wiele odmian *rauszu* ze względu na czynnik sprawczy (*rausz*: *adrenalinowy, amfetaminowy, intelektualny, kofeiny, konsumpcyjny, literacki, miłosny, morderczy, nacjonalistyczny, narkotykowy/narkotyczny, wiosenny*).

37 Z tego: *SMS-a* (zapis z łącznikiem bez spacji) – 369 wystąpień, *smsa* (zapis bez łącznika) – 78, *sms -a* (zapis z łącznikiem i poprzedzającą go spacją) – 5.

38 Z tego: *SMS-a* – 406 wystąpień, *smsa* – 149, *sms -a* – 3.

39 W nawiasach podano liczby wystąpień form z samogłoską tematyczną *ó* (przed ukośnikiem) i form z samogłoską tematyczną *o* (po ukośniku).

40 Gdy się okazało, że frekwencja dopełniacza na *-u* również w korpusie pełnym jest zerowa, wyjątkowo – ze względu na bardzo dużą liczbę wszystkich wystąpień segmentów *vis/ViS/VIS* i *visa/Visa/VISA* (odpowiednio: 5670 i 3563) – zrezygnowano ze zliczania stamtąd pozostałych form (form na *-a* w D.lp. oraz form obu typów rodzajowych, czyli na *-a* i na *-ø*, w B.lp.).

WALKOWER 'przyznanie zwyczajstwa w wyniku niestawienia się przeciwnika albo niedo- pełnienia przez niego wa- runków regulaminu'	z	30	27	30	128	-a/-u	M <sub>3</sub> /M <sub>2</sub>
WILCZOMLECZ 'roślina zielna lub drzewia- sta podobna do kaktusa, zawierająca w pędach sok mleczny, zwykle trujący, = ostromlecz'	p	15	6	1	7	-a/-u	M <sub>3</sub> /M <sub>2</sub>
ZIMOWIT 'roślina zielna'	p	24	4	0	6	-a/-u	M <sub>3</sub>

Na 62 prezentowane jednostki 43 przeszły pozytywnie weryfikację zarówno pod względem dwurodzajowości M<sub>2</sub>/M<sub>3</sub>, jak i oboczności -a/-u (niezacienione wiersze w tabeli), z tego: większość (26) – już w podkorpusie zrównoważonym (prymarnie wykorzystywanym), pozostałe (17) – dopiero w korpusie pełnym.

Dwie jednostki – GLOBUS i GNIOT – przeszły tę weryfikację połowicznie w innym jeszcze sensie (niż przez odwołanie się tylko do korpusu pełnego), to znaczy tylko w jednym z dwóch odrębnie ujętych i liczonych znaczeń. Zgodnie z przyjętym tutaj warunkiem gramatycznej spójności analizowanych jednostek należy odrębnie ujmować jednostki o pokrewnym znaczeniu, ale różniące się fleksyjnie bądź rodzajowo. Uzyskane dane korpusowe przemawiają jednak tylko za rozdzieleniem jednostek GLOBUS<sub>1</sub> i GLOBUS<sub>2</sub>, a nie GNIOT<sub>1</sub> i GNIOT<sub>2</sub>, gdyż deficyty frekwencji znaczenia GNIOT<sub>1</sub> nie pozwalają wyrokować o jego specyfice gramatycznej.

Ogólnie brak wystąpień którejs z czterech przeszukiwanych form oznacza, że nie potwierdza się oboczność -a/-u w D.lp. lub dwurodzajowość M<sub>2</sub>/M<sub>3</sub> danej jednostki (lub ani jedna, ani druga<sup>41</sup>). W tabeli takie jednostki notowano w zacienionym wierszu. Trudno jednak mówić tu wprost o sfałszykowaniu przyjętej za słownikami kwalifikacji gramatycznej, zwłaszcza w odniesieniu do rzeczowników o ogólnie niskiej frekwencji.

Na podstawie przeprowadzonych badań korpusowych nie można potwierdzić żadnego wyjątkowego związku -u (D.lp.) : M<sub>2</sub> (-u albo jako nieoboczna końcówka jednostek dwurodzajowych M<sub>2</sub>/M<sub>3</sub>, albo jako końcówka jednostek o niepoświadczonym rodzaju M<sub>3</sub>). Z racji minimalnej frekwencji potencjalnym tylko wyjątkiem w przebadanym zbiorze jednostek pozostaje rzeczownik SWEET: w NKJP ma formę na -u jako jedyne wystąpienie D.lp. i zarazem męskozwierzęcą formę B.lp. jako jedyne poświadczenie rodzajowe.

W modelującym opisie słownikowym, zwłaszcza z ambicjami do uprawiania jakiejś polityki językowej (w duchu normatywnym, prosystemowym czy nawet glottodydaktycznym),

41 W skrajnym wypadku nie było żadnego wystąpienia z wszystkich czterech uwzględnianych typów form (HALOBIONT).

prędzej można pominąć elementy (warianty) skrajnie mniejszościowe jednostek o ogólnie wyższej frekwencji. Największe dysproporcje liczby wystąpień zarówno między obydwoma końcówkami D.lp., jak i między oboma alternantami rodzajowymi w B.lp. są jednakowego typu, to znaczy z jednej strony *-a* dominuje nad *-u*: MIKST – 193 : 1, SMS/SMS – 452 : 9 ( $\approx 50,2 : 1$ ), GIMBUS – 63 : 2 ( $= 31,5 : 1$ ), DYNGUS/ŚMIGUS-DYNGUS – 27 : 1, BUS – 223 : 15 ( $\approx 14,9 : 1$ ), z drugiej strony M<sub>3</sub> dominuje nad M<sub>2</sub>: BEFSZTYK – 74 : 1, PAINTBALL – 40 : 1, MIGDAŁ ‘narząd ciała’ – 161 : 5 ( $= 32,2 : 1$ ), ŁYK – 742 : 25 ( $\approx 29,7 : 1$ ), FILET – 45 : 2 ( $= 22,5 : 1$ ), FANT – 22 : 1. Największe przewagi odwrotne, czyli *-u* nad *-a* oraz M<sub>2</sub> nad M<sub>3</sub>, są skromniejsze: PIRUET – 47 : 3 ( $\approx 15,7 : 1$ ), OMLET – 25 : 2 ( $= 12,5 : 1$ ), RAUSZ – 31 : 3 ( $\approx 10,3 : 1$ ); w B.lp.: PLATFUS – 25 : 2 ( $= 12,5$ ), MIKST – 40 : 5 ( $= 8$ ), ESEMES – 218 : 28 ( $\approx 7,8 : 1$ ), KONDON – 22 : 3 ( $\approx 7,3 : 1$ ).

Prezentując tego typu wyniki, trzeba oczywiście pamiętać o zróżnicowanych rangach frekwencyjnych badanych jednostek oraz o dwoistości źródła uzyskanych danych, czyli o rozdzwieku między wersją pełną i zrównoważoną NKJP.

Co ciekawe jednak, największe sumy frekwencji wszystkich czterech liczonych typów form (dopełniacz na *-a*, dopełniacz na *-u*, biernik M<sub>2</sub>, biernik M<sub>3</sub>) uzyskano z podkorpusu zrównoważonego: SMS/SMS – 1586 (odpowiednio: 452 + 9 + 558 + 567), ŁYK – 818 (26 + 25 + 25 + 742), POST – 689 (95 + 196 + 157 + 241), ŚWIERK – 382 (210 + 95 + 11 + 66), BUS – 370 (223 + 15 + 102 + 30), ESEMES – 286 (37 + 3 + 218 + 28)<sup>42</sup>. Wyniki uzyskane dla tych właśnie jednostek można uważać za statystycznie najbardziej wiarygodne (istotne).

Największe sumy frekwencji (w korpusie pełnym) jednostek „odrzuconych” były znacznie niższego rzędu: PARCH w znaczeniu botanicznym – łącznie 70 wystąpień form biernikowych i dopełniaczowych (lecz ani jednej formy D.lp. na *-u*)<sup>43</sup>, PLATFUS i KONDON – po 44, CIENKUSZ – 36, ZIMOWIT – 34. Jeżeli chodzi o duże (większe) frekwencje o zbliżonych wartościach, uderza równowaga rodzajowa (M<sub>2</sub> : M<sub>3</sub>) jednostek: SMS/SMS (558 : 567) i FART (74 : 73), ale więcej przykładów można podać na frekwencyjnie wyrównaną oboczność *-a/-u*: CHIP/CZIP (51 : 58), HIACYNT (18 : 20), FILET (15 : 12), ŁYK (26 : 25), PAINTBALL (24 : 23), WALKOWER (30 : 27).

Łatwiej oczywiście wskazywać frekwencyjne tendencje (przewagi i mniejszości) niż precyzyjnie określić liczby i proporcje wystąpień. Do żadnych z prezentowanych w tym artykule liczb, nie tylko bezwzględnych, nie należy się więc przywiązywać.

Mamy nadzieję, że przedstawione tu analizy i wyniki nie tylko zwiększą zainteresowanie jednostkami dwurodzajowymi M<sub>2</sub>/M<sub>3</sub> z obocznością końcówek D.lp., ale także stanowią będą jeden z przyczynków do dalszych refleksji i dyskusji nad zagadnieniami z szerokiego zakresu metodologii badań korpusowych (zwłaszcza na styku gramatyki i leksykologii) ze szczególnym uwzględnieniem ich dialektycznych powiązań z opracowaniami leksykograficznymi. Z jednej strony badania korpusowe są bowiem sterowane słownikami, ale z drugiej

42 Dopiero kolejny wynik w tym rankingu pochodzi z korpusu pełnego: MIGDAŁ ‘narząd ciała’ – 277 (106 + 5 + 5 + 161). Z sumą powyżej stu mamy jeszcze następujące jednostki (w nawiasie zaznaczono typ źródła: k – korpus pełny, z – (pod)korpus zrównoważony; nie podawano już – dostępnych w tabeli – liczb wystąpień poszczególnych typów form): WALKOWER – 215 (z), FART – 214 (z), CHIP/CZIP – 192 (z), KOKOS – 174 (p), PIRUET – 170 (p), NIEFART – 160 (p), MLECZ – 157 (p), GIMBUS – 112 (z), BEFSZTYK – 108 (z).

43 Łącznie w obu pokrewnych znaczeniach (z których każde miało dwa analogiczne podznaczenia: ‘choroba’ i ‘objaw choroby’) liczba wystąpień wyniosła 88 (36 + 0 + 42 + 10).

strony przyczyniają się do poprawienia opisów jednostek leksykalnych zastanych w słownikach. Niektóre z przyjętych tu założeń i proponowanych rozwiązań mogą się wszak wydawać kontrowersyjne, niedopracowane czy nazbyt uproszczone: porównywalność wyników zwichnięta rozdwojeniem korpusowego źródła danych (kompromis w korzystaniu z pełnej i zrównoważonej wersji NKJP) czy niejednolite podejście do wieloznaczności (wygodne w starciu z korpusowym żywiołem odchodzenie od ujęć polisemicznych na rzecz pojemnej formuły monosemicznej, ale też łączne zliczanie niektórych jednostek mimo ich wykazanej polisemii).

Konfrontacja kilkudziesięciu jednostek leksykalnych z korpusem wiąże się z taką porcją szczegółowej problematyki semantycznej (definicyjnej), która nie mogła się pomieścić w ramach jednego artykułu, tym bardziej o prymarnym ukierunkowaniu gramatycznym. Nie sposób jednak było tutaj uniknąć konkretnych rozstrzygnięć co do gramatyczno-semantycznego „spasowania” badanych jednostek (w świetle korpusu), choć pewne dodatkowe szczegóły mogły być odnotowywane już tylko w trybie glossarycznym. Zdaliśmy sobie przy tym sprawę, że niektóre jednostki (m.in. FANT) zasługiwałyby na odrębne studium, i to niezależnie od niejednoznacznych charakterystyk gramatycznych (dwurodzajowości).

W następnym artykule (Wojdak, w druku) powrócimy do *stricte* gramatycznych rozważań na temat sposobów ujęcia relacji między obocznością *-a/-u* a dwurodzajowością M2/M3, jako gotowe dane wykorzystując prezentowane tutaj frekwencje korpusowe (tylko jednostek pozytywnie zweryfikowanych zarówno pod względem oboczności końcówek D.lp., jak i alternacji rodzajowej w B.lp.).

## Bibliografia

- Bańko M. 2001: *Z pogranicza leksykografii i językoznawstwa*, Wydział Polonistyki Uniwersytetu Warszawskiego, Warszawa.
- Bańko M. 2002: *Wykłady z polskiej fleksji*, Wydawnictwo Naukowe PWN, Warszawa.
- Bańko M. 2018: *Korzyści i problemy wynikające z „nadmiaru” danych językowych i z wielości metod ich analizy*, „Biuletyn Polskiego Towarzystwa Językoznawczego” LXXIV, s. 23–30.
- EPLar: *Encyklopedia powszechna. Encyklopedyczny słownik języka polskiego od a do z. Uniwersalna encyklopedia od A do Z*, red. T. Michałowski, M.D. Krajewska, M. Gumkowski, Larousse Polska, Wrocław 2006.
- Górski R. 2012: *Zastosowanie korpusów w badaniu gramatyki*, [w:] A. Przepiórkowski, M. Bańko, R.L. Górski, B. Lewandowska-Tomaszczyk (red.), *Narodowy Korpus Języka Polskiego*, Wydawnictwo Naukowe PWN, Warszawa, s. 291–299.
- Górski R., Łaziński M. 2012: *Reprezentatywność i zrównoważenie korpusu*, [w:] A. Przepiórkowski, M. Bańko, R.L. Górski, B. Lewandowska-Tomaszczyk (red.), *Narodowy Korpus Języka Polskiego*, Wydawnictwo Naukowe PWN, Warszawa, s. 25–36.
- ISJP: *Inny słownik języka polskiego*, red. M. Bańko, t. 1–2, Wydawnictwo Naukowe PWN, Warszawa 2000 [w wersji elektronicznej jako *Słownik szkolny języka polskiego PWN*].
- NKJP: *Narodowy Korpus Języka Polskiego* (online: [www.nkjp.pl](http://www.nkjp.pl), dostęp: 1 marca 2020).
- PSWP: *Praktyczny słownik współczesnej polszczyzny*, t. 1–50, red. H. Zgólkowa, Kurpisz, Poznań 1994–2005.
- SAJW: *Słownik apelatywnych jednostek wielorodzajowych*, [w:] P. Wojdak, *Wielorodzajowość wśród rzeczowników współczesnej polszczyzny*, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin 2013.
- SGJP: Z. Saloni, M. Woliński, R. Wołosz, W. Gruszczyński, D. Skowrońska, *Słownik gramatyczny języka polskiego*, wyd. 3 online, Warszawa 2015 (online: <http://sgjp.pl>, dostęp: 1 marca 2020).
- USJP: *Uniwersalny słownik języka polskiego*, red. S. Dubisz, t. 1–6, Wydawnictwo Naukowe PWN, Warszawa 2006 (dokument elektroniczny na płycie).

- Wojdak P. 2013: *Wielorodzajowość wśród rzeczowników współczesnej polszczyzny*, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin.
- Wojdak P. 2020: *Oboczność -a/-u w D.lp. a dwurodzajowość M2/M3 – antysystemowe „skrzyżowanie”*, „Język Polski” C, z. 1, s. 47–56.
- Wojdak P. (w druku): *Dwurodzajowe jednostki rzeczownikowe M2/M3 z obocznością -a/-u w D.lp. Modelowanie relacji między końcówkami a alternantami rodzajowymi na podstawie frekwencji*, „Język Polski”.
- WSJPDun: *Język polski. Współczesny słownik języka polskiego*, red. B. Dunaj, Langenscheidt Polska, Warszawa 2009 (oraz: LexLand, Knurów) (dokument elektroniczny na płycie).
- WSJP PAN: *Wielki słownik języka polskiego PAN*, red. P. Żmigrodzki (online: <http://www.wsjp.pl/>, dostęp: 1 marca 2020).
- WSPP: *Wielki słownik poprawnej polszczyzny*, red. A. Markowski, Wydawnictwo Naukowe PWN, Warszawa 2008 (wyd. 1, 4. dodruk).
- WSWO: *Wielki słownik wyrazów obcych PWN*, red. M. Bańko, Wydawnictwo Naukowe PWN, Warszawa 2008.
- Zbróg P. 2014: Recenzja w przewodzie habilitacyjnym P. Wojdaka (<https://www.ck.gov.pl/promotion/id/259/type/l.html>, dostęp: 1 marca 2020).
- Żmigrodzki P. 2014: [rec.] *Piotr Wojdak, Wielorodzajowość wśród rzeczowników współczesnej polszczyzny*, „Język Polski” XCIV, z. 3, s. 281–284.

## Summary

---

### The *-a/-u* variance in GEN sing. and the M2/M3 double gender in the light of dictionary-driven corpus studies

Keywords: noun, masculine gender, double gender, inflectional ending, singular genitive, variance, corpus.

The article presents the results of corpus research (based on the PELCRA NKJP search engine) regarding the frequency of GEN. and ACC. sing. forms of noun units, which, in the light of at least some contemporary Polish dictionaries, are characterized by the M2/M3 double gender, as well as the *-a/-u* variance in GEN sing. Out of the investigated 62 units – considered with admission of formal variance and polysemy (which, however, do not affect their grammatical coherence) – 43 were positively verified both in terms of the M2/M3 double gender and the *-a/-u* variance in GEN sing., of which: 26 – already in the balanced sub-corpus, 17 – only in the full corpus. The presentation of the results was preceded by a discussion of the ‘origin’ of the initial list of the double gender units with *-a/-u* variance and the rules for counting their frequency in NKJP. The article encourages further reflection and discussion of the issues within the wide range of corpus research methodologies, especially at the intersection of grammar and lexicology, and with particular emphasis on their connections with lexicographic sources (but also intentions). The author is aware that some of the assumptions or solutions adopted in the article (split of the corpus data source, a heterogeneous approach to polysemy) may be subject to controversy, and the semantic issues related to the examined units could not be exhausted.