

Uniwersytet Łódzki
Wydział Filologiczny
Instytut Filologii Polskiej i Logopedii

Rafał Maćkowiak

Socjolekt polskich graczy wideo

Dysertacja napisana pod kierunkiem
dr hab. prof. UŁ Edyty Pałuszyńskiej
w Zakładzie Lingwistyki Stosowanej i Kulturowej

Łódź 2018

SPIS TREŚCI

WSTĘP.....	5
------------	---

CZĘŚĆ TEORETYCZNA

1. ŚRODOWISKO POLSKICH GRACZY WIDEO.....	9
1.1. CHARAKTERYSTYKA BRANŻY GIER WIDEO I ŚRODOWISKA GRACZY WIDEO.....	9
1.2. WEWNĘTRZNY PODZIAŁ ŚRODOWISKA POLSKICH GRACZY WIDEO.....	12
1.3. BADANIA GROZNAWCZE W POLSCE.....	14
2. BADANIA NAD ZRÓŻNICOWANIEM JĘZYKOWYM POLSZCZYZNY.....	18
2.1. POCZĄTEK BADAŃ W RAMACH DIALEKTOLOGII.....	18
2.2. BADANIA SOCJOLINGWISTYCZNE.....	22
2.3. KLASYFIKACJE ODMIAN JĘZYKOWYCH POLSZCZYZNY.....	33
3. BADANIA ILOŚCIOWE W POLSCE.....	42
4. USTALENIA TERMINOLOGICZNE.....	59
4.1. JĘZYK.....	59
4.2. SOCJOLEKT.....	62
4.3. WYRAZ, LEKSEM I JEDNOSTKA LEKSYKALNA.....	63
4.4. POLE LEKSYKALNE.....	70

CZĘŚĆ METODOLOGICZNA

5. CHARAKTERYSTYKA MATERIAŁU.....	73
5.1. KRYTERIA POZYSKANIA MATERIAŁU.....	73
5.2. ZAKRES POZYSKANEGO DO BADAŃ MATERIAŁU.....	75
5.3. SKRÓTOWE NAZWY POSZCZEGÓLNYCH CZĘŚCI KORPUSU.....	77
6. METODOLOGIA POZYSKIWANIA I PRACY Z MATERIAŁEM.....	78
ETAP 0. UWAGI TECHNICZNE.....	78
ETAP 1. POZYSKIWANIE EDYTOWALNYCH MATERIAŁÓW TEKSTOWYCH Z CZASOPISM DRUKOWANYCH.....	80
PRACE PRZYGOTOWAWCZE.....	80
PRACA Z PROGRAMEM ABBYY FINEREADER 12 PROFESSIONAL EDITION.....	81
GRUPOWANIE PLIKÓW MIESIĘCZNYCH W KWARTAŁY.....	87
ZMNIEJSZENIE WAGI/ROZMIARU PLIKÓW.....	88
ETAP 2. POZYSKIWANIE EDYTOWALNYCH MATERIAŁÓW TEKSTOWYCH Z INTERNETU.....	90
PRACA Z PROGRAMEM HTRACK WEBSITE COPIER.....	90
PORZĄDKOWANIE ZGROMADZONEGO MATERIAŁU.....	93
ŁĄCZENIE MAŁYCH PLIKÓW W WIĘKSZE – TWORZENIE WERSJI FINALNYCH.....	97
ETAP 3. PRZYGOTOWANIE MATERIAŁU LEKSYKALNEGO DO PRACY Z KORPUSAMI TEKSTOWYMI.....	98
ETAP 4. PRACA Z KORPUSAMI TEKSTÓW.....	103
SPÓSÓB WYSZUKIWANIA WYRAZÓW.....	104

OPRACOWANIE FORMUŁ WYSZUKIWANIA	105
POŁĄCZENIE FORMUŁ WYSZUKIWANIA	106
PROBLEM PIERWSZY – ZNAKI DIAKRYTYCZNE	108
PROBLEM DRUGI – RODZINY WYRAZÓW	110
PROBLEM TRZECI – KRÓTKIE WYRAZY	120
PROBLEM CZWARTY – WIELOZNACZNOŚĆ	122
ETAP 5. STWORZENIE PÓL LEKSYKALNYCH	125

CZĘŚĆ ANALITYCZNA

7. POLA LEKSYKALNE A ICH ZŁOŻONOŚĆ	135
8. KORPUSY TEKSTOWE A LICZBA WSZYSTKICH WYRAZÓW GRAFICZNYCH	138
9. WYRAZY GRAFICZNE ŚRODOWISKA GRACZY A PLATFORMY BADAWCZE	144
10. WYRAZY GRAFICZNE ŚRODOWISKA GRACZY A PODZIAŁ KWARTALNY	149
11. WYRAZY GRAFICZNE ŚRODOWISKA GRACZY WIDEO A WSZYSTKIE WYRAZY GRAFICZNE KORPUSÓW	152
12. WYRAZY GRAFICZNE A CZĘŚCI MOWY	155
13. SYSTEM 1.0	157
13.1. WPROWADZENIE DO SYSTEMU 1.0	157
13.2. ZASADY SYSTEMU 1.0	163
13.3. WYNIKI SYSTEMU 1.0	167
14. ZŁOŻONOŚĆ A SYSTEM 1.0	170
15. POLA LEKSYKALNE A SUMY WYRAZÓW GRAFICZNYCH	180
15.1. ZESTAWIENIE NAJBOGATSZYCH I NAJUBOŻSZYCH W WYRAZY GRAFICZNE PÓL LEKSYKALNYCH – ANALIZA POSZCZEGÓLNYCH CZĘŚCI KORPUSU PLATFORMY ZIELONEJ	180
15.2. SUMA NAJBOGATSZYCH I NAJUBOŻSZYCH W WYRAZY GRAFICZNE PÓL LEKSYKALNYCH – ANALIZA CAŁEGO KORPUSU PLATFORMY ZIELONEJ	185
15.3. ZESTAWIENIE NAJBOGATSZYCH I NAJUBOŻSZYCH W WYRAZY GRAFICZNE PÓL LEKSYKALNYCH – ANALIZA Z UWZGLĘDNIENIEM PODZIAŁU NA KORPUSY KONSOLOWCÓW ORAZ KOMPUTEROWCÓW PLATFORMY ZIELONEJ	188
15.4. ZESTAWIENIE NAJBOGATSZYCH I NAJUBOŻSZYCH W WYRAZY GRAFICZNE PÓL LEKSYKALNYCH – ANALIZA Z UWZGLĘDNIENIEM PODZIAŁU NA KORPUSY ZAWIERAJĄCE TEKSTY CZASOPISM, NEWSÓW ORAZ WPISÓW NA FORACH INTERNETOWYCH PLATFORMY ZIELONEJ I ŻÓŁTEJ	190
15.5. ZESTAWIENIE NAJBOGATSZYCH I NAJUBOŻSZYCH W WYRAZY GRAFICZNE PÓL LEKSYKALNYCH – ANALIZA POSZCZEGÓLNYCH CZĘŚCI KORPUSU PLATFORMY CZERWONEJ	193
15.6. SUMA NAJBOGATSZYCH I NAJUBOŻSZYCH W WYRAZY GRAFICZNE PÓL LEKSYKALNYCH – ANALIZA CAŁEGO KORPUSU PLATFORMY CZERWONEJ	196
15.7. ZESTAWIENIE NAJBOGATSZYCH I NAJUBOŻSZYCH W WYRAZY GRAFICZNE PÓL LEKSYKALNYCH – ANALIZA Z UWZGLĘDNIENIEM PODZIAŁU NA KORPUSY KONSOLOWCÓW ORAZ KOMPUTEROWCÓW PLATFORMY CZERWONEJ	200
15.8. ZESTAWIENIE NAJBOGATSZYCH I NAJUBOŻSZYCH W WYRAZY GRAFICZNE PÓL LEKSYKALNYCH – ANALIZA Z UWZGLĘDNIENIEM PODZIAŁU NA KORPUSY ZAWIERAJĄCE TEKSTY CZASOPISM ORAZ NEWSÓW PLATFORMY CZERWONEJ	202

16. ZŁOŻONOŚĆ A SUMA WYRAZÓW GRAFICZNYCH W CAŁYM KORPUSIE	205
17. POLA WYRAZOWE A CZĘSTOTLIWOŚĆ WYSTĄPIEŃ WYRAZÓW GRAFICZNYCH	207
17.1. ANALIZA WYKRESÓW PLATFORMY ZIELONEJ I CZERWONEJ W ODNIESIENIU DO PÓL NAJBOGATSZYCH W WYRAZY GRAFICZNE	209
17.2. ANALIZA WYKRESÓW PLATFORMY ZIELONEJ I CZERWONEJ W ODNIESIENIU DO PÓL NAJUBOŻSZYCH W WYRAZY GRAFICZNE	219
17.3. ANALIZA WYBRANYCH WYKRESÓW PLATFORMY ZIELONEJ I CZERWONEJ W ODNIESIENIU DO PÓL, KTÓRE OSIĄGAJĄ ŚREDNIE WARTOŚCI NASYCENIA WYRAZAMI GRAFICZNYMI (NIE SĄ ANI NAJUBOŻSZE, ANI NAJBOGATSZE)	223
17.4. WYBRANE WYKRESY WSZYSTKICH PLATFORM W ODNIESIENIU DO KORPUSÓW KONSOLOWCÓW ORAZ KOMPUTEROWCÓW	227
17.5. WYBRANE WYKRESY WSZYSTKICH PLATFORM W ODNIESIENIU DO KORPUSÓW CZASOPISM ORAZ NEWSÓW	230
18. SYSTEM 2.0	234
18.1. SYSTEM 2.0 W ODNIESIENIU DO SUMY WYRAZÓW GRAFICZNYCH POSZCZEGÓLNYCH PÓL W CAŁYM KORPUSIE – PLATFORMA ZIELONA I CZERWONA	234
18.2. SYSTEM 2.0 NA TLE INNYCH ZESTAWIEŃ I POŁĄCZEŃ POSZCZEGÓLNYCH CZĘŚCI KORPUSU – WSZYSTKIE PLATFORMY	248
ZAKOŃCZENIE	252
BIBLIOGRAFIA	260

WSTĘP

Steve Jobs, założyciel firmy Apple, twierdził, że przyszłość szeroko rozumianej elektroniki tkwi w połączeniu technologii z humanistyką. Podobnie twierdził jego mentor – Edwin Land (założyciel Polaroida – firmy produkującej aparaty fotograficzne), który w jednym z wywiadów podkreślił znaczenie ludzi działających na styku humanistyki i nauk ścisłych [Isaacson 2011]. Zdanie to stało się jedną z myśli przewodnich życia Jobsa, co znalazło symboliczne odzwierciedlenie na rzekomo realnie istniejącym zdjęciu skrzyżowania ulic Technologii i Sztuk Wyzwolonych. Przedstawiało ono drogowskazy tych ulic i często pojawiało się podczas prezentacji nowych produktów firmy.



Steve Jobs na tle zdjęcia, które przedstawia skrzyżowanie ulic Technologii i Sztuk Wyzwolonych
Źródło: www.pl.pinterest.com/pin/83457399314940903/visual-search/?x=16&y=12&w=530&h=397

Sadzę, że przyszłość humanistyki także jest usytuowana na tym samym skrzyżowaniu. Bo czy można dzisiaj uprawiać badania w zakresie humanistyki bez korzystania z pomocy jakiegokolwiek technologii, która coraz częściej wkracza do naszej codzienności, a dla wielu osób jest jej nieodzownym elementem? Uważam, że to pytanie jest już dzisiaj retoryczne.

Moja praca także ulokowana jest na styku dwóch interdyscyplinarnych dziedzin: socjolingwistyki i językoznawstwa komputerowego. Socjolingwistyka to połączenie lingwistyki i socjologii, a językoznawstwo komputerowe to symbioza językoznawstwa i informatyki. Można nawet pokusić się o stwierdzenie, że jest to praca multidyscyplinarna.

Dysertacja zbudowana jest z trzech głównych części: teoretycznej, metodologicznej oraz analitycznej.

W pierwszej części pracy – teoretycznej, opisuję, jak wyglądał rozwój branży gier wideo oraz pokrótce analizuję stan polskiego środowiska graczy wideo, następnie przedstawiam rozwój dwóch dziedzin naukowych, które stanowią trzon tej pracy, to jest socjolingwistyki i językoznawstwa korpusowego. Opis pierwszej dziedziny potraktowałem możliwie szeroko, czyli pokazałem, jak ewoluowało zainteresowanie polskich badaczy takim tematem, jak zróżnicowanie językowe polszczyzny. Zacząłem od dialektologii, bo w niej, moim zdaniem, miały początek badania nad wariantami językowymi polszczyzny, a następnie stopniowo zbliżałem się do współczesnej socjolingwistyki. W części tej przyjąłem porządek chronologiczny (z nielicznymi odstępstwami od tego), aby lepiej zobrazować rozwój i modyfikacje pewnych poglądów badawczych. Na koniec przedstawię i omówię wybrane terminy, których będę używał często w tej pracy.

W drugiej części pracy – metodologicznej, bardzo szczegółowo pokazałem, jak tworzyłem korpusy tekstów oraz, jak z nimi pracowałem. Część ta zawiera dokładne opisy postępowania, które doprowadzić mają do pozyskania surowego materiału badawczego do dalszej analizy ilościowej. W części tej kierowałem się przede wszystkim przydatnością, ponieważ chciałem, aby opisy przedstawione w tej pracy posłużyły innym badaczom, którzy szukają prostego sposobu na stworzenie korpusu i pracę na materiale z niego pozyskanym.

W części trzeciej – analitycznej, pokazałem wyniki moich badań. Jest to praca ilościowa, więc umieściłem w tej części wiele tabel, zestawień oraz wykresów, które pokazują, jak zebrany materiał prezentuje się z uwzględnieniem pewnych obostrzeń wynikających z badań ilościowych.

Warto jeszcze zaznaczyć, że do pracy dodany został obszerny aneks, w którym zawarte zostały definicje analizowanych jednostek leksykalnych (wraz z przykładami użycia, w których zachowano oryginalną pisownię) oraz różnego rodzaju zestawienia tabelaryczne i wykresy, do których będę odwoływał się w dalszej części pracy.

Dysertacja ma na celu pokazanie, że istnieje wewnętrzne zróżnicowanie leksyki na taką, która związana jest tylko z pewną grupą graczy wideo oraz taką, która istnieje ponad wewnętrznymi podziałami środowiska graczy wideo, czyli jest wspólna wszystkim graczom.

Zakładam, że wyrazy środowiska graczy tworzą większe zbiory, czyli pola leksykalne. Pola te różnią się pod względem struktury i liczebności elementów. Dzięki badaniom

korpusowym mogą także wziąć pod uwagę dodatkowy parametr, a mianowicie częstotliwość występowania (reprezentujących leksemy) wyrazów graficznych w tekstach.

Wszystkie powyższe czynniki różnicują poziom ważności pól, a tym samym jednostek leksykalnych, które są w nich zawarte, co zostanie wykazane w części analitycznej.

Praca ma potwierdzić lub obalić powyższe przypuszczenia oraz dostarczyć odpowiedzi na inne pytania, które wystąpią w czasie analizy zebranego do badań materiału.

Ponadto celem dodatkowym tej pracy jest opracowanie i zaprezentowanie prostej, możliwej do powtórzenia przez innych badaczy, metody pozyskania tekstów do korpusów oraz przedstawienie, w jaki sposób można z nimi pracować.

Zanim jednak przejdę do części teoretycznej i przedstawię rozwój branży gier wideo, to chciałbym wyjaśnić nietypowe połączenie wyrazowe, które wystąpiło już w tytule dysertacji, czyli *gracz wideo*, z którym w literaturze zarówno naukowej, popularnonaukowej czy beletrystycznej dotychczas nie zetknąłem się. Zostało ono przeze mnie stworzone z konieczności nazwania całej grupy graczy, czyli ludzi grających w gry wideo, niezależnie od używanego sprzętu. Wyrażenie *gracz wideo* łączy zakres dotychczasowych i ustabilizowanych terminów *gracz komputerowy*¹ i *konsolowy*² oraz dodatkowo eksponuje nie tyle sam sprzęt, na którym się gra, ile rodzaj gry (różny np. od gier planszowych czy w karty). Człon podrzędny/określający (*wideo*) jest doprecyzowaniem obiektu czynności. Funkcjonuje on już w połączeniach takich, jak: *gry wideo*, *kamera wideo* czy *kaseta wideo*.

¹ Gracz, który gra na komputerze.

² Gracz, który gra na jednej konsoli/ kilku konsolach. Najczęściej posiadają wiele konsol. *Konsola* to „elektroniczne urządzenie multimedialne umożliwiające granie w gry wideo, a także – w niektórych modelach – słuchanie muzyki, oglądanie filmów, przeglądanie internetu” [www.sjp.pwn.pl/slowniki/konsoli.html, dostęp: 6.04.2017].

Forma w dopełniaczu liczby mnogiej – *konsol* wielu osobom może wydać się błędna, gdyż także słowniki proponują, by używać – *konsoli*. W środowisku graczy bardziej rozpowszechniona jest jednak forma *konsol*, czyli wersja z końcówką \emptyset , dlatego też w swojej pracy zdecydowałem się na tę formę. Dopuszcza ją także Mirosław Bańko. Na zadane mu w poradni językowej pytanie, które brzmiało następująco: „Konsola – jak brzmi dopełniacz liczby mnogiej? Słowniki twierdzą, że *konsoli*. Wydaje mi się, że bardziej czytelna jest forma *konsol*” odpowiedział tak: „Na kilkadziesiąt wystąpień słowa *konsoli* w Korpusie Języka Polskiego PWN tylko jedno reprezentuje liczbę mnogą, pozostałe zaś są formami liczby pojedynczej – *tej konsoli*. Korpus zawiera też kilka przykładow na dopełniacz liczby mnogiej – *tych konsol*. Tak więc forma zgodna z Pani wycuciem językowym dominuje w tekstach współczesnej polszczyzny. Podaje ją też (implicite, bo poprzez tabelę tworzenia form) *Inny słownik języka polskiego* PWN. Szkoda, że inne słowniki w tym względzie pobiły, podając formę rzadszą jako jedyną” [http://sjp.pwn.pl/poradnia/haslo/konsola;3437.html, dostęp: 5.06.2017].

CZEŚĆ

TEORETYCZNA

1. ŚRODOWISKO POLSKICH GRACZY WIDEO

1.1. CHARAKTERYSTYKA BRANŻY GIER WIDEO I ŚRODOWISKA GRACZY WIDEO

Branża gier wideo to w dzisiejszych czasach jedna z najszybciej rozwijających się gałęzi przemysłu rozrywkowego. Korporacje takie, jak: Microsoft, Sony czy Nintendo inwestują coraz większe środki pieniężne w produkcję sprzętu do gier (np. komputerów, konsol, tabletów) oraz w produkcję gier na różne platformy. Początki branży gier wideo sięgają lat 50-tych XX wieku. Pierwsza graficzna gra wideo powstała w 1952 roku, a było nią kółko i krzyżyk. Gra stworzona została na maszynę o nazwie Electronic Delay Storage Automatic Calculator (EDSAC). Była ona dostępna tylko dla niewielkiej grupy użytkowników [Mańkowski 2010: 11]. Inaczej natomiast było z pierwszą grą komercyjną, którą była Spacewar. W roku 1962 gra dodawana była do komputera PDP-1, którego łącznie sprzedano około 50 egzemplarzy [Mańkowski 2010: 13–14]. Natomiast za pierwszą konsolę do gier uważa się Magnavox Odyssey, która była jednocześnie też pierwszą konsolą wprowadzoną do masowej produkcji, a miało to miejsce w roku 1972. W Polsce pierwsza gra wideo powstała na komputer Odra 1003 (wyprodukowany on został we Wrocławiu) w 1962 roku, była nią Marienbad. Gra polegała na tym, że generowała cztery rzędy zapalek, a gracz na przemian z komputerem wyciągał po jednej zapalce z wybranego rzędu. Przegrywał ten, komu została ostatnia z zapalek. Gra została tak zaprojektowana, że zawsze wygrywał komputer.

Tak wyglądały początki branży gier wideo, której prawdziwy rozkwit nastąpił dopiero w 1983 roku, kiedy to swoją premierę miała konsola Nintendo Entertainment System. Na tę konsolę ukazała się najbardziej znana i popularna gra na świecie, czyli Super Mario Bros. Obecnie świat gier wideo zdominowany jest przez konsole, które wydawane są przez Microsoft, Nintendo i Sony. Każda z tych firm co kilka lat tworzy nową platformę do gier. Sony do tej pory wydało takie konsole, jak: PlayStation, PlayStation 2, PlayStation 3 i PlayStation 4. Microsoft stworzył konsole: Xbox, Xbox 360 i Xbox One. Natomiast Nintendo, które najdłużej istnieje w branży gier wideo, wyprodukowało takie konsole, jak: Nintendo Entertainment System, Super Nintendo Entertainment System, Nintendo 64, GameCube, Wii, Wii U i Nintendo Switch. Oczywiście wymieniałem wyłącznie konsole tych firm, które do dnia dzisiejszego zajmują się branżą gier wideo. Poza nimi konsole produkowała także np. firma Sega, która swoją ostatnią konsolę wydała w roku 1998, po czym zajęła się wyłącznie

produkcją gier na konkurencyjne platformy. Natomiast w nieco dalszej przeszłości konsole wydawały jeszcze takie firmy, jak np.: Bandai (konsola WonderSwan Color), Atari (konsola Atari 2600), czy też założona przez polskiego emigranta, łodzianina – Jacka Tramiela firma Commodore (konsola Amiga CD32). Część z tych firm zbankrutowała, część natomiast, podobnie jak Sega, przestała produkować konsole i zajęła się wyłącznie wydawaniem gier.

W Polsce duże zainteresowanie grami konsolowymi zaczęło się wraz z premierą konsoli Pegasus, która była podrobioną wersją platformy Nintendo Entertainment System. Kolejnym punktem przełomowym dla polskich graczy konsolowych było wyprodukowanie przez firmę Sony platformy do gier Playstation, która w Polsce stała się synonimem określenia konsola (gdy ktoś mówił, że posiada w domu konsolę, to wiadomo było, że chodzi o sprzęt Sony).

Zainteresowanie grami wideo w naszym kraju z roku na rok jest coraz większe. Widać to na przykład po wzroście sprzedaży czasopism związanych z grami (takimi jak „CD-Action” czy „PSX Extreme”) oraz po wynikach przeprowadzonej w 2014 r. ankiety „Jestem graczem”, w której wzięło udział aż 79 543 osób. Część z tych osób jest graczami *hardkorowymi*³, a część *casualowymi*⁴, jednak wszyscy oni, w mniejszym lub większym stopniu, posługują się charakterystycznym dla środowiska graczy socjolektem. Wiele z tych osób wychowało się i dorastało w świecie, w którym komputery nie były czymś nowym, a dostęp do internetu był lub stawał się powszechny. Upowszechnienie internetu ułatwiło komunikację między graczami i było czynnikiem sprzyjającym rozprzestrzenianiu się specjalistycznych określeń związanych z grami. Zauważyć można, że osoby te przenoszą pewne słownictwo, do tej pory zarezerwowane dla rozmów między graczami, do swojego języka potocznego, czy też zawodowego (na przykład jedną z relacji z zawodów sportowych Poznań Open 2014 autor zatytułował następująco: „W Poznaniu wróciłem na swój *level*”. Użył tutaj określenia *level*, wcześniej zarezerwowanego dla środowiska graczy komputerowych/konsolowych⁵).

Także inne elementy związane z grami wideo (np. znani bohaterowie, muzyka czy styl ubierania się) coraz częściej spotkać można poza ekranami telewizorów, monitorów czy też telefonów komórkowych. Tematyka związana z grami coraz częściej towarzyszy nam w życiu codziennym (w tzw. realu), efektem tego są na przykład coraz liczniej powstające „knajpy dla graczy”⁶. Istnieją one już w takich miastach jak: Poznań, Łódź, Kraków, Warszawa i Wrocław. „Knajpa dla graczy” jest połączeniem salonu gier oraz pubu. Osoba odwiedzająca taki lokal

³ Gracze nałogowi, którzy grają bardzo dużo i w tytuły bardzo wymagające.

⁴ Gracze, którzy grają w tytuły łatwe i niewymagające doświadczenia w tej czynności.

⁵ Zdanie to odnotowano na stronie: www.zwoleja.blog.onet.pl/2014/07/22/w-poznaniu-wrocilem-na-swoj-level-kamil-majchrzak [dostęp: 8.04.2017].

⁶ Zdecydowałem się zostawić tę potoczną nazwę, ponieważ jest ona popularniejsza w środowisku graczy.

może z przyjaciółmi zagrać w grę na konsoli oraz skosztować alkoholu. „To coś dla osób, które czują się nieswojo w rzekomo popularnych klubach, gdzie liczy się hajs, wypolerowane buty i dyskotekowe hity. Oto najnowsza moda, która radzi sobie coraz lepiej w Polsce” [Kujawa 2014a]. W tego typu lokalach bardzo wyraźnie słychać posługiwanie się charakterystycznymi dla graczy określeniami. Osoba, która ich nie zna, nie zrozumie treści danego komunikatu. Wizyta w takich miejscach jeszcze dobitniej pokazuje, jak liczne jest środowisko graczy i jak rozpowszechniona jest ich leksyka.

W ostatnim czasie w Polsce zaobserwować można także coraz większe zainteresowanie e-sportem, czyli sportem elektronicznym. Gracze, podobnie jak w tradycyjnym sporcie (w piłce nożnej, skokach narciarskich czy pływaniu), rywalizują ze sobą o tytuł najlepszego zawodnika. Jedni z nich wyspecjalizowali się w strzelaninach (gra akcji, w której gracz eliminuje przeciwników przy pomocy broni palnej) inni na przykład w bijatykach (gra, w której gracz eliminuje przeciwników walcząc wręcz lub przy pomocy broni białej), jedni grają na komputerach, natomiast inni na konsolach, ale wszyscy oni nazywają siebie e-sportowcami (lub e-atletami). Często ze zdobyciem tytułu najlepszego zawodnika wiąże się dodatkowo otrzymanie odpowiedniej nagrody, zazwyczaj pieniężnej.

„W Polsce z profesjonalnego gamingu może utrzymać się około 50 osób” [Kujawa 2014b]. Nie jest to duża liczba, gdyż w naszym kraju, w przeciwieństwie do krajów zachodnich, e-sport jest jeszcze czymś nowym.

Jak w każdym sporcie, także i w tym organizowane są zawody, które przyciągają niekiedy większą widownię niż tradycyjne zawody sportowe, a ludzie oglądający mecze e-sportowe muszą doskonale znać charakterystyczne dla graczy słownictwo, chociażby po to, by zrozumieć, co mówią komentatorzy podczas takich wydarzeń.

O dużym zainteresowaniu sportem elektronicznym świadczy także to, że w Polsce, a konkretniej w kędzierzyńsko-kozielskim Zespole Szkół Technicznych i Ogólnokształcących powstała klasa o profilu e-sportowym, w której kształceni są przyszli e-sportowcy.

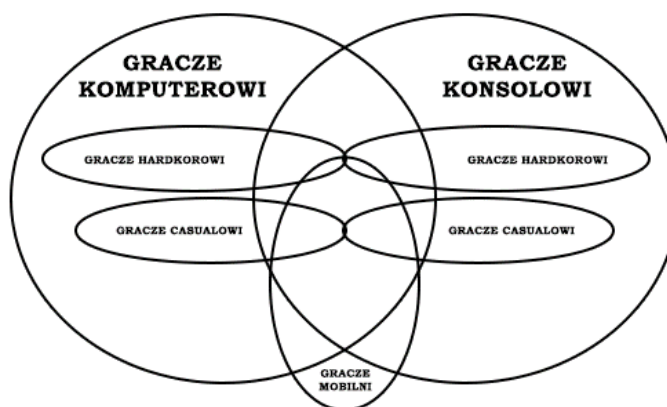
Coraz większe zainteresowanie grami w naszym kraju może wynikać także z tego, że polskie firmy zajmujące się produkcją gier powoli zaczynają być doceniane na świecie. Przykładem może być studio CDProject Red, które odpowiedzialne jest za wyprodukowanie serii gier o Wiedźminie (na tę chwilę powstały trzy części, a w planach są kolejne). Seria ta została bardzo dobrze przyjęta przez większość czasopism branżowych, ale co ważniejsze – także przez graczy. Pierwsza część tej serii została oceniona przez portal Gamerankings na 81,71% (najgorszą oceną na tym portalu może być 0%, natomiast najlepszą 100%), a przez portal Metacritic na 81 punktów (najgorsza oceną na tym portalu może być 0 punktów,

natomiast najlepszą 100 punktów). Drugą część gry portal Gamerankings ocenił na 87,97 %, a portal Metacritic na 88 punktów, natomiast część trzecia oceniona została najlepiej, bo Gamerankings ocenił ją na 92%, a Metacritic na 93 punkty. Cała seria sprzedała się w olbrzymim nakładzie 25 mln sztuk.

1.2. WEWNĘTRZNY PODZIAŁ ŚRODOWISKA POLSKICH GRACZY WIDEO

Środowisko polskich graczy wideo nie stanowi jednolitego pod względem zainteresowań zbioru ludzi, lecz jest to bardzo duża grupa osób o różnorodnych zamiłowaniach. Umieszczona obok grafika prezentuje uproszczony obraz tego środowiska (patrz grafika 1.). Zilustrowałem podział graczy ze względu na sprzęt, na jakim grają oraz ze względu na to, ile czasu poświęcają na tę czynność

Grafika 1. Podział środowiska graczy wideo.



Źródło: opracowanie własne

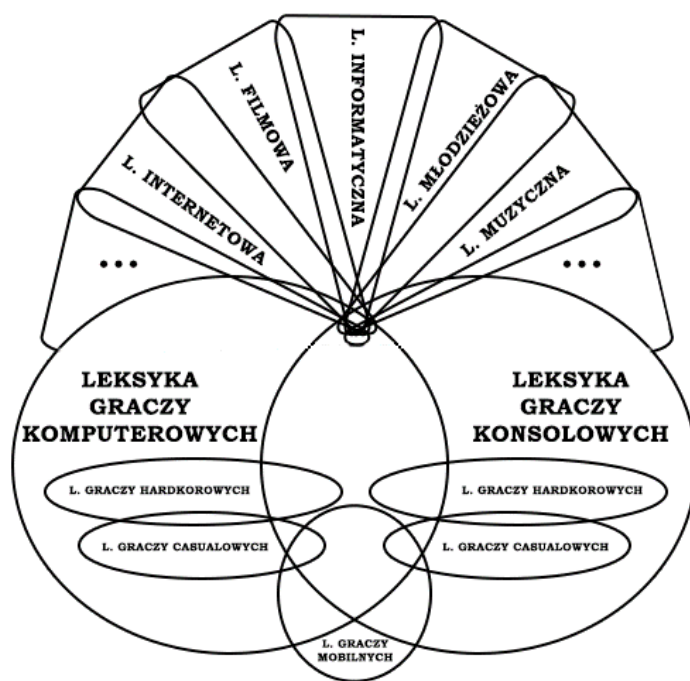
(grają okazjonalnie lub nałogowo). Trzon społeczności stanowią *gracze komputerowi* i *konsolowi*. Poza nimi wyróżniłem mniejszy, ale ciągle rozrastający się zbiór *graczy mobilnych*. Grają oni na urządzeniach przenośnych typu smartfon czy tablet. Dodatkowo w zbiorze *graczy komputerowych* i *konsolowych* wyszczególnić można *graczy hardkorowych* (gracze tzw. nałogowi, którzy potrafią grać nawet po kilkanaście godzin dziennie) i *casualowych* (gracze tzw. niedzielni, czyli grający rzadko). Zbioru *graczy mobilnych* nie można poddać takiemu zróżnicowaniu, ponieważ urządzenia mobilne są stosunkowo nowymi platformami do gier i trudno wyróżnić tam *graczy hardkorowych*. Ilustracja pokazuje także, że istnieją gracze, którzy grają wyłącznie na jednym rodzaju sprzętu (np. na *konsolach*), ale są także osoby, które grają na wszystkich możliwych platformach do gier, znajdują się one w centrum prezentowanej grafiki, w miejscu połączenia wszystkich zbiorów. Nie bez znaczenia jest także wielkość elipsy symbolizującej *graczy mobilnych* oraz rozmieszczenie, czyli to, że jej większa część znajduje się w zasięgu innych zbiorów (*graczy konsolowych* i *komputerowych*). *Gracze mobilni* to w dużej mierze osoby, które grają już na innym sprzęcie, a gry na tabletach czy telefonach komórkowych traktują jako dodatek, np. w sytuacji podróży, gdy nie mogą zagrać na swoim ulubionym sprzęcie.

Taki podział środowiska wymusza pewne zróżnicowanie używanej leksyki. Aby zaprezentować teoretyczny podział słownictwa graczy, powyższy wykres trzeba nieco zmodyfikować. Na poniższym wykresie (grafika 2.) pokazane zostały zbiory leksykalne, które bezpośrednio wynikają z podziału środowiska graczy (patrz grafika 1.).

O ile prawdziwość podziału środowiska graczy wideo (patrz grafika 1.) można potwierdzić, ponieważ ma on odzwierciedlenie w rzeczywistości i wynika z obserwacji, o tyle podział słownictwa tego środowiska jest tylko hipotezą badawczą, którą udowodnić będzie można dopiero po przeprowadzeniu dokładnej analizy językowej.

Na załączonej obok grafice zobrazowano, że w zbiorze leksyki graczy wideo wyróżnić można trzy główne grupy: leksykę graczy komputerowych, leksykę graczy konsolowych i leksykę graczy mobilnych. Wszystkie trzy zbiory częściowo nachodzą na siebie, ilustrując, że istnieje pewien zbiór leksyki znany wszystkim graczom. Zróżnicowanie leksykalne jest konsekwencją wyboru różnych platform do gier. Słownictwo zmuszone było dostosować się do zupełnie innych realiów. Przykładowo

Grafika 2. Leksyka środowiska graczy wideo.



Źródło: opracowanie własne

dla *komputerowców* (*graczy komputerowych*) i *graczy mobilnych* ważne będą minimalne wymagania sprzętowe, które są potrzebne, aby gry na danym sprzęcie uruchomiły się. Dla *graczy mobilnych* wymagania mają mniejsze znaczenie, bo technologicznie telefony/tablety są do siebie bardziej podobne niż komputery, co nie zmienia faktu, że także się liczą. Natomiast dla *konsolowców* (*graczy konsolowych*) ten aspekt nie będzie miał znaczenia, gdyż konsole to urządzenia dostosowane do tego typu aktywności, czyli do grania (w odróżnieniu od konsol każdy komputer jest nieco inaczej zbudowany).

Także gracze *hardkorowi* i *casualowi* posługują się trochę innym słownictwem. Założyć można, że gracze *casualowi* w swojej mowie używają zdecydowanie mniej określeń specjalistycznych, w porównaniu z graczami *hardkorowymi*. Dodatkowym aspektem zróżnicowania leksykalnego jest to, że słownictwo środowiska graczy nie istnieje w izolacji

i czasami asymiluje pewne elementy z innych odmian specjalistycznych/środowiskowych języka polskiego, takich jak np. leksyka filmowa, internetowa czy informatyczna. Przykładowo *sequel*, który pierwotnie używany był w odniesieniu do kontynuacji jakiegoś filmu, obecnie odnosi się także do kontynuacji gry wideo.

Leksyka środowiska polskich graczy wideo nieustannie zmienia się, część jednostek zanika, wychodzi z użycia, nowe zastępują je, natomiast inne są absorbowane do języka polskiego, np. z angielskiego. Dużą rolę odgrywają tutaj czynniki zewnętrzne, które pełnią funkcję aktywatorów zmian językowych w danej społeczności.

Moje badania wpisują się w szeroko rozumiany nurt badań środowiska graczy wideo. Wielu badaczy (np. M.B. Garda i P. Grabarczyk [2014] lub np. K. Jędrasiak [praca doktorska]) wydziela nawet osobną dziedzinę badań nazwaną groznawstwem, dlatego ja też się nim posłużę, nazywając kolejną część pracy. Badania te zyskują na popularności, ale wciąż brakuje kompletnego zbadania socjolektu tej grupy społecznej.

1.3. BADANIA GROZNAWCZE W POLSCE

Biorąc pod uwagę skalę zjawiska, jakim jest granie, dziwi fakt, że wciąż brakuje językoznawczego opracowania socjolektu graczy. Pisze o tym między innymi Kamila Gądek: „W Polsce obecnie brak badań językoznawczych dotyczących takich zjawisk związanych z grami komputerowymi, jak chociażby (...) komunikacji w grach, przede wszystkim sieciowych, specyficznego użycia języka przy opisie interakcji gracza ze światem gry czy wykształcania się słownictwa związanego z grami” [Gądek 2009a]. Mimo że autorka napisała ten artykuł 9 lat temu, ciągle nie ma kompletnej językowej analizy socjolektu polskich graczy wideo. Kamila Gądek w swojej pracy magisterskiej dokonała analizy leksyki graczy komputerowych na podstawie tekstów dotyczących komputerowych gier przygodowych. Autorka ograniczyła się wyłącznie do zbadania leksyki graczy komputerowych oraz tylko do tekstów o grach przygodowych.

Poza pracami Kamili Gądek⁷ fragmentaryczne/przyczynkarskie opracowania leksyki graczy znaleźć można w kilku innych pracach językoznawczych. W artykule *Elementy socjolektu graczy komputerowych* [2005] autorka (Dominika Urbańska-Galanciak) skupiła się jedynie na scharakteryzowaniu języka internetowego środowiska użytkowników gier

⁷ Warto tu przytoczyć jeszcze takie artykuły Kamili Gądek, jak: *Metody opisu interakcji: gracz – postać – elementy świata gry w języku graczy komputerowych* [2009b] i *Cudzysłów jako wyraz niepewności i dystansu w tekstach dotyczących gier komputerowych* [2009c].

komputerowych, przyjrzała się jego leksyce oraz składni. W artykule *Komunikacja w świecie gier komputerowych* [2005] autorzy (Dorota Chmielewska-Łuczak i Tomasz Smejliś) przedstawili, jak wygląda i na jakich zasadach odbywa się dialog gracza z maszyną i z *avatarem*⁸, dialog twórcy z graczem, dialog gracza z twórcą oraz dialog między graczami. Przytoczyli i wyjaśnili także wiele terminów związanych z grami. Krzysztof Inglot w artykule *Nazwy własne a lokalizacja gier elektronicznych fantasy – o potrzebie przekładu oraz wybranych tendencjach w tłumaczeniu gier anglojęzycznych na rynek polski i niemiecki* [2013] przedstawił, jak wygląda proces tłumaczenia na język polski różnych nazw własnych, takich, jak np.: nazwy bohaterów, w tym potworów i tytuły gier.

Poza artykułami na uwagę zasługują także trzy pozycje książkowe⁹: *Wirtualny plac zabaw. Gry sieciowe i przemiany kultury współczesnej* [2006] Mirosława Filiciaka, *Fabularyzowane gry komputerowe w przestrzeni humanistycznej* [2012] Rafała Kochanowicza oraz *Homo players. Strategie odbioru gier komputerowych* [2009] Dominiki Urbańskiej-Galanciak. W pierwszej pracy autor poświęcił jeden podrozdział (*Komunikacja werbalna – mechanizmy komunikacyjne oraz slang graczy*) na wyjaśnienie kilkunastu terminów, głównie skrótów, którymi posługują się użytkownicy internetowej gry *EverQuest*¹⁰. Natomiast w drugiej pracy jej autor zwraca uwagę na to, że leksyka graczy szybko ewoluuje i przeistacza się, a w efekcie powstają nowe terminy, inne natomiast wychodzą z użycia. Autor wyjaśnia także wiele terminów związanych z grami.

Na zakończenie omawiania pozycji książkowych trzeba wspomnieć, że najobszerniejszym opracowaniem zagadnienia (socjolekt polskich graczy wideo) jest zaledwie trzydziestostronicowy rozdział „język cyberrozrywki”, który umieszczony został w pracy Dominiki Urbańskiej-Galanciak pt. *Homo players. Strategie odbioru gier komputerowych*. Autorka wyszczególnia w tym rozdziale trzy podrozdziały: *Komputerowo mediatyzowana komunikacja*, *Tendencje rozszerzające – język jako czynnik wyodrębniający zbiorowość* oraz *Tendencje upraszczające – język jako środek jednoczenia wspólnoty*. W pierwszym podrozdziale D. Urbańska-Galanciak starała się odpowiedzieć na pytanie, jakie jest miejsce socjolektu w grupie społecznej, cytuje więc E. Sapira, który zauważył, że grupy społeczne związane wspólnym interesem wykazują „tendencje do wytwarzania pewnych osobliwości

⁸ *Avatar* to „postać, którą kieruje gracz” [Chmielewska-Łuczak, Smejliś 2005: 327].

⁹ Książek opisujących środowisko graczy wideo jest oczywiście więcej (np. *Gry wideo. Zarys poetyki* [2016] Piotra Kubińskiego czy *Kultura gier komputerowych* [2011] Johna Dovey’a i Helen W. Kennedy), ale pomijają one szczegółowy opis socjolektu środowiska graczy wideo.

¹⁰ Gra internetowa z gatunku RPG (*role-playing game*). Gracz steruje stworzoną przez siebie postacią; rozwija ją, dzięki czemu może pokonywać coraz to mocniejszych przeciwników. Gra powstała w 1999 roku.

mowy, które pełnią symboliczną funkcję wyodrębniania danej grupy z liczniejszej” [Sapir 1978, cyt. za Urbańska-Galanciak 2009: 107]. Autorka podkreśla także, że „część wyrazów powstająca pierwotnie w języku graczy przenoszona jest, często przez nich samych, do języka potocznego i używana w innych okolicznościach” [Urbańska-Galanciak 2009: 113]. W drugim podrozdziale (*Tendencje rozszerzające – język jako czynnik wyodrębniający zbiorowość*) D. Urbańska-Galanciak stwierdza, że „język cyberrozrywki jest (...) językiem hermetycznym, dostępnym dla osób dobrze zorientowanych w świecie wirtualnych zabaw” [Urbańska-Galanciak 2009: 115]. Z tym stwierdzeniem niestety nie mogę się zgodzić, gdyż coraz więcej osób ma do czynienia z grami i ich język siłą rzeczy traci swoją hermetyczność. Te osoby nieustannie komunikują się ze sobą, dlatego terminy początkowo zarezerwowane dla środowiska graczy, przestają takimi być. Autorka pisała książkę przeszło osiem lat temu, wtedy gry były zdecydowanie mniej popularnym sposobem spędzania wolnego czasu, a leksyka była bardziej hermetyczna. W tym podrozdziale dokonana została także szczątkowa charakterystyka leksyki graczy. Autorka wyszczególniła takie grupy, jak np. zapożyczenia, neologizmy czy skróty i wymieniła po kilka przykładów leksemów pasujących do tych grup. W ostatnim podrozdziale (*Tendencje upraszczające – język jako środek jednoczenia wspólnoty*) badaczka pisze, że „socjolekt nie tylko wyodrębnia zbiorowość, która go tworzy, ale też jednoczy jednostki posługujące się nim” [Urbańska-Galanciak 2009: 125]. W dalszej części autorka analizuje wybrane wypowiedzi graczy z forum internetowego Gry-Online¹¹.

Pozycji tej poświęciłem najwięcej miejsca podczas omawiania literatury związanej ze środowiskiem graczy wideo, gdyż ten trzydziestostronicowy rozdział, napisany przez kulturoznawcę stanowi, jak dotąd, najobszerniejsze opracowanie leksyki graczy.

Udało mi się także dotrzeć do nieopublikowanej pracy magisterskiej *Język komunikacji uczestników gier sieciowych* autorstwa Jędrzeja Chruściela, w której umieszczone zostały wyjaśnienia wielu terminów używanych w grach sieciowych.

Jeśli natomiast chodzi o opracowania internetowe, to należy wspomnieć o *Słowniku gracza*, który znajduje się na stronie internetowej www.gry-online.pl/slownik-gracza.asp [dostęp: 26.05.2017]. Definicje w tym słowniku „koncentrują się na hasłach używanych zarówno w artykułach publikowanych na łamach serwisu informacyjnego gry-online.pl, jak i rozmowach na naszym forum dyskusyjnym”¹². Jest to słownik stworzony przez graczy i dla graczy. W sposób nieskomplikowany wyjaśnia występujące w języku graczy pojęcia. Cały

¹¹ Forum znaleźć można na stronie internetowej: www.gry-online.pl/S008.asp [dostęp: 25.05.2017].

¹² Cytat ten pochodzi ze strony: www.gry-online.pl/slownik-gracza.asp [dostęp: 26.05.2017].

słownik zawiera tylko 365 pozycji, co stanowi, według mojego rozeznania, ułamek całego zbioru specjalistycznej leksyki środowiska graczy.

Moje badania plasują się na pograniczu badań groznawczych i socjolingwistycznych, dlatego w kolejnej części omówię rozwój dziedzin związanych ze zróżnicowaniem językowym polszczyzny. Zacznę od dialektologii, która dała podwaliny do wytworzenia socjolingwistyki.

2. BADANIA NAD ZRÓŻNICOWANIEM JĘZYKOWYM POLSZCZYZNY

2.1. POCZĄTEK BADAŃ W RAMACH DIALEKTOLOGII

Początki badań nad zróżnicowaniem językowym polszczyzny przypadają na drugą połowę XIX w., kiedy to zaczęło się „systematyczne badanie gwar” [Dunaj 1986: 15]. Wiąże się ono ściśle z pierwszymi polskimi dialektologami. Stanisław Urbańczyk w *Zarysie dialektologii polskiej* podaje, że pierwszym polskim dialektologiem był Lucjan Malinowski. W centrum jego zainteresowań był dialekt śląski [Urbańczyk 1981: 14]. I tutaj pojawia się pewna wątpliwość, czy Malinowski rzeczywiście był pierwszy? W artykule Kazimierza Nitscha *Próba ugrupowania gwar polskich*, który wydany został dużo wcześniej niż tekst Urbańczyka, bo w 1910 roku, przeczytać można, że Lucjan Malinowski był twórcą polskiej dialektologii, natomiast autor nie twierdzi, jakoby był on pierwszym dialektologiem [Nitsch 1910: 336]. Sięgając jeszcze do wcześniejszych artykułów, natrafić można na pracę Drwęcy¹³ *Rzecz o dyalektach mowy polskiej Wincentego Pola* [1869], w której autor pisze, że niejaki Wincenty Pol przebadał „cały nowy materiał nietknięty jeszcze zupełnie, ani przez historyków, ani grammatyków (...) on przeto pierwszy wprowadza nas tutaj na szersze szranki (...) mowy polskiej” [Drwęca 1869: 3]. Wynika stąd, że pierwszym dialektologiem nie był, jak twierdzi Urbańczyk, Lucjan Malinowski, ale właśnie Wincenty Pol. I choć, jak twierdzi K. Nitsch, praca Pola zdumiewała brakiem podstawy naukowej [Nitsch 1911: 19], której moim zdaniem nie miał skąd pojąć, bo wtedy dopiero rodziła się polska dialektologia, to nic nie odbiera mu tytułu pierwszego polskiego dialektologa¹⁴. Warto też dodać, że Pol z wykształcenia był geografem, więc, moim zdaniem, miał podstawy do rozpoczęcia tego typu badań. Argumentem wzmacniającym tezę o pierwszeństwie Pola jest to, że jego praca wydana została w 1869 roku, czyli można przypuszczać, że autor pracował nad jej przygotowaniem rok lub kilka lat wcześniej, natomiast Malinowski w roku 1869 odbywał dopiero „kilkumiesięczną podróż naukową po Śląsku” [Nitsch 1911: 18] i przygotowywał swoją pracę. Reasumując, w roku

¹³ Przypuszczam, że jest to pseudonim naukowy autora. W ten sposób podpisał się pod pracą.

¹⁴ Wcześniejsze od polskich badań dialektologicznych były badania etnograficzne ludów na terenie Polski. To w nich można znaleźć pierwociny dotyczące pewnych zjawisk językowych, np. w pierwszym tomie (pt. *Pieśni ludu polskiego*) serii etnograficznej – *Lud. Jego zwyczaje, sposób życia, mowa, podania, przysłowia, obrzędy, gusła, zabawy, pieśni, muzyka i tańce* Oskara Kolberga umieszczona została informacja na temat mazurzenia [Kolberg 1857: IX–X]. Niemniej jest to tylko wzmianka, a całość pracy nie można potraktować jako opracowania dialektologicznego.

1869, za sprawą pracy Wincentego Pola *Rzecz o dialektach mowy polskiej*, miała swój początek polska dialektologia, z której następnie, po wielu latach, wyrosła polska socjolingwistyka.

Na początku istnienia polskiej dialektologii głównym problemem badawczym stało się ustalenie liczby polskich dialektów oraz to, jaki jest ich stosunek do *wariantu ogólnego języka polskiego*.

Wincenty Pol wyróżnił siedem dialektów: *chrobacki* (z trzema odcieniami: właściwy chrobacki, szlązki i krakowski), *wielkopolski* (z trzema odcieniami: Odraków, Pałużan i Kujawiaków), *małopolski* (z trzema odcieniami: sanocki, sandomierski i lubelski), *mazowiecki* (z sześcioma odcieniami: podlaski, pogranicza litewskiego, Gbórow, Warmii polskiej, Kaszubów i właściwy Warszawie), *białoruski* (trzy odcienie: właściwy, Czarnoruski i Piński), *litewsko-polski* (brak odcieni) i *opolski* (cztery odcienie: Wołyński, Poberecki, Podolski i Polan kijowski) [Pol 1869]. Według autora dialekt małopolski jest najbliższy mowie Polaków, to „najczystszy odcień mowy polskiej” [Pol 1869: 28].

Podział Pola uwzględniał ukształtowanie geograficzne Polski („dialekty będą się tak dzieliły, jak się obszary siedmiu głównych rzek dzielą” [Pol 1869: 13]) oraz zróżnicowanie językowe, np. autor zwraca uwagę na terytoria geograficzne kraju, w których ludność mazurzy, oraz na takie, w których ono nie występuje. Podkreślić trzeba, że nie wyróżniono w tym podziale, jako osobnego, dialektu śląskiego, jest on jedynie odcieniem *chrobackiego*.

Dialekt śląski, jak już wspomiano, był w centrum zainteresowania Lucjana Malinowskiego, który znacząco przyczynił się do poznania tego dialektu języka polskiego. W *Listach z podróży etnograficznej po Szląsku* pisał „chodząc ode wsi do wsi, zapisując powieści i pieśni, ciągle między ludem, miałem sposobność przypatrzeć się blisko jego bytowi i stosunkom” [Malinowski 1871: 289].

Pod koniec XIX wieku w „Rozprawach i sprawozdaniach z posiedzeń Wydziału Filologicznego Akademii Umiejętności” opublikowanych zostało wiele wyników badań dialektologicznych, takich, jak np.: *Spostrzeżenia nad właściwościami językowymi w mowie górali bieskidowych, z dodatkiem słowniczka wyrazów góralskich* J. Kopernickiego [1875], *Gwara Lasowska w okolicy Tarnobrzega. Studium dialektologiczne* S. Matusiaka [1880] czy *Przyczynek do gwary zakopańskiej* W. Kosińskiego [1884]. Wszystkie one niestety dotyczyły tylko niewielkich obszarów Polski, najczęściej autorzy ograniczali się do zbadania gwary występującej w obrębie jednej lub kilku wsi. W tym okresie pojawiły się też pierwsze zainteresowania odmianą języka polskiego, występującą w dużych miastach. Jako przykład

może posłużyć artykuł L.W. Szczerbowicza *Gwara warszawska w społecznym języku literackim* [1890], który rozpoczyna się następująco:

Nieprawidłowości, o których mam mówić, są albo nowotworami, w dotychczasowym języku wykształconym i literackim nieznanym, albo formy i zwroty, odmienne od uświęconych w literaturze. Może znajdzie się ktoś taki, co zechce bronić tych nabytków, opierając się na tej zasadzie, że język żyjący wciąż bogaci się i rozwija. Wszelako, według naszego zdania, wzrost i pomnażanie się niezawsze jest bogaceniem się i rozwojem, a mianowicie w tym wypadku, gdy mnożenie się nowotworów językowych, najczęściej zupełnie zbytecznych, idzie w parze z zaniedbaniem bogatego słownika, dostatecznie urobionego i ogładzonego na podstawie odwiecznej mowy ludu, przez Kochanowskiego, Skargę, Sniadeckich (sic!)¹⁵ i wzorowych pisarzy naszego stulecia (...)

[Szczerbowicz 1890: 202].

W przytoczonym powyżej fragmencie autor ujawnia pejoratywny stosunek do innych niż literackie odmiany języka. W tamtych czasach gwara często postrzegana była przez elity, ale i pozostałe grupy społeczne jako coś zepsutego, z czym trzeba walczyć poprzez edukację [Nitsch 1911: 1]. Był to gorszy wariant „języka literackiego”. Na szczęście znaleźli się też tacy badacze, którzy podejmowali się badania gwar czy też dialektów i publikowali wyniki swoich prac.

Na początku XX w. opublikowany został artykuł Nitscha *Próba ugrupowania gwar polskich* [1910], w którym autor zaprezentował dialektalny podział terytorium Polski, uwzględniając różnice językowe między danymi regionami. Autor zaznacza, że jest to „bezwzględnie pierwsza próba tego rodzaju” [Nitsch 1910: 21], natomiast po dokładniejszym zapoznaniu się z literaturą, należy stwierdzić, że jest to druga naukowa próba dialektalnego podziału Polski (pierwsza była W. Pola), a nawet trzecia uwzględniając *Słownik gwar polskich* Jana Karłowicza [1900–1911]. Pomijam natomiast, ze względu na nienaukowy charakter, książkę popularnonaukową Karłowicza *Lud. Rys ludoznawstwa polskiego* [1906], w której autor podzielił, a następnie scharakteryzował, m.in. pod względem językowym, poszczególne regiony Polski. Jego praca została bardzo negatywnie oceniona przez Kazimierza Nitscha, który zarzucał autorowi podawanie fałszywych informacji oraz brak fachowości [Nitsch 1911]. Pierwszeństwa nie można odebrać pracy Nitscha, jeśli chodzi o podział terytorium Polski z uwzględnieniem wyłącznie cech fonetycznych. Jest to absolutnie pierwszy tego typu podział, ponieważ wszyscy poprzednicy łączyli cechy językowe z wiedzą geograficzną i/lub historyczną, często zaczynając od podziału geograficznego, a następnie dokładając cechy

¹⁵ Pozostawiłem pisownię oryginalną.

językowe. Nitsch natomiast postąpił odwrotnie, a więc wyszedł od cech fonetycznych, a następnie przeszedł do geografii Polski.

Nitsch zaproponował podział, uwzględniając dwie cechy fonetyczne: mazurzenie i fonetykę międzywyrazową. Podzielił więc Polskę na mazurzącą (Mazowsze, Małopolska z Łęczyckiem i Sieradzkim oraz Śląsk) i niemazurzącą (Wielkopolska z Kujawami i ziemią Dobrzyńską) oraz uwzględniając fonetykę międzywyrazową: na południowo-zachodnią (fonetyka międzywyrazowa udźwięczniająca) i północno-wschodnią (fonetyka międzywyrazowa ubezdźwięczniająca) [Nitsch 1910]. Podział ten okazał się na tyle trafny, że stanowi obecnie podstawy polskiej dialektologii¹⁶.

Rok później Nitsch opublikował *Mowę ludu polskiego* [1911]. Pierwszą część pracy autor poświęcił terminologii. Między innymi zastanawiał się nad pojęciami „*dialekt*” i „*język*” oraz ich stosunkiem do siebie. Kluczowe tutaj także jest pojęcie „*język indywidualny*”¹⁷, czyli „*język*” pojedynczego człowieka. Autor uważa, że jest tyle języków, ilu jest mówiących ludzi na świecie. I nie chodzi tutaj o te osoby, które mówią w danej chwili, ale o to, co poszczególni ludzie na świecie są w stanie powiedzieć, natomiast „*język*” lub „*dialekt*”, to tylko zbiory „*języków indywidualnych*”. Jakie więc są wzajemne relacje tych dwóch terminów? Według Nitscha są to pojęcia względne i stoją do siebie w stosunku identycznym, jak rodzaj do gatunku, przy czym trzeba pamiętać, że „każdy rodzaj w stosunku do jakiegokolwiek pojęcia nadrzędnego staje się gatunkiem” [Nitsch 1911: 9]. Dla autora „*dialektem*” może być „*mowa*” jakiejś wioski, zestawiona z „*mową indoeuropejską*”, która jest w tym przypadku „*językiem*”, ale może ona także stać się „*dialektem*”, gdy tylko zestawimy ją z jakąś większą grupą. Dzisiaj, moim zdaniem, takie rozumienie tych pojęć znacznie utrudniłoby pracę, ponieważ każdy badacz mógłby przyjąć zupełnie inną perspektywę postrzegania terminów *dialekt* i *język*. Wprowadziłoby to tylko niepotrzebny zamęt. Aczkolwiek termin „*język*” ciągle bywa używany w wielu różnych znaczeniach. W dalszej części pracy Nitsch rozwija, zaproponowaną w artykule z 1910 roku, koncepcję podziału dialektalnego Polski.

Kolejną dużą i ważną klasyfikację zaproponował dopiero w 1966 roku Antoni Furdal. Skupia się on na klasyfikacji polskich „*dialektów miejskich*”, które sam rozumie jako warstwę

¹⁶ W powyższym akapicie zachowano pisownię oryginalną.

¹⁷ Gdy w pracy przytaczam jakiś termin, to używam kursywy, aby go bardziej wyróżnić. Pewną część terminów, zwłaszcza w pierwszej części pracy, dodatkowo wzięłem w cudzysłów. Uczyniłem tak, ponieważ autorzy prac używają wielu różnych terminów i często w skrajnie różnych znaczeniach, że termin posiadający taką samą formę graficzną, w różnych tekstach może mieć zupełnie inne znaczenie. Często z terminami lub terminami w znaczeniu zaproponowanym przez autora (czasami także z autorskimi połączeniami wyrazowymi) nie zgadzam się, wtedy umieszczam go w cudzysłowie, aby zasygnalizować odbiorcy, że to nie jest termin używany przeze mnie, ale przez autora przytaczanego tekstu.

pośrednią między „gwarami ludowymi” a „dialektem kulturalnym” lub, jak pisze w innym miejscu artykułu, ułożone są one między „gwarami” a „językiem literackim” [Furdal 1966: 155]. Autor używa wymiennie terminów „język literacki” i „dialekt kulturalny”, co nieco burzy porządek wypowiedzi. Część kluczową jego artykułu stanowi podział „dialektów miejskich”, który uwarunkowany jest historycznie (autor uwzględnił wydarzenia historyczne, które wpłynęły na podział terytorialny Polski, np. zabory czy też rozbiecie dzielnicowe). Wyróżnionych zostało dziewięć „dialektów miejskich”: warszawski, północnomalopolski („zaliczymy tu Radom, Łódź, także chyba Kielce i Lublin” [Furdal 1966: 157]), krakowski, śląski, poznański, pomorski (tutaj miasta takie, jak: Grudziądz, Tczew, Starogard), mazowiecki (miasta nieleżące w bezpośrednim sąsiedztwie Warszawy), wileński i lwowski. Klasyfikacja jest początkiem zainteresowań tego badacza podziałami języka polskiego, kolejnym jego krokiem będzie zaproponowanie własnej klasyfikacji wszystkich odmian polszczyzny.

Jak wynika z powyższych ustaleń, dialektologia dała impuls do rozwoju badań nad zróżnicowaniem polszczyzny. Pokazała, że naukowiec nie może poprzestać na kulturze elitarniej, ale powinien wsłuchiwać się w mowę wszystkich rodaków. Z przeświadczeń tych zrodziła się socjolingwistyka. Która jest dziedziną ciągłych poszukiwań i opisu kolejnych pokoleń badaczy.

2.2. BADANIA SOCJOLINGWISTYCZNE

Naturalną konsekwencją zainteresowań odmianami dialektalnymi języka polskiego było to, że po jakimś czasie badacze zainteresowali się również wariantami językowymi grup społecznych/socjalnych. Początkowo publikowano średnio jedną pracę na kilka lat, potem wraz z pojawieniem się nauk grup społecznych, intensywność badań zwiększała się.

W 1903 roku opublikowana została praca Karola Estreichera pt. *Szwargot więzienny*. Jest to, według mojego rozeznania, pierwsza duża praca, która w całości dotyczyła wariantu językowego określonej grupy społecznej. Już wcześniej Estreicher podejmował pierwsze próby przyjrzenia się temu wariantowi języka polskiego. Na łamach „Rozmaitości” (dodatku do „Gazety Lwowskiej”) w 1859 roku (w numerach 12, 13 i 14) Mateusz Szreniewa Sartini (redaktor gazety) opublikował artykuł Estreichera *Język złoczyńców*, w którym krótko scharakteryzowane zostało środowisko złoczyńców oraz przedstawiono, wraz z omówieniem, ich szcątkową twórczość artystyczną. Estreicher dołączył także „słowniczek polsko-złodziejski” [Estreicher 1859 (14): 107], w którym zdefiniował 391 leksemów tego środowiska. Po publikacji artykułu i jego licznych przedrukach (np. w „Gazecie Polskiej” czy „Kurierze

Warszawskim”¹⁸) kilka osób (np. więzień polityczny Władysław Czaplicki) skontaktowało się z autorem, w celu uzupełnienia jego słowniczka. Prace Estreichera, zarówno artykuł, jak i wersja książkowa, miały wymiar nie tylko archiwizatorski, ale także praktyczny, bo jak sam autor pisze „ogłoszenie zbioru było użytecznym dla praktyki kryminalnej” [Estreicher 1903: 9]. Prace jego były także wykorzystywane przez władze policyjne. Reasumując, była to praca pionierska w dziedzinie badania mowy grup społecznych. Jednak w owym czasie dominowały jeszcze badania dialektologiczne. Kolejna duża praca związana z odmianami językowymi wybranych grup społecznych pojawiła się dopiero w 1951 roku, a była nią książka Henryka Ułaszyna *Język złodziejski. La langue des voleurs*. Jest ona pokłosiem wygłoszonych w latach 1926–1933 wykładów dotyczących właśnie „języka złodziejskiego”. Uczęszczali na nie głównie policjanci, kryminolodzy i inne osoby związane z prawem. Spisywali oni notatki, które następnie były rozpowszechniane jako skrypty, ponieważ poznanie „języka przestępców” było kluczowe do rozwiązania wielu zagadek kryminalnych (np. odczytania przechwyconego listu jakiegoś przestępcy). Po zakończeniu serii wykładów zaczęto namawiać Ułaszyna, żeby je opracował i wydał. Z genezą pracy związana jest anegdota z wątkiem łódzkim. Otóż do wydania pracy zachęcali badacza nawet ówczesny sędzia Sądu Najwyższego Jan Bossowski oraz wiceprokurator Sądu Administracyjnego C. Matysiak. Ułaszyn uległ ich namowom i wydał wykłady w formie książkowej, ale jak sam zaznaczył we wstępie, solidniej nie mógł ich dopracować, ponieważ ma „75 lat życia za sobą i 5 lat ciężkiego chorowania, jako skutków dwóch wojen światowych (pierwszą też ciężko odchorowałem) i łódzkich warunków mego życia” [Ułaszyn 1951: 13].

Na początku książki H. Ułaszyn ustosunkowuje się do takich terminów, jak: *język, gwara, dialekt, żargon i szwargot*. Dwa ostatnie uważa za nienaukowe. Nieco dalej przedstawiony został podział społeczeństwa, uwzględniający różnice językowe, a przebiegał on w dwóch kierunkach: poziomym (terytorialnym, geograficznym) i pionowym (nawarstwienia społeczne, zawodowe, ideologiczne). Z kierunkiem poziomym łączyć należy takie terminy, jak: *gwara, dialekt czy narzecze* (ostatnie dwa można traktować jako synonimy), zaś z pionowym głównie *język* (czasami też *gwara*). Ułaszyn uważa też, że *dialekt/narzecze* są jednostkami większymi niż *gwara*.

W dalszej części autor charakteryzuje i grupuje słownictwo „języka przestępców”, wyszczególniając takie grupy, jak nazwy: złodziei, kradzionych przedmiotów, schowków (miejsc, z których się kradnie), dokonywanych kradzieży, narzędzi pomocniczych, wrogów

¹⁸ Artykuł na przestrzeni lat wielokrotnie zmieniał swój tytuł, więc znaleźć go można pod takimi tytułami, jak np. *Gwara złoczyńców czy Szwargot złoczyńców*.

złodziei, sytuacji kradzieżowych, sytuacji związanych z aresztem i sądem oraz z więzieniem. Dominują w tym „języku” różnego rodzaju zapożyczenia (np. niemieckie, żydowskie czy ruskie). Elementem kończącym tę pracę jest historia „języków złodziejskich” na świecie (np. we Francji czy Anglii).

Sześć lat później, w 1957, ukazała się trzecia duża praca dotycząca wariantu językowego grupy społecznej, tym razem ochweśników, czyli wędrownych przekupniów lub po prostu kupców. Praca zatytułowana została *Żargon ochweśników*, a jej autorką była Wanda Budziszewska. Na temat wariantu języka polskiego tej grupy, do momentu wydania książki Budziszewskiej, powstał jedynie krótki artykuł autorstwa Stanisława Górki *Skulscy „ochweśnicy”* [1901], w którym scharakteryzowano „obraźników” (osoby handlujące obrazami) i dodano „maleńki słowniczek” [Górka 1901: 4] wyjaśniający 194 wyrazy używane przez to środowisko.

Celem Budziszewskiej było zebranie możliwie dużego materiału i przeanalizowanie go, a przede wszystkim zarchiwizowanie, by nie uległ całkowitemu zapomnieniu, ponieważ w tamtym czasie ten wariant polszczyzny zaczął już zanikać [Budziszewska 1957: 6]. Autorka opisała podstawowy zasób słów tego „żargonu”, jego słowotwórstwo, głosownię (obecnie używa się terminu fonetyka – przyp. autora) oraz składnię, a także podała znaczenie wyrazów i ich odmianę. Na końcu pracy dodany został słownik ochweśnicki. Warto w tym miejscu zauważyć, że częstym elementem strukturalnym prac socjologicznych jest słownik leksyki badanej grupy społecznej.

Dopiero w 1972 roku, zaczyna wyodrębniać się, jako osobna gałąź językoznawstwa, socjolingwistyka. W tym roku Ludwik Zabrocki publikuje pracę *Z teorii socjolingwistyki*, pierwszy znany mi polski artykuł, w którym autor cały tekst poświęcił nowo wyodrębnionej dziedzinie. Przedstawiono w nim, jak wyglądał rozwój socjolingwistyki na świecie, przytoczono teorie zagranicznych socjolingwistów, np. Bernsteina¹⁹ oraz z niektórymi dyskutowano, np. z twierdzeniem Grossego i Neuberta, że socjolingwistyka jest częścią pragmatyki. Zabrocki przedstawia także cel badawczy socjolingwistyki, która według niego bada „warianty języka, uwarunkowane odpowiednim statusem socjalnym” [Zabrocki 1972: 18], a także zastanawia się, jaki jest stosunek socjolingwistyki do innych dyscyplin naukowych, takich jak socjologia języka czy pragmatyka. Według niego socjolingwistyki nie można

¹⁹ Zabrocki przytacza między innymi teorię kodów restrykcyjnych Bernsteina, według której kody restrykcyjne są charakterystyczne dla klas robotniczych. A co z tym się wiąże, to że dzieci z tych środowisk mają większe trudności w uczeniu się, w odróżnieniu od dzieci z wyższych klas społecznych [Zabrocki 1972: 22]. W nowszych pracach o teorii Bernsteina *kod restrykcyjny* określany jest *kodem ograniczonym*.

traktować jako synonimu socjologii języka, ponieważ inny jest punkt wyjścia obu dziedzin. Socjologowie języka wychodzą od grupy społecznej, po czym docierają do odpowiedniego wariantu językowego, natomiast socjolingwiści na początku przyglądają się danemu wariantowi językowemu, po czym docierają do grupy społecznej, która go używa. Przy czym trzeba podkreślić, że zjawiska językowe rzadko kiedy całkowicie będą pokrywać się z granicami jednej grupy, najczęściej będą interferowały między różnymi środowiskami. Zabrocki twierdzi, że pragmatyka ma za zadanie wskazanie tych środków językowych, które mają wywołać u odbiorcy właściwy skutek pożądany przez nadawcę. Powinien on przy tym uwzględnić informacje, które dostarczane są przez badania socjolingwistyczne. Socjolingwistyka to nauka o „stanach statycznych”, czyli jej wyniki wcale nie muszą, aczkolwiek mogą, być wykorzystane w komunikacji językowej, natomiast „pragmatyka wyrasta z samego procesu komunikacji” [Zabrocki 1972: 23]. Nie można więc uznać socjolingwistyki za część pragmatyki.

W tym samym roku ukazuje się artykuł Władysława Kaniuki *Tendencje rozwojowe współczesnej socjolingwistyki* [1972]. Jest on wart odnotowania, ponieważ tutaj, według moich rozeznań, po raz pierwszy wprowadzony został termin „*socjolekt*”. A użyty on został w definicji socjolingwistyki, która według autora zajmuje się badaniem „*socjolektów*” i „*dialektów*”. Nieco dalej przytoczone zostało zadanie R. Grossego i A. Neuberta, które autor podziela, że dialektologię trzeba uznać za część składową socjolingwistyki. Uważam, że dialektologia i socjolingwistyka mają ze sobą wiele wspólnego, ale nie można zdegradować dialektologii do subdyscypliny socjolingwistyki. Dialektologia dała początek badaniom nad zróżnicowaniem polszczyzny i to z niej wyrosła socjolingwistyka. Widać to w stosowanej terminologii, np. „*gwara*” – termin czysto dialektologiczny na oznaczenie wariantu językowego panującego na danym terenie, przejęty został przez socjolingwistykę i odnosił się już nie tylko do wariantu terytorialnego, lecz także, co jest moim zdaniem błędem, do wariantu socjalnego języka. Przywołać tu należy także podział polszczyzny na: pionowe (społeczne) i poziome (terytorialne) odmiany językowe, pisał o nich m.in. Henryk Ułaszyn (patrz strona 23). Teoretycznie, gdyby przyjąć za Kaniuką, że socjolingwistyka jest dziedziną nadrzędną wobec dialektologii, to obejmowałaby ona swoim zakresem badania dialektologiczne oraz badania nad społecznymi wariantami polszczyzny. Tym samym trzeba by było stworzyć nową nazwę dla dziedziny, która przejęłaby zakres badawczy socjolingwistyki. Uważam więc, że lepiej zostać przy rozgraniczeniu Zabrockiego, a więc że socjolingwistyka bada socjalne, a tak dialektologia terytorialne warianty języka polskiego, a obecny stosunek tych dwóch dyscyplin do siebie jest równorzędny.

Dopiero od roku 1972, można twierdzić, że badania socjolingwistyczne zaczęły się oddzielać od dialektologicznych, a socjolingwistyka coraz bardziej usamodzielniała się. Jednak tak naprawdę minie jeszcze kilka lat, zanim rzeczywiście wśród badaczy zostanie uznana za pełnoprawną dyscyplinę językoznawczą.

W 1974 roku Stanisław Grabias zaproponował funkcjonalną klasyfikację „*socjalnych wariantów języka*” [1974]. Wyszczególnił on trzy ich główne cechy, na podstawie dotychczasowych opracowań różnych odmian socjalnych, które występują w języku polskim („*żargon marynarzy*”, „*język łowiecki*”, „*żargon ochweśnicki*”, „*język złodziejski*”, „*gwara partyjna*”, „*tajny język dzieci*”, „*gwara studencka*”). Pierwszą jest zawodowość, która przejawia się głównie w obecności leksyki specjalistycznej. Ważne tutaj są takie aspekty, jak: rodzaj leksyki (np. czy spełnia warunek ekonomiczności), jej ilość względem ogólnopolskich wariantów leksykalnych występujących w danym wariantcie języka oraz jej charakter, czyli to, czy wiąże się tematycznie z działalnością danej społeczności. Jako przykład ograniczonej realizacji zawodowości autor podaje „*gwarę studencką*”, w której tylko 1/3 całego słownictwa związana jest z uczelnią i studiowaniem. Dobrą realizacją tej cechy jest „*język łowiecki*”, w którym prawie wszystkie wyrazy dotyczą realiów myśliwskich.

Kolejną cechą jest tajność, nie rozumiana jednak jako niezrozumiałość, ale jako celowe działanie nadawcy.

Jako przykłady środków formalnych stosowanych w celu uzyskania tajności, Grabias podaje: maskowanie wyrazów poprzez zamianę formantów np. *gluporny* (*glupi*) czy *czarnośny* (*czarny*), dodawanie cząstek, które zniekształcają wyrazy, np. *byco* (*co*), *bynic* (*nic*) i skracanie wyrazów, np. *gra* (*granica*), *wa* (*Warszawa*). Charakter tajny ma np. „*gwara złodziejska*”, natomiast półtajny (krypto-zabawowy) „*tajny język dzieci*”. Gwarze studenckiej natomiast nie można przypisać cechy tajności.

Ostatnią cechą jest ekspresywność. Cecha ta najbardziej uwydatniona została w „*gwarze studenckiej*”, w której np. występują długie ciągi synonimiczne – wyraz *jeść* ma 91 synonimów, a różnią się one tylko stopniem ekspresywności. Przykładem wariantu socjalnego, w którym niski jest stopień ekspresywności może być „*gwara więzienna*”.

W podsumowaniu artykułu autor twierdzi, że wszystkie socjalne warianty polszczyzny można podzielić na dwie grupy:

1. odmiany o prymarnej funkcji profesjonalno-komunikatywnej, czyli
 - a) zawodowe odmiany języka o charakterze jawnym (dobór środków językowych podporządkowany jest w nich sprawnemu i ścisłemu przekazywaniu myśli) i

b) żargony – genetycznie tajne odmiany języka (środki językowe podporządkowane są zachowaniu tajności);

2. odmiany o prymarnej funkcji ekspresywnej:

a) gwary środowiskowe o charakterze jawnym (środowiskowe warianty natury kolokwialnej, w których dobór środków językowych podporządkowany jest ekspresywności);

b) warianty językowe o charakterze krypto-zabawowym (np. tzw. tajny język dzieci).

[Grabias 1974: 31]

Rok 1977 był przełomowy dla polskiej socjolingwistyki, ponieważ wtedy premierę miał pierwszy numer czasopisma „Socjolingwistyka”, pod redakcją Władysława Lubasia. Badacz ten powszechnie uważany jest za ojca polskiej socjolingwistyki. Można przypuszczać, że jeśli już istnieje czasopismo, które w nazwie ma termin „*socjolingwistyka*”, to od tego momentu można ją uważać za pełnoprawną dyscyplinę językoznawczą. Wielu badaczy tak uważało (np. Zabrocki), ale niestety, co zaskakujące, inaczej twierdził sam redaktor czasopisma – Stanisław Lubaś, który już w drugim tomie opublikował swój artykuł pt. *Socjolingwistyka jako metoda badawcza* [1979], natomiast w czwartym numerze utwierdził czytelnika w tym, że jest ona tylko metodą, bo opublikował pracę *Socjolingwistyka – metoda interdyscyplinarna* [1982]. Niezależnie od terminologicznych kontrowersji, były to ważne dla językoznawstwa (programowe) artykuły. W pierwszej pracy, z 1979 roku, autor dokonuje przeglądu definicji socjolingwistyki, socjologii języka i socjologii mowy, pokazując, że nie są one, jak twierdzili niektórzy ówczesni badacze, synonimami nazywającymi tę samą gałąź nauki.

W 1982 roku w Łodzi odbyła się ważna dla socjolingwistyki konferencja językoznawcza, której tematem głównym był stosunek dialektologii do innych dziedzin naukowych, w tym do socjolingwistyki. Jej pokłosiem był wydany w 1986 roku tom z serii „Acta Universitatis Lodzianensis”. W jednym z artykułów opublikowanych w tym tomie Bogusław Dunaj zestawia dialektologię z socjolingwistyką [Dunaj 1986]. Umieszcza w nim stwierdzenie, które utwierdza mnie w przekonaniu, że socjolingwistyka wyrasta z dialektologii, a mianowicie, że „dobra znajomość wariantów regionalnych i form wywodzących się z podłoża gwarowego jest nie tylko przydatna, ale wręcz nieodzowna w badaniach socjolingwistycznych” [Dunaj 1986: 19]. Natomiast Maria Kamińska, autorka artykułu *Powiązania dialektologii z socjolingwistyką* [1986], staje w obronie dialektologii jako osobnej dziedziny językoznawczej. Jest to odpowiedź, między innymi, na artykuł Kaniuki, który twierdzi, że dialektologia jest częścią socjolingwistyki. Co ciekawe, M. Kamińska popiera zdanie Lubasia, który twierdzi, że socjolingwistyka jest jedynie metodą badawczą. Autorka pisze także, że dialektologia ma jeszcze przed sobą do zbadania wiele aspektów języka polskiego, np. frazeologię, składnię czy stylistykę ludową [Kamińska 1986].

Artykuł, który zasługuje na szczególną uwagę i z którym powinni zapoznać się wszyscy socjolingwiści to *Język polski i jego odmiany* Karola Dejny [1987]. Autor zwraca uwagę na fakt, że badacze bardzo często używają terminu „język” w znaczeniu potocznym, lub w znaczeniu, do którego lepiej pasuje termin „mowa” (definicja poniżej). Według Dejny takie zachowanie mogło być akceptowalne do momentu wprowadzenia przez Saussure rozróżnienia na język (*la langue*) i jego realizację (*la parole*) w obrębie mowy (*la langage*), którą Dejna rozumie jako „zespół wszystkich zjawisk, procesów i wytworów porozumiewania się ludzi między sobą za pomocą wyrazów” [Dejna 1987: 37]. Niedopuszczalne według niego jest użycie terminu „język” w takich połączeniach wyrazowych, jak np.: „język naukowy”, „język złodziejski”, „język żeglarski” etc. „Językami” nie powinno się także nazywać stylistycznych odmian mowy.

W pracy wybrzmiewa jedno, kluczowe zdanie:

„obowiązuje w Polsce jeden system językowy – polski język”

[Dejna 1987: 41]

Odmianek polszczyzny nie powinno nazywać się „językami”. Niestety, mimo że artykuł napisany został dawno temu, to ciągle badacze mieszają terminologię, i nadużywają terminu „język”, widać, to już w tytułach artykułów np.: *Język użytkowników CB-radia a dawna gwara szoferska w kontekście przemian językowych i społeczno-cywilizacyjnych w Polsce* [Jaros 2012], *Język sportu w ujęciu leksykografów. Analiza terminograficzna wybranych słowników z terminologią sportową* [Murrmann 2015] czy *O języku kierowców autokarów* [Bartłomiejczyk 2012].

Na uwagę zasługuje także artykuł Wilkonia opublikowany rok później, w którym autor proponuje własną typologię socjolektów. Wyróżnia pięć typów, bo, jak sam twierdzi, wcześniejsze klasyfikacje są niepełne (ma tu na myśli głównie pracę Grabiasa):

1. socjolekty dużych i znaczących grup społecznych, tworzących duże skupiska na określonym terytorium, np. socjolekt środowisk górniczych, język włóknarzy łódzkich itp. (...)
2. socjolekty grup zinstytucjonalizowanych i zorganizowanych, idzie o takie grupy jak: wojsko, organizacje paramilitarne, marynarze, harcerstwo i inne (...)
3. socjolekty środowisk powstałych na zasadzie dobrowolnych zrzeszeń, towarzystw, klubów, można w tej grupie wymienić m.in. język grup sportowych język wędkarzy, taterników, a także i język łowiecki (...)
4. socjolekty środowisk przestępczych i bliżej nie określonych środowisk lumpenproletariackich, kryminogennych, tworzących tzw. margines społeczny (...)
5. socjolekty środowisk młodzieżowych; idzie o gwary i uczniowskie, studenckie, młodzieżowe grupy punków itp.

[Wilkon 1988: 90–93]

W 1987 roku Paweł Kupiszewski opublikował artykuł [Kupiszewski 1987], w którym przyjrzał się „językowi żołnierskiemu”. Autor skupił się na analizie języka tych żołnierzy, którzy odbywali dwuletnią służbę zasadniczą. Kupiszewski zaznacza, że to nie zawodowi żołnierze, a właśnie grupa poborowych tworzy „język” środowiska żołnierskiego. Dzieje się tak, gdyż jest to grupa najliczniejsza – średnio na jednego żołnierza zawodowego przypada 20–30 niezawodowych. Dzięki analizie „języka” tego środowiska, autor stwierdził, że w warstwie leksykalnej dominują takie grupy tematyczne, jak: nazwy sprzętu i uzbrojenia, nazwy osobowe, słownictwo związane z tradycją żołnierską i nazwy czynności, a także frazeologizmy. Kontynuatorem badań Kupiszewskiego był Jerzy Głowacki, który w 1990 roku opublikował artykuł *O języku żołnierskim* [Głowacki 1990], w którym zaznaczył, że ważną rolę w tym środowisku odgrywają konstrukcje wyrażające polecenie lub nakaz.

Trzy lata później, czyli w 1990 roku Ewa Kołodziejek zainteresowała się z kolei „gwarą marynarską” i opublikowała artykuł *O gwarze marynarzy i subkulturze marynarskiej*. Warto na początku omawiania tej pracy dodać, że autorka posługuje się w niej terminem „*dialektologia socjalna*”. Według niej zakres dziedziny to „gwary środowiskowe, zawodowe czy potoczny język miast” [Kołodziejek 1990], czyli definicyjnie termin *dialektologia socjalna* pokrywa się z terminem socjolingwistyka. Jak widać, autorka była świadoma genetycznych związków obu dziedzin. W początkowych fragmentach pracy autorka zaznaczyła, że badanie „*gwary marynarskiej*” jest specyficzne, ponieważ wiele jej odmian istnieje bardzo krótko (na przykład podczas jednego rejsu) Poza tego typu odmianami istnieje, wymykający się ciągłym zmianom, podstawowy zasób leksyki, który jest wspólny wszystkim wariantom „*gwary marynarzy*”. Dalej Kołodziejek charakteryzuje „subkulturę marynarską” oraz opisuje jej zwyczaje i wierzenia. Autorka podkreśla także silny wpływ języka angielskiego na to środowisko, a także, zwłaszcza wśród starszego pokolenia, wpływ języka niemieckiego (przede wszystkim w warstwie leksykalnej).

Kilka lat później, bo w 1995 roku, Kwiryna Handke publikuje dużą pracę na temat „*języka familijnego*”. Jest to, zdaniem autorki, ważny, ale dotychczas niezbadany dostatecznie, wariant polszczyzny. Postulat jego analizy wygłosił w 1917 roku J. Zborowski, jednak nie został on dostrzeżony, a temat ten "przez blisko siedemdziesiąt lat nie był kontynuowany" [Handke 1995: 15]. Dopiero w 1983, za sprawą konferencji *Cechy regionalne w języku familijnym* oraz wynikających z niej publikacji pokonferencyjnych (*Regionalizmy w języku familijnym i Polszczyzna regionalna Pomorza*), ponownie zainteresowano się „*językiem familijnym*”. Krokiem milowym w tej dziedzinie badań była dopiero monografia Kwiryny Handke *Polski język familijny* [1995], w której autorka przyjrzała się językowemu obrazowi

przestrzeni społecznej rodziny, różnym relacjom rodzinnym (np. mąż: żona czy rodzice: dzieci) a także omówiła wybrane tradycje rodzinne w Polsce. Do publikacji dodany został także obszerny aneks z tekstami, które są realizacjami języka familijnego. Przedstawiono także krótką analizę wybranego słownictwa.

Rok 1997 przyniósł publikację książkową Stanisława Grabiasa pt. *Język w zachowaniach społecznych*, która w moim odczuciu stanowi ważną pozycję wśród literatury językoznawczej, ponieważ jest podsumowaniem dotychczasowych badań w zakresie socjolingwistyki. Jest to pierwsza, znana mi praca tego typu. Grabias na początku książki przytacza terminologię socjolingwistyczną, o której nie ma dobrego zdania. Twierdzi, że obecnie panuje „mętlik terminologiczny, który wynika z braku stabilnej klasyfikacji socjolektów” [Grabias 1997: 141]. Autor zwraca także uwagę, że niestosowne jest używanie terminu „język” w bardzo szerokim i często bardzo różnym znaczeniu. Nie podoba mu się także to, że nadużywany przez badaczy jest termin „gwara”. Zjawisko to tłumaczy tym, że polscy językoznawcy, zwłaszcza w początkowej fazie zainteresowań zróżnicowaniem językowym polszczyzny, czerpali z wiedzy i badań rosyjskich językoznawców, u których zamiast socjolingwistyki, wytworzyła się dialektologia socjalna, która z kolei podlegała dialektologii i korzystała z jej terminologii. Dodatkowym atutem terminu „gwara” było to, że podkreślał on podległość względem czegoś większego, w tym przypadku wobec „języka polskiego ogólnonarodowego”. S. Grabias podziela zadanie J. Bartmińskiego, że obecnie w większości miejsc Polski „gwara” straciła status głównego środka komunikacji, a stała się wariantem, który często używany jest tylko w kontaktach rodzinnych. „Gwary” zrównały się z socjolektami, które są także podrzędne wobec „polszczyzny ogólnej”. W dalszej części pracy Grabias dokonuje przeglądu różnych socjolektów i sprawdza, jak wygląda pod względem ilościowym rozkład leksyki w danych socjolektach, czyli jakie słownictwo i z jakiego pola tematycznego dominuje w danym wariacie socjalnym.

W 1999 roku Halina Zgółkowa wydała artykuł pt. *Język subkultur młodzieżowych*. Autorka twierdzi, że środowisko młodzieżowe, decydując się na stworzenie danej subkultury, dokonuje wewnętrznego podziału według rodzaju słuchanej wśród młodzieży muzyki (np. metal, disco-polo czy rock) a także zainteresowań konkretnymi ruchami religijnymi (np. satanizmem czy sektami powiązаныmi z buddyzyzmem lub hinduizmem). Subkultury młodzieżowe prawie zawsze są ze sobą w konflikcie. Demonstrowanie wrogości do innych grup widać w ich quasi-twórczości, która często prezentowana jest na murach budynków przy użyciu graffiti. Ich teksty są zazwyczaj bardzo wulgarne, ale często też podszyte ironią lub

sarkazmem. Do zabiegów językowych charakterystycznych dla subkultur młodzieżowych H. Zgólkowa zalicza także częste używanie superlatywów.

W 2004 roku Małgorzata Kasperczak publikuje artykuł *Język kobiet – język mężczyzn* i tym samym rozpoczyna zainteresowanie wśród polskich badaczy biolektami (zwłaszcza kobiet i mężczyzn). Autorka w swoich badaniach poszukiwała odpowiedzi na pytanie: czy istnieją różnice między współczesnymi biolektami języka polskiego. Aby je znaleźć przeprowadziła badanie ankietowe, w którym wzięło udział 200 respondentów. Wyniki pokazały, że istnieją „różnice między językiem kobiet a językiem mężczyzn, które mogłyby mieć podłoże biologiczne i być wyznacznikiem biolektów współczesnego języka polskiego” [Kasperczak 2004 : 72]. Podobne badania zostały przeprowadzone przez M. Karwatowską i J. Szpyrę-Kozłowską, a następnie opublikowane w obszernej pracy poświęconej zróżnicowaniu językowemu kobiet i mężczyzn – *Lingwistyka płci. Ona i on w języku polskim* [2005]. Autorki także przeprowadziły ankietę, z tą różnicą, że nie wśród 200 losowych respondentów, ale w dobranej grupie badawczej, w której było 80 uczniów szkół średnich (40 chłopców i 40 dziewczyn). Badania obejmowały takie zagadnienia, jak: skojarzenia z rzeczownikami *chłopak* i *dziewczyna*, łączliwość tych wyrazów z przymiotnikami, określenia równoważne lub bliskie tym rzeczownikom oraz zdefiniowanie wyrazów. Z badań wynika, że dziewczęta posiadają znacznie bogatsze słownictwo, ponieważ podały dwukrotnie więcej skojarzeń niż chłopcy. Także więcej wymieniły przymiotników, które łączą się z rzeczownikami *chłopiec* i *dziewczynka* oraz synonimów do nich. W definicjach respondentów zaobserwowano, że bardziej cenią oni przedstawicieli swojej płci, natomiast z lekceważeniem odnoszą się do płci przeciwnej. Badania po raz kolejny pokazały, że istnieją różnice w języku *chłopców* i *dziewczynek*. Dalszą część książki autorki poświęciły takim tematom, jak: seksizm w języku polskim czy tzw. „niewidzialność kobiet” we współczesnych tekstach (np. regulaminach, przepisach czy ustawach).

W 2008 roku Tomasz Piekot wydał publikację *Język w grupie społecznej. Wprowadzenie do analizy socjolektu*, w której zebrał wszystkie informacje dotyczące polskich badań socjolektalnych i ustosunkował się do nich. Była to druga, po Stanisławie Grabiasie (patrz strona poprzednia), tego typu publikacja. W głównej części pracy autor wyszczególnia, na jakich zasadach do tej pory były wydzielane poszczególne odmiany językowe. Na przestrzeni lat podczas tworzenia klasyfikacji uwzględniano takie kryteria, jak:

1. Kryterium ograniczania dostępu. Biorąc pod uwagę to kryterium, wyróżnia się odmiany językowe, które dostępne są wszystkim użytkownikom języka polskiego oraz takie, do których mają dostęp tylko wybrani ludzie.

2. Kryterium terytorialne. Różnicuje się tutaj socjolekty w zależności od tego, na jakim podłożu powstały (np. według Klemensiewicza „języki zawodowe” powstały na gruncie „języka ogólnego”, A. Furdal zaś rozróżnia „język zawodowy miejski” od „wiejskiego”).

3. Kryterium semantyczne. W kryterium tym ważny jest stopień precyzji znaczeniowej (np. w klasyfikacji odmian polszczyzny A. Furdala podzielono „języki wyspecjalizowane” na: „naukowe” (w praktycznym, potocznym użyciu) i „zawodowe”).

4. Kryterium socjologiczne. Uwzględniając to kryterium, dokonuje się charakterystyki danej grupy na podstawie cech charakterystycznych, zwłaszcza pozajęzykowych (np. T. Skubalanka w obrębie „mówionego ogólnopolskiego stylu potocznego” wyróżnia takie podgrupy, jak: „j. rodzinny”, „j. biologiczny”, „j. zawodowy” i „j. socjalny”).

5. Kryterium tajności. W tym kryterium ważna jest opozycja *tajność* i *jawność*. Niektóre grupy celowo czynią swój „język” tajnym, aby informacje były niedostępne dla osób postronnych (przykładem może być „gwara więzienna”. Więcej przykładów znaleźć można w klasyfikacji Grabiasa, patrz strony 26–27).

6. Kryterium funkcjonalne. Klasyfikuje się tutaj socjolekty, ze względu na funkcje, jakie pełnią (zaliczyć tu można np. „socjolekty ekspresywne”).

Po wymienieniu powyższych kryteriów i krótkim omówieniu ich, autor dodaje, że klasyfikacja odmian polszczyzny może opierać się na jednym tylko kryterium, ale może wykorzystywać ich także wiele.

T. Piekot proponuje, aby wprowadzić do socjolingwistyki dodatkowe kryterium, takie jak „*stosunek grupy do rzeczywistości*”. Chodzi tutaj o postawę danej grupy do otaczającej ją rzeczywistości. Grupa może akceptować istniejące w społeczeństwie normy, zakazy czy też nakazy (T. Piekot nazywa taką grupę „*socjolektem tradycyjnym*”), może okazywać bierność wobec zasad społecznych (wtedy mówimy o „*socjolektach ludycznych*”) oraz może być do nich nastawiona negatywnie („*socjolekty kontestacyjne*”). Wyżej omówione fragmenty pracy Piekota stanowią jej trzon, natomiast w drugiej części publikacji, autor pokazuje, jak może wyglądać analiza socjolektu. Autor postuluje, aby początkowo ustalić stan badań nad socjolektem, następnie sprawdzić sposoby komunikacji w grupie, a dopiero po tym przejść do analizy językowego obrazu świata i słownictwa badanego socjolektu.

W tym samym roku, czyli w 2008 (9–11 października) odbyło się VII Forum Kultury Słowa, którego głównym tematem były „języki zawodowe i środowiskowe”. Pokłosiem tego wydarzenia była wydana dwa lata później publikacja, w której wydane zostały artykuły prelegentów oraz spisane dyskusje panelowe. Tadeusz Zgółka [2010], jeden z prelegentów, przedstawił parametry, które pomagają w odróżnianiu od siebie pewnych odmian językowych.

Tymi parametrami są: rodzaj kanału, rodzaj kontaktu oraz rodzaj sensu komunikacyjnego. Autor pokazał, jak po uwzględnieniu tych parametrów różnicuje się „język fachowy” od „żargonu” i „gwary środowiskowej”. Autor odniósł się także do poprzednich klasyfikacji i dostrzegł, że żadna z nich nie uwzględniła jednego ważnego wariantu polszczyzny, a mianowicie języka migowego. Ponadto autorzy poprzednich podziałów starali się w większości utrzymać podział dychotomiczny, a według Zgółki, usilne obstawanie przy takim podziale może prowadzić do pewnych niekonsekwencji.

Dwa lata później Barbara Pędzich opublikowała artykuł (*Jak powstaje środowiskowa odmiana języka (na przykładzie leksyki paralotniarskiej)* [2010]), w którym pokazała, jak powstaje socjolekt, a zobrazowała to przykładami leksyki paralotniarskiej. Publikacja ta była zaczątkiem do większej pracy pt. *Jak powstaje socjolekt? Studium słownictwa paralotniarzy* [2012], w której autorka przedstawiła analizę językową tytułowego socjolektu i stwierdziła, że dominują w nim zapożyczenia, zarówno wewnętrzne, jak i zewnętrzne oraz neologizmy i frazeologizmy.

Badania socjolektów cieszą się obecnie coraz większą popularnością wśród językoznawców. Widać to po zwiększającej się liczbie publikacji na ten temat. Pozytywnie należy przyjąć tendencję do stabilizacji odrębnej terminologii socjolektalnej. Cieszy mianowicie fakt, że coraz więcej osób decyduje się na wykorzystanie w swojej pracy terminów takich, jak „socjolekt” lub „profesjolekt”, zamiast określeń typu „żargon” czy „gwara”.

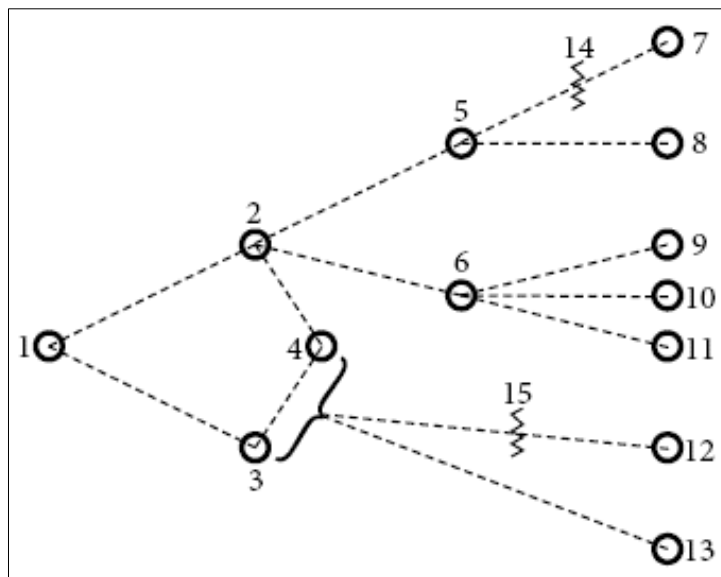
W powyższych dwóch częściach pominięto ważny element związany z badaniami nad odmianami językowymi polszczyzny, czyli ich różne klasyfikacje. Temu zagadnieniu zostanie poświęcony następny fragment pracy.

2.3. KLASYFIKACJE ODMIAN JĘZYKOWYCH POLSZCZYZNY

Na kanwie wielu badań dialektologicznych i socjolingwistycznych powstawały liczne klasyfikacje odmian polszczyzny. Zaczęto je sporządzać już w połowie XX wieku.

Pierwszą propozycję podziału wariantów języka polskiego opublikował Zenon Klemensiewicz w pracy *O różnych odmianach współczesnej polszczyzny* [1953].

Schemat 1. Przegląd odmian językowych Z. Klemensiewicza



- | | | |
|------------------|---------------------|-------------------------------|
| 1. j. narodowy | 6. j. literacki | 11. j. normatywno-dydaktyczny |
| 2. j. ogólny | 7. j. konwersacyjny | 12. j. reg. konwersacyjny |
| 3. j. regionalny | 8. j. monologowy | 13. j. reg. artystyczny |
| 4. gwara miejska | 9. j. artystyczny | 14. języki zawodowe |
| 5. j. potoczny | 10. j. naukowy | 15. gwary zawodowe |

Źródło: Klemensiewicz 1953: 42²⁰

Z uwagi na to, że jest to pierwsza tego typu klasyfikacja, doczekała się wielu omówień, a każdy kolejny badacz tego tematu musiał się do niej ustosunkować.

Witold Taszycki i Tadeusz Milewski we wspólnym artykule zarzucają autorowi powyżej zaprezentowanej klasyfikacji, że niepotrzebnie wprowadza nową terminologię, jest ona ich zdaniem niepotrzebna, bo lepiej byłoby użyć terminologii tradycyjnej, np. zamiast wprowadzać określenie „język ogólny” pozostać przy „języku literackim”, a zamiast terminu „język potoczny” lepiej użyć „forma mówiona” [Taszycki, Milewski 1956: 408]. Rzeczywiście współcześnie nie używa się wielu zaproponowanych przez Klemensiewicza terminów, ale bardzo trafne było wprowadzenie do klasyfikacji terminu „język ogólny”, który stał się jednym z podstawowych w polskim językoznawstwie. Wątpliwości autorów (Milewskiego i Taszyckiego) budzi także brak konkretnej definicji terminu „język”²¹, który przez Klemensiewicza używany jest w wielu różnych znaczeniach, na przykład w takich

²⁰ Zdecydowałem się w tym miejscu zaprezentować jedynie klasyfikację Klemensiewicza, pominąłem natomiast podobną, opublikowaną dwa lata później klasyfikację Urbańczyka [1956: 24], ponieważ, jak pisze A. Furdal [1973: 8] te klasyfikacje są ze sobą niemal zbieżne, gdyż autorzy mieli podobne stanowisko wyjściowe. Ponadto klasyfikacja Klemensiewicza była pierwszą tego typu pracą w polskim językoznawstwie, dlatego uznałem, że warto ją przedstawić w tekście głównym mojej pracy.

²¹ Taszycki i Milewski w artykule definiują „język”, jako „system norm regulujących komunikację między jednostkami, system złożony z czterech mniejszych systemów częściowych, a mianowicie systemu fonologicznego (forma dźwiękowa), słownikowego (wyrazy), morfologicznego (afiksy) i składniowego (środki budowy zdania)” [Taszycki, Milewski 1956: 409].

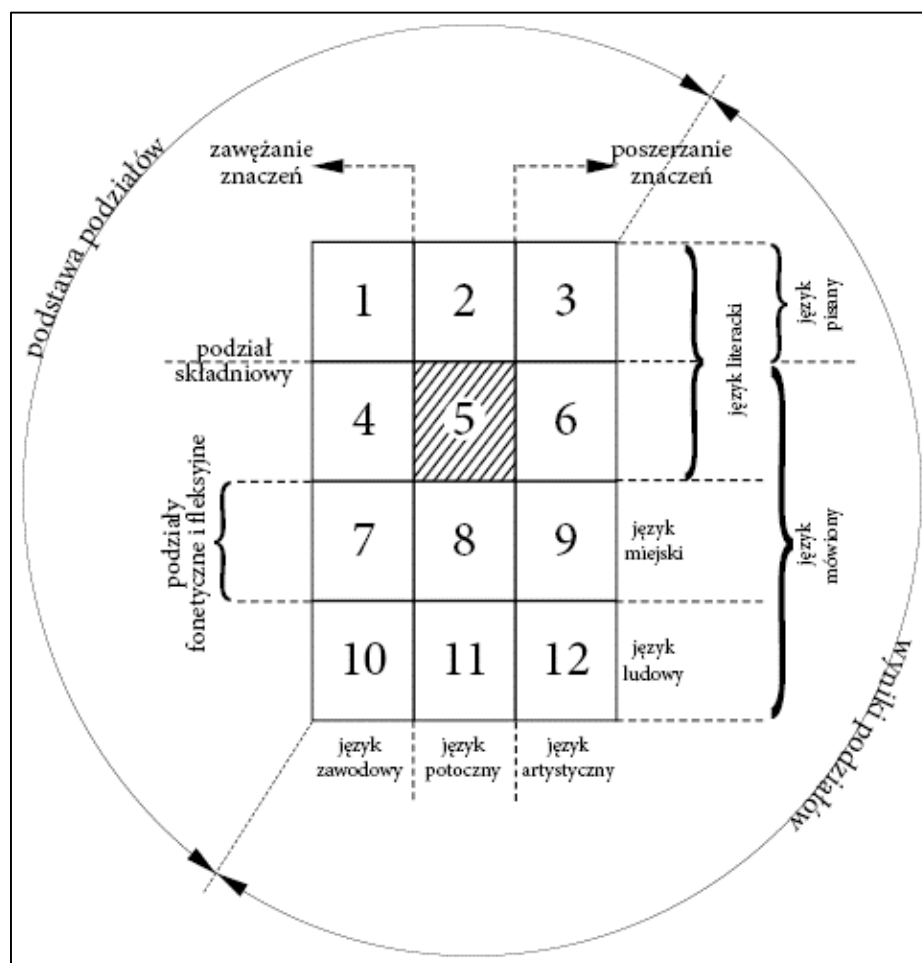
połączeniach wyrazowych, jak: „język ogólny”, „język polski narodowy” czy „język regionalny”. Według nich, nie można pisać w „języku polskim narodowym”, a jedynie w jednej z jego odmian. Zalecają, zamiast niezbyt konkretnego terminu, jakim jest „język polski narodowy”, używać „polska grupa dialektyczna” [Taszycki, Milewski 1956].

Teresa Skubalanka podkreśla, że błędem tej klasyfikacji jest to, że nie uznaje nadrzędności stylu mówionego i pisanego względem innych [Skubalanka 1976: 261], natomiast Danuta Buttler [1982: 17] zaznacza, że autor niepotrzebnie utożsamia ze sobą tak różne terminy, jak „język potoczny” i „język mówiony”. Dalej autorka podaje przykłady, które realizują się w mowie, ale nie są reprezentantami polszczyzny potocznej. Są to: wykład, przemówienie publiczne czy dyskusja naukowa.

Zalety tej klasyfikacji wymienia Antoni Furdal [1973: 8], który zaznacza, że publikacja tej klasyfikacji to ważna innowacja metodologiczna, ponieważ połączono w jedną całość zakresy badawcze stylistyki z dialektologią regionalną i socjalną.

W 1973 zostaje opublikowane, ważne dla mojej pracy, dzieło językoznawstwa polskiego, jaką była klasyfikacja odmian polszczyzny Antoniego Furdala.

Schemat 2. Klasyfikacja odmian współczesnego języka polskiego A. Furdala



- 5 – język potoczny literacki mówiony; najważniejsza społecznie odmiana językowa, ośrodek naszego układu i miernik w odniesieniu do innych odmian,
- 2 – język potoczny literacki pisany,
- 8 – język potoczny miejski, oczywiście mówiony,
- 11 – język potoczny ludowy, także mówiony.

Pion wyspecjalizowany:

- 1 – język naukowy pisany,
- 4 – język naukowy mówiony,
- 7 – język zawodowy miejski,
- 10 – język zawodowy wiejski.

Pion artystyczny:

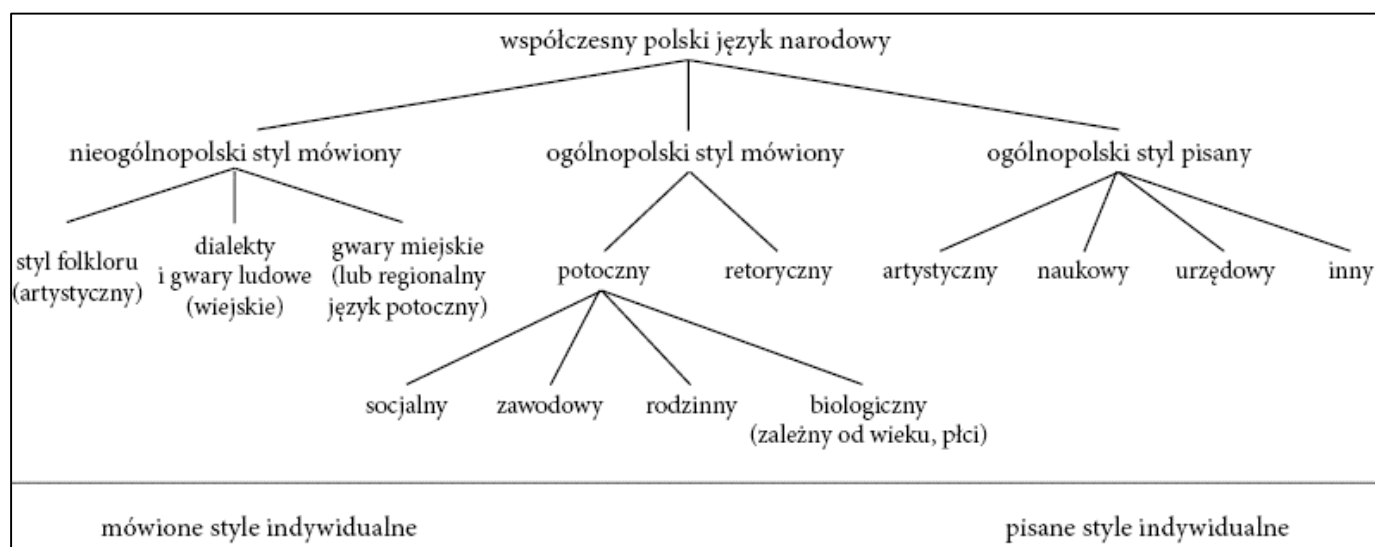
- 3 – język literacki artystyczny prymarnie pisany,
- 6 – język artystyczny na szczeblu literackim, ale mówiony, np. język improwizacji poetyckiej; w zasadzie tożsamy z pisaniem, a więc jako osobna odmiana nie istnieje,
- 12 – język artystyczny ludowy,
- 9 – język artystyczny miejski; w warunkach polskich chyba nie istnieje jako osobna odmiana, albo istnieje tylko w zaczątkach.

Źródło: Furdal 1973: 54

Podobnie jak poprzednia klasyfikacja tego typu (patrz strona 34), także i ta doczekała się wielu nawiązań w innych pracach oraz była przedmiotem dyskusji. Badacze w kolejnych pracach odnosili się do niej. Danuta Buttler na przykład zwraca uwagę autorowi klasyfikacji, że za cechę konstytutywną „*języka potocznego*” błędnie uważa pośredniczenie w codziennych kontaktach pomiędzy członkami danej społeczności [Buttler 1982: 18]. Według niej nie wszystkie codzienne rozmowy odbywają się z wykorzystaniem „*języka potocznego*”, bo np. dialogi petenta z urzędnikiem czy też pacjenta z lekarzem często wychodzą poza granice potoczności i są bardziej formalnymi realizacjami danego wariantu polszczyzny. Aleksander Wilkoń podkreśla oryginalność tej klasyfikacji, która polega na wykorzystaniu kryteriów semantyczno-leksykalnych oraz gramatycznych (składni, fleksji i fonologii). Wykorzystane zostały one przy rozróżnianiu takich odmian, jak: „*język artystyczny*”, „*potoczny*” czy „*naukowy*”. Badacz zarzuca natomiast autorowi, że zatarta została hierarchia pewnych odmian, a także, że występują tzw. „*problematyczne odmiany*” np. „*język artystyczny literacki mówiony*”. Dodaje także, że Furdal pominął kilka ważnych wariantów, takich jak np. „*język administracyjny*” [Wilkoń 1987: 18.]. W klasyfikacji tej na uwagę zasługuje wyróżnienie, jako najważniejszego z „*języków*”, „*języka potocznego literackiego mówionego*”. Jest to moim zdaniem najważniejsza zasługa tej klasyfikacji.

Kolejna klasyfikacja odmian polszczyzny została zaprezentowana przez Teresę Skubalanek w artykule *Założenia analizy stylistycznej*, który ukazał się w 1976 roku.

Schemat 3. Klasyfikacja odmian współczesnej polszczyzny T. Skubalanki



Źródło: Skubalanka 1976: 261

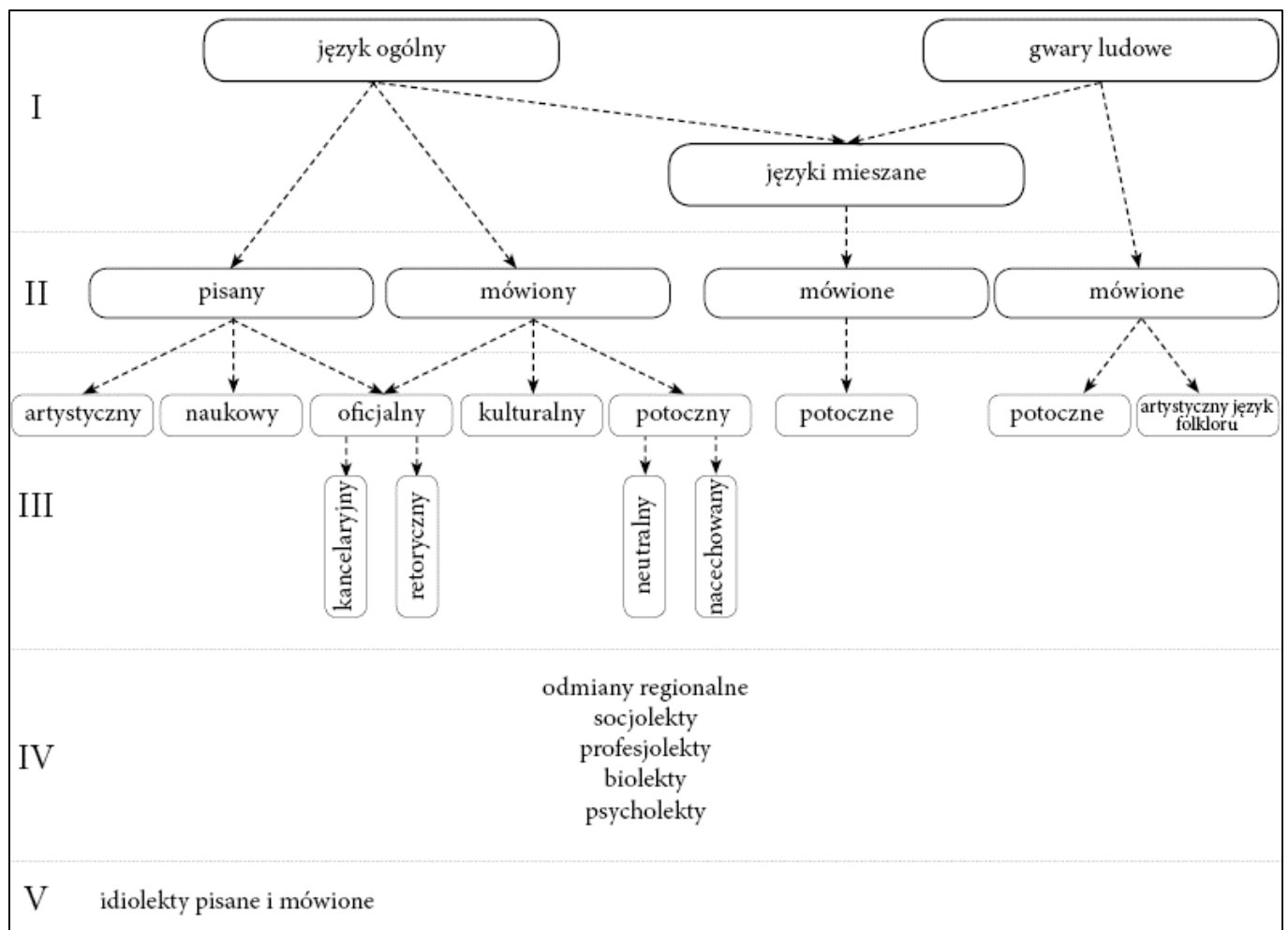
Autorka zaprezentowanej powyżej klasyfikacji zdecydowała się na wprowadzenie do niej terminu *styl*, który według niej oznacza po prostu konkretyzację języka [Skubalanka 1976: 254]. Warto podkreślić, że decydując się na ten termin, unika ona wielokrotnego używania, jak to miało miejsce w przypadku poprzednich klasyfikacji, określenia „język”, które to w tej klasyfikacji zostało użyte jedynie do nazwania wariantu zawierającego w sobie wszystkie pozostałe – „współczesnego polskiego języka narodowego” oraz w utartym połączeniu wyrazowym – „język potoczny”.

Nowością względem wcześniejszych klasyfikacji jest także to, że jako główne rozróżnienie wariantów polszczyzny występuje tu styl mówiony i styl pisany, dopiero w następnej kolejności usytuowane zostały inne style. Niepokojąca, ze względu na brak konkretyzacji lub chociażby podania kilku przykładów, jest natomiast kategoria „inne”, bo tak naprawdę nie wiadomo, co może ona zawierać.

W roku 1987 ukazała się książka Aleksandra Wilkonia – *Typologia odmian językowych współczesnej polszczyzny*, w której poza samą typologią autor wprowadził nowe terminy. Wiele z nich przyjęło się w polskim językoznawstwie. Wilkoń zauważył, że „język” jest połączeniem trzech czynników: społecznego, biologicznego i psychicznego. Podczas prac badawczych nad konkretną odmianą języka wymaga się oddzielenia od siebie tych czynników, a co się z tym wiąże, wyodrębnienia trzech odmian językowych: *socjolektu*, *biolektu* i *psycholektu*. Wilkoniowi zawdzięczamy wprowadzenie do polszczyzny dwóch ostatnich terminów. Autor przywołuje także typologię Teresy Skubalanki, która umieściła w niej „styl

biologiczny”, czyli odmianę językową uzależnioną od takich parametrów, jak wiek i płeć. Dla Wilkonia lepszym określeniem na „styl biologiczny” jest krótki *biolekt*. Natomiast za *psycholekty* Wilkoń uważa te realizacje językowe, w których odbijają się pewne stany chorobowe (głównie psychiczne) np. schizofrenia. W dalszej pracy autor wprowadza także pojęcie *profesjolektu*, a jest to „potoczność + słownictwo specjalne” [Wilkoń 1987: 101]. Tomasz Piekot przypisuje także Wilkoniowi wprowadzenie do języka polskiego terminu *socjolekt* [Piekot 2008: 31], ale jak już wcześniej wykazałem, to nie Wilkoń pierwszy wprowadził go do polszczyzny (patrz strona 25). Elementem najważniejszym tej publikacji jest autorska typologia odmian współczesnej polszczyzny.

Schemat 5. Typologia odmian językowych współczesnej polszczyzny A. Wilkonia

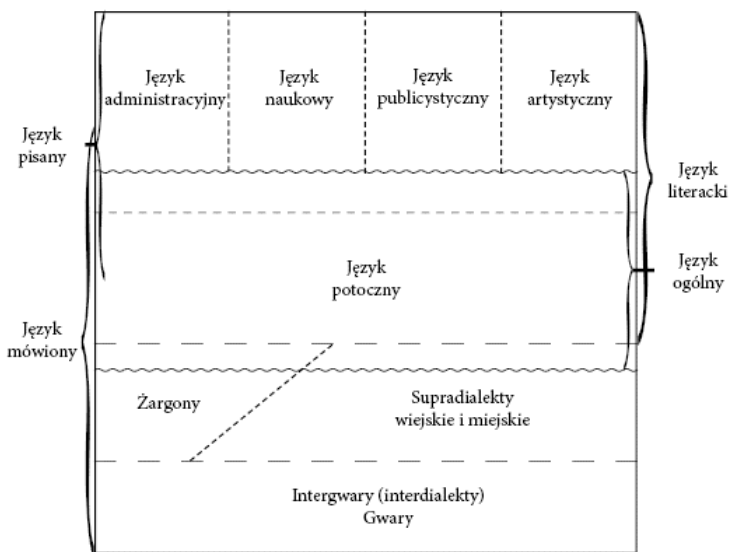


Źródło: Wilkoń 1987: 111

Autor, identycznie jak Teresa Skubalanka, jako drugi stopień podziału wprowadził rozróżnienie na warianty pisane i mówione oraz oddzielił od części głównej klasyfikacji *idiolekty*.

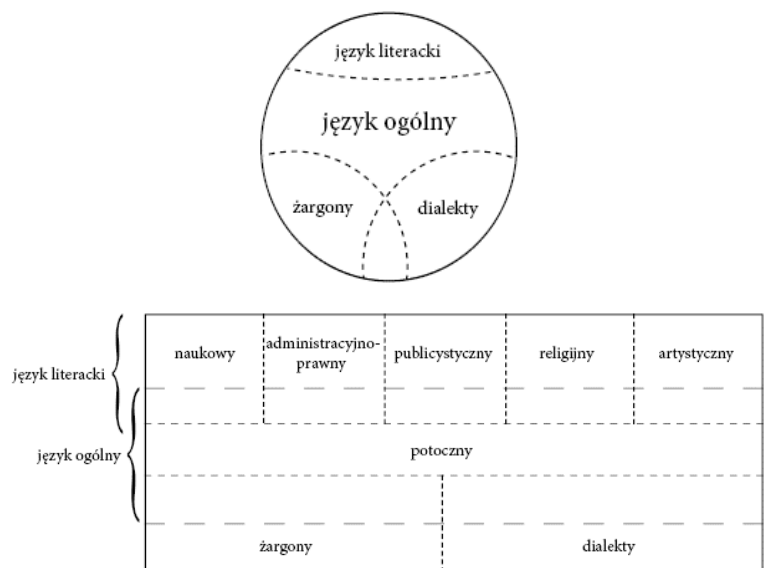
Jedną z nowszych klasyfikacji odmian współczesnej polszczyzny opracował Stanisław Gajda. Została ona opublikowana w tomie z serii *Najnowsze dzieje języków słowiańskich* wydanym w 2001 roku (patrz schemat 7).

Schemat 6. Pierwsza klasyfikacja S. Gajdy



Źródło: Gajda 1982: 54

Schemat 7. Druga klasyfikacja S. Gajdy



Źródło: Gajda 2001: 212

Kluczowym fragmentem artykułu, w którym opublikowana została klasyfikacja, jest według mnie stwierdzenie, że „język literacki przeznaczony do zaspokojenia „wyższych” potrzeb społecznych egzystuje w cieniu języka ogólnego” [Gajda 2001: 212]. Autor tym samym opowiada się za współistnieniem tych terminów, które przez wielu badaczy na przestrzeni lat były używane wymiennie na oznaczanie tej samej odmiany polszczyzny. Według Gajdy, uzasadnione jest użycie obu terminów, ponieważ istnieją dwie normy skodyfikowane: wyższa (reprezentowana przez „język literacki”) oraz niższa (przez „język potoczny” lub szerzej przez „język ogólny”).

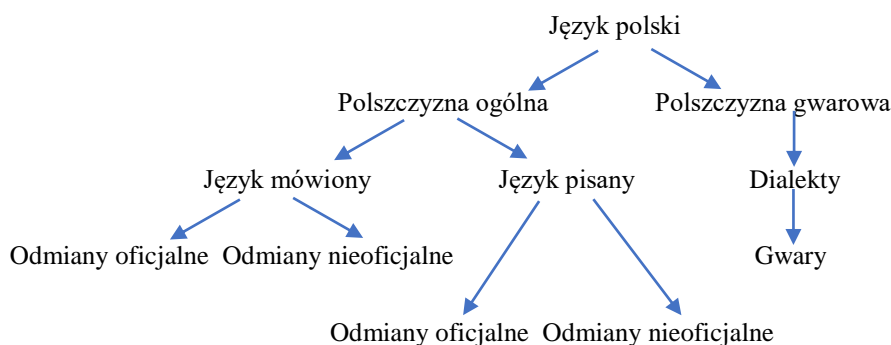
Na schemacie numer 7 widać, że S. Gajda podziela zdanie wielu badaczy, iż w centrum polszczyzny jest „język potoczny”, który stanowi także jądro „języka ogólnego”, do którego zaliczone zostały także niektóre „żargony”, „dialekty”, a także pewne realizacje takich „języków”, jak: „naukowy”, „administracyjno-prawny”, „publicystyczny”, „religijny” czy

„artystyczny”. Grupa tych „języków” została przez Gajdę, przynajmniej w większej części, zaklasyfikowana jako reprezentacja „języka literackiego”.

Dla porównania powyżej przytoczyłem także wcześniejszą klasyfikację odmian polszczyzny autorstwa S. Gajdy, która opublikowana została w 1982 roku (patrz schemat 6). Widać, że jest ona w pewnych aspektach nieco bardziej skomplikowana, ponieważ autor wprowadził np. rozróżnienie na „język mówiony” i „pisany”, czego w późniejszej klasyfikacji nie ma. Starsza klasyfikacja jest także nieco uboższa od nowszej, ponieważ autor nie uwzględnił w niej np. „języka religijnego”. Reasumując, stwierdzić trzeba, że klasyfikacja Gajdy przez kilkanaście lat znacznie ewoluowała. Autor zrewidował pewne poglądy na temat języka polskiego, a część pomysłów po prostu przeniósł w niezmienionej wersji do nowszej klasyfikacji. Wątpliwości mogą budzić niektóre, użyte przez Gajdę w nowej klasyfikacji, terminy. Autor zamiast użyć współczesnej terminologii, decyduje się na terminy, które wychodzą z użytku, np. zamiast „socjolektów” (o których, co ciekawe, autor pisze w swoim artykule) zdecydował się na „żargony”. Nadmiernie także eksploatuje termin „język”.

W dalszej części pracy autor porusza sprawę istnienia „dialektów” w obecnej polszczyźnie i stwierdza, że zmieniły one swoją postać, nie wymarły, ale stały się nowymi funkcjonalnymi odmianami polszczyzny. Nastąpiło tutaj swoiste połączenie dziedzin dialektologii i socjolingwistyki. Dodaje także, że minęły czasy, w których ludzie spędzali całe swoje życie w jednym miejscu, a dzisiejszy człowiek często zmienia miejsce swego pobytu i, co się z tym łączy, zmienia się także jego „język” oraz grupa społeczna, do której należy. Gajda także dodaje, że w tych czasach „zaznacza się tendencja do zacierania ostrych strukturalnych granic międzyodmianowych i do mieszania się odmian” [Gajda 2012: 208–209].

Ostatnia ze znanych mi klasyfikacji odmian językowych polszczyzny została przedstawiona w *Wielkim słowniku poprawnej polszczyzny PWN* pod redakcją A. Markowskiego. Redaktor słownika wyróżnił dwa kryteria podziału polszczyzny: ze względu na zasięg użycia oraz ze względu na użytkowników. Ze względu na pierwsze kryterium podział wygląda następująco:



Ze względu na kryterium użytkowania podział ulega dalszemu uszczegółowieniu:



Schemat 8. Klasyfikacja A. Markowski

Źródło: Markowski 2006: 1586–1591

Przyglądając się propozycji podziału polszczyzny według A. Markowskiego można zauważyć, że (podobnie jak w klasyfikacjach innych badaczy) autor nadużywa terminu „język”. Ponadto stosuje termin „gwara” w odniesieniu do tego, co dzisiaj raczej nazywane jest *socjolektem* lub *profesjolektem*.

Pomijając jednak stosowanie przestarzałej terminologii, należy stwierdzić, że typologia odmian polszczyzny zaproponowana w NSPP osiągnęła stadium modelowania najbliższe rzeczywistemu odczuciu użytkowników. Stało się tak dzięki zhierarchizowaniu poziomów podziału i zróżnicowaniu kryteriów.

3. BADANIA ILOŚCIOWE W POLSCE

W mojej pracy ważną rolę odgrywają badania korpusowe, które w linii prostej wywodzą się z szeroko rozumianych badań ilościowych. Początki badań ilościowych były trudne, gdyż naukowcy nie dysponowali podstawowymi narzędziami ułatwiającymi i przyspieszającymi pracę, takimi, jak chociażby komputer.

Poniżej przedstawiam chronologiczny rozwój badań ilościowych i statystycznych, które doprowadziły w rezultacie do wykształcenia się odrębnej gałęzi językoznawstwa, zwanej obecnie badaniami korpusowymi.

Za pierwszą polską pracę językoznawczą związaną ze statystyką²², a konkretnie ze statystyką leksykalną, uważa się artykuł Witolda Lutosławskiego *Sur une nouvelle méthode pour déterminer la chronologie des dialogues de Platon* opublikowany w 1896 roku w Paryżu. Informację o pierwszeństwie tego artykułu podaje w jednej ze swoich prac Jadwiga Sambor. Artykuł Lutosławskiego był próbą zbadania dzieł Platona w celu ustalenia ich chronologii. Autor wykorzystał statystykę do zbadania leksyki używanej przez filozofa w kolejnych utworach, przy założeniu, że w początkowych pracach była ona uboższa, natomiast w ostatnich jego tekstach bogatsza. Artykuł w całości napisany został po francusku. Nasuwa się więc pytanie, czy pracę Lutosławskiego można uznać za pierwszy **polski** językoznawczy artykuł, w którym wykorzystano statystykę? Nie budzi wątpliwości, zakładając prawdziwość twierdzenia Jadwigi Sambor, że Lutosławski był pierwszym polskim **badaczem**, który połączył statystykę z badaniami nad językiem. Moim zdaniem jednak jego pracy nie można uznać za pierwszy **polski** językoznawczy artykuł, w którym użyto statystyki, ponieważ po pierwsze – nie był on napisany po polsku, co moim zdaniem jest już pierwszą przesłanką, aby zastanowić się nad prawdziwością twierdzenia o pierwszeństwie, po drugie – nie dotyczył polskiej problematyki, nie nawiązywał do polskich realiów, tylko do greckich (jak powszechnie wiadomo, Platon był filozofem pochodzącym z Aten).

Pierwsze typowo polskie opracowania, w których uwzględniono badania ilościowe, znaleźć można dopiero w pracach Witolda Doroszewskiego, który m.in. zajmował się

²² Takie informacje można znaleźć w pracy Jadwigi Sambor *Badania statystyczne nad słownictwem (na materiale „Pana Tadeusza”)* [1969].

dialektologią, a z tą dziedziną badawczą silnie łączy się statystyka. W 1934 r. badacz opublikował artykuł, w którym scharakteryzował „mowę” mieszkańców wsi Staroźreby²³.

Jednym z pierwszych autorów, którzy wykorzystali statystykę nie tylko jako element dodatkowy do badań, uzupełnienie ich, ale jako główne narzędzie badawcze był Jan Czekanowski, który w 1947 roku opublikował monografię *Polska synteza slawistyczna w perspektywie ilościowej*. Czekanowski był antropologiem, więc pisząc monografię, w której dominują zagadnienia językowe, musiał skorzystać z pracy jakiegoś znawcy tematu. Zdecydował się wykorzystać wyniki badań Tadeusza Lehra-Spławińskiego, który w kilku swoich pracach zaproponował pogrupowanie różnych rodzin językowych z uwzględnieniem występujących w nich identycznych procesów gramatycznych. W efekcie powstał wykaz zjawisk językowych (gramatycznych), które wspólne są dla różnych grup rodzin językowych [Lehr-Spławiński 1924, 1946]. Czekanowski do swoich badań z tej listy wybrał tylko te elementy gramatyczne, które występują w „*narzeczach balto-słowiańskich*”, a następnie skonfrontował je z innymi „*narzeczami*”, a także z całą grupą języków indoeuropejskich. W rezultacie tych badań powstało wiele zestawień tabelarycznych, które prezentują różnice, ale także i punkty wspólne w budowie językowej poszczególnych rodzin wyrazowych lub też pojedynczych języków.

W połowie XX w. dużym problemem było to, w jaki sposób badać słownictwo w długich lub wielu różnych tekstach, np. w całej twórczości danego pisarza, aby można było odpowiednio dobrze scharakteryzować leksykę występującą w tekstach. Problem ten postanowił rozwiązać Władysław Kuraszkiewicz, kiedy pracował w zespole 30 osób nad *Słownikiem polszczyzny XVI wieku*. Na jednym ze spotkań grupy badawczej, po tym, jak dokonał odpowiednich obliczeń, podzielił się swoim spostrzeżeniem, że skatalogowanie całego słownictwa tekstów tego okresu, zajęłoby im mniej więcej 20 lat. Zaproponował więc, aby wybrać do analizy tylko 20% całego materiału, czyli zbadać i opracować co piątą stronę danego tekstu. Propozycję swoją umotywował, przytaczając wyniki badań nad *Biblią Zofii* autorstwa M.R. Mayenowej, która wykazała, że z 20% tekstów wyzyskać można około 60% niepowtarzalnych wyrazów, które wykorzystane są w tekście. Zespół zgodził się, aby zastosować tę metodę w początkowych etapach pracy nad słownikiem, z tym zastrzeżeniem, że

²³ Jak już wspominałem, autor wykorzystał w artykule badania ilościowe. Wcześniej zebrał materiał fonetyczny podczas rozmów z mieszkańcami. Ze wszystkich mieszkańców Staroźrebów tylko pięcioro ostatecznie uczestniczyło w badaniu. Doroszewski ograniczył się do tej liczby, ponieważ gdyby zebrał większy materiał, to jak sam pisze „rzecz przybrałaby takie rozmiary, na jakie mi niestety czas nie pozwala. Pozostają zatem na materiale ograniczonym” [Doroszewski 1934: 250]. Autor w pracy scharakteryzował poszczególne cechy wymowy każdego z mieszkańców, a następnie wyciągnął ogólne wnioski dotyczące ich mowy.

małe i ważne dla polszczyzny teksty uwzględniane będą w całości lub przynajmniej w 50%. Po wykonaniu dodatkowych obliczeń, na podstawie analizy słownictwa takich tekstów, jak *Wizerunek własny żywota człowieka poczciwego* i *Figliki* Mikołaja Reja oraz *Fraszki* Jana Kochanowskiego, Kuraszkiewicz poszedł jeszcze krok dalej i zaproponował, aby teksty:

- do 100.000 liter badać w całości,
- od 100.000 do 200.000 liter badać w 50% (co 2 stronę),
- od 200.000 do 300.000 liter badać w 33% (co 3 stronę),
- a dłuższe niż 300.000 liter badać w 20% (co 5 stronę)

[Kuraszkiewicz, Łukaszewicz 1951: 15]

Powyżej przedstawione zasady ostatecznie przyjęto podczas początkowych prac nad *Słownikiem polszczyzny XVI wieku*.

Kuraszkiewicz nie poprzestał na zaproponowanej powyżej metodologii pracy z tekstami, dalej interesował się statystyką i rozwijał swoje badania nad słownictwem. Najważniejszy rezultat jego prac został opublikowany w artykule z 1958 roku pt. *Statystyczne badania słownictwa polskich tekstów XVI wieku*. W pracy tej na przykładzie większej liczby tekstów niż w przytoczonym powyżej artykule, autorstwa takich osób, jak np. Mikołaj Sęp-Szarzyński, Łukasz Górnicki, Biernat z Lublina, Sebastian Klonowic, Mikołaj Rej i Jan Kochanowski, pokazał, w jaki sposób można możliwie obiektywnie ocenić obfitość słownictwa tychże pozycji. Z tym jednak zastrzeżeniem, że metoda ta lepiej sprawdza się przy tekstach długich. Kuraszkiewicz wykorzystał w niej stworzony przez siebie wzór na wyliczenie wskaźnika k (od pierwszej litery nazwiska Kuraszkiewicz):

$$y=k\sqrt{x}$$

gdzie x , to liczba liter w całym tekście, a y to liczba haseł, czyli niepowtarzalnych użyć danych wyrazów. W artykule na wielu przykładach udowodnił, że owa metoda działa. I tak, np. *Rytmy* Sępa-Szarzyńskiego, które zbudowane zostały z 29.000 liter, a użyto w nich 1.700 niepowtarzalnych wyrazów (haseł) po dokonaniu obliczeń, mają współczynnik k na poziomie 10,0. Po zestawieniu tego tekstu np. z *Trenami* Kochanowskiego, które miały współczynnik k na poziomie 8,8, widać, że Sęp-Szarzyński operował w tym utworze bogatszym słownictwem niż Kochanowski w *Trenach*. Kuraszkiewicz w artykule potwierdza także tezę, oczywiście przy wykorzystaniu statystyki, że teksty wierszowane operują bogatszym słownictwem niż prozaiczne, a wymuszone jest to przez ich formę.

Zaproponowana przez Kuraszkiewicza metoda badawcza była punktem przełomowym w historii polskiej statystyki językowej, ale nie tylko polskiej, bo także zagraniczni badacze

zainteresowali się jego wzorem w taki sposób, że modyfikowali i usprawniali sposoby wyliczania współczynnika k . Kuraszkiewicz tym samym, jako jeden z nielicznych polskich badaczy, zostawił ślad w światowych badaniach językowych, które oparte są na statystyce.

W tym samym roku (1958) Stefan Hrabec, jeden ze współautorów *Słownika polszczyzny XVI wieku*, opublikował artykuł, w którym opisał, jakie są postępy w pracy nad słownikiem. Zaznaczył przy tym, że podczas początkowych badań zespół korzystał z metody zaproponowanej przez Kuraszkiewicza, czyli analizowali materiał tylko z pewnej części tekstów, jednak w dalszej części artykułu zaznacza, że „późniejsze doświadczenia, jak też chęć pokazania indywidualności twórczych, kazały nam (zespołowi – przyp. R.M.) się z tego sposobu opracowywania materiałów wycofać” [Hrabec 1958: 237]. Prawdopodobnie metoda Kuraszkiewicza była zbyt nowoczesna, jak na owe czasy i dlatego nie spotkała się ze zrozumieniem. Współcześnie badanie próbek korpusów, ze względu na ich ogrom, jest już koniecznością.

W 1959 roku, czyli rok po publikacji artykułu, w którym Kuraszkiewicz zaproponował wyliczanie współczynnika k , metoda ta doczekała się krytyki ze strony Kazimierza Budzyka [1959]. Zauważył on, że zmieniając metodę ekscerpcji materiału, np. z całościowego badania tekstu na korzystanie tylko z co piątej strony (20% całości), wskaźnik ulega zmianie, a w konsekwencji jest mniej korzystny dla tekstów, które zbadane zostały w całości. Budzyk dostrzegł także, że autor nie ustalił do tej pory wysokości współczynnika błędu, który jest niemalże obowiązkowy przy jakichkolwiek badaniach statystycznych. Badacz uważa, że wykorzystanie wyników badań analizy ilościowej pozyskanych metodą Kuraszkiewicza może być jedynie użyteczne w celu sprawdzenia, jaka była rola pisarza w tworzeniu „języka literackiego” lub w jakim stopniu twórca go opanował. Nie można natomiast na podstawie wyników badań statystycznych słownictwa ocenić walorów artystycznych utworu, bo nie wynikają one tylko z bogactwa czy też zróżnicowania używanej leksyki.

Mimo że Kazimierz Budzyk negatywnie odniósł się do zastosowania metod statystycznych w badaniach dzieł literackich, to nie porzeczono ich podejmować. W 1962 roku Maria Kniaginiowa przy wykorzystaniu metod ilościowych zbadała przydawkę w twórczości Stefana Żeromskiego. Zebrawszy teksty autora, podzieliła je na cztery kategorie, a mianowicie na dramaty, powieści, prozy poetyckie i publicystykę. Z każdej z nich, przy pomocy tablic liczb losowych Vielrosego²⁴, wylosowała 200 stron, a z każdej strony wybrała po jednym wypowiedzeniu. Metoda takiego losowania nazwana została losowaniem dwuwarstwowym,

²⁴ Tablice, na których Egon Vielrose wynotował losowe liczby. Ideą tablic było to, żeby w czasach, w których nie było komputerów, wytworzyć możliwie obiektywne narzędzie do obiektywnej pracy z materiałem.

gdzie pierwsza warstwa to strony, druga to zdania. Celem autorki artykułu było potwierdzenie hipotezy, że poszczególne gatunki, nawet w twórczości tego samego pisarza, różnią się istotnie co do wykorzystania przydawek. Autorka w toku analizy potwierdziła swoje przypuszczenie, a największą różnicę w użyciu przydawek wykazała między publicystyką i dramatem.

W 1965 roku ukazał się tom o przewrotnym tytule *Poetyka i matematyka*. Jego redaktorką była M.R. Mayenowa, której artykuł otwiera tę publikację. Badaczka we wstępie pracy podkreśliła, że w ówczesnych czasach panuje ogólne przekonanie, oczywiście z nielicznymi wyjątkami, że „statystyka jest znakomitą drogą do fałszywych wniosków” [Mayenowa 1965: 10]. Mayenowa nie zgadza się z tym stwierdzeniem i staje w obronie statystyki. Według badaczki trzeba jej używać z rozwagą, a nie opierać się wyłącznie na danych liczbowych i ślepo im wierzyć. Ważny w tego typu badaniach, poza danymi ilościowymi, jest także czynnik ludzki. Statystyka jest, zdaniem Mayenowej, idealnym uzupełnieniem i dowodem potwierdzającym niektóre statystyczne intuicje badawcze. Przydaje się ona w badaniu stylu, który według niej jest wynikiem wyborów środków synonimicznych. Ważnym mechanizmem statystyki językowej jest także opracowywanie tekstów w taki sposób, aby ich wyniki można było ze sobą zestawiać i porównywać. Tekst Mayenowej jest swoistym manifestem na temat przydatności badań statystycznych w językoznawstwie.

Rok później (1966) powstała pierwsza duża analiza całościowa ważnego dla Polski dzieła – „Pan Tadeusz”, w której to analizie zastosowano wyłącznie badania statystyczne. Autorką monografii była Jadwiga Sambor. Badaczka dodatkowo zestawiała wyniki badań uzyskane z epepei z opracowaną przez R. Morgenthalera grecką leksyką Nowego Testamentu [1958]. Praca Sambor to obszerna, licząca 160 stron monografia, w której badaczka przyjrzała się tekstowi epepei w celu statystycznego przeanalizowania „wyrazów”²⁵ o różnych częstotliwościach występowania (zarówno rzadkich, jak i częstych). Badaczka wyliczyła częstość ich użycia, a wyniki przedstawiła w obszernych tabelach. Efekty pracy pokazały, że np. w *Panu Tadeuszu* hapax-, dis- i trislegomeny, czyli wyrazy, których częstotliwość w tekście wynosi 1, 2 lub 3 użycia, zajmują do 80% całego zbioru słownictwa, jest to wynik znacznie wyższy niż w np. analogicznych badaniach Nowego Testamentu. We wnioskach Sambor napisała, że:

PT (Pan Tadeusz – przyp. R.M.) to tekst poetycki o luźnej strukturze leksykalnej, ogromnej przewadze wyrazów rzadkich i dominujących małych częstościach, tekst NT (Nowego Testamentu – przyp. R.M.) charakteryzuje się słownictwem ubogim

²⁵ Autorka analizowała sumy wyrazów graficznych w tekście.

o znakomitej przewadze wyrazów o wysokich częstościach, przy czym między wyrazami zachodzą dość silne zależności

[Sambor 1969: 118]

Przełomowy, ponieważ pokonano kolejną granicę w badaniach statystycznych, był rok 1971. To właśnie wtedy opublikowana została pierwsza praca statystyczna, która nie bazowała na materiale pisanym, a na mówionym. Autorką pracy była Maria Zrębina, która postanowiła sprawdzić, jakie są najpopularniejsze wyrazy polszczyzny mówionej. Monografia ta była odpowiedzią na zapotrzebowanie ówczesnej nauki, ponieważ język mówiony był słabo opracowany, nie tylko statystycznie, ale pod każdym innym względem. Zarębina próbą badawczą objęła 10.000 „wyrazów”²⁶, czyli, jak pisze autorka, tyle samo, ile wynosiła próba tekstu potocznego pisanego w nieopublikowanych badaniach M. Kniaginiowej i W. Pisarka, które dostępne były wyłącznie pracownikom Ośrodka Badań Prasoznawczych w Krakowie, w tym Zarębinie. Badaczka często w swojej pracy odwoływała się do tych badań i zestawiała je z własnymi.

Konkludując, badaczka stwierdziła, że słownik polszczyzny mówionej jest znacznie uboższy od jej wariantu pisanego. W polszczyźnie mówionej dominują czasowniki, natomiast w pisanej rzeczowniki. Najczęstszym wyrazem w polszczyźnie mówionej jest „wyraz”²⁷ *być* (ze wszystkimi swoimi formami, np. *jest, będzie, byłbym* itd.), na drugim miejscu uplasował się *nie*, a na trzecim *ten (ta, to)*.

Rok 1972 przynosi pierwszą w Polsce, syntetyzującą dotychczasowe badania statystyczne, pracę autorstwa Jadwigi Sambor – *Słowa i liczby. Zagadnienia językoznawstwa statystycznego* [1972]. Bodźcem do powstania pracy była seria wykładów ze statystyki językoznawczej, które autorka prowadziła dla studentów filologii polskiej na Uniwersytecie Warszawskim. W pierwszej części pracy przedstawione zostały zasady budowania list frekwencyjnych, które niezaprzeczalnie są bardzo ważnym elementem badań statystycznych, bo często to od ich stworzenia zaczyna się analizę, a dopiero potem przechodzi się do obliczeń. W pracy przedstawiono także i krótko scharakteryzowano wydane do tej pory słowniki frekwencyjne, ale tylko zagraniczne, ponieważ w Polsce do tej pory nie było jeszcze takiej pracy. W drugiej części natomiast pokazano, jak zastosować metody statystyczne głównie w językoznawstwie oraz np. przy ustalaniu chronologii tekstów czy też ich autorstwa. Ostatnim, ale równie ważnym elementem tej pracy jest dodatek *ABC o metodach statystycznych*, gdzie w przystępny sposób i możliwie prostym językiem omówiono

²⁶ Dzisiaj należałoby użyć terminu *wyraz graficzny*.

²⁷ Według współczesnej terminologii byłby to raczej termin *leksem*.

najważniejsze metody używane w badaniach statystycznych. Ta część książki przeznaczona jest dla osób, które – jak pisze autorka – nie znają zupełnie wyższej matematyki.

W tym samym roku (1972) zbadana została frekwencja wyrazów w prasie. Analizę tę przeprowadził Walery Pisarek, który już we wstępnych częściach pracy zadeklarował, że jest zwolennikiem metod ilościowych w badaniach stylistycznych (językoznawczych). Pisarek już na początku podkreśla, co moim zdaniem jest warte odnotowania, ponieważ także współcześni badacze popełniają ten błąd, że nie można badań ilościowych rozumieć jako opozycji do badań jakościowych.

Jeżelibyśmy koniecznie chcieli przymiotnikowi „ilościowy” ‘oparty na liczeniu’ przeciwstawić przymiotnik o odwrotnym znaczeniu, mógłby to być przymiotnik subiektywny, czyli ‘oparty na poczuciu podmiotu poznającego’. A więc ilościowy = oparty na liczeniu = obiektywny wersus subiektywny = oparty na jednostkowym poczuciu.

[Pisarek 1972: 13]

Reasumując, o badaniach jakościowych możemy mówić zarówno w odniesieniu do badań subiektywnych, jak i ilościowych, jeśli tylko wyniki zostały dobrze i sumiennie opracowane.

Pisarek w swojej pracy zdecydował się zbadać trzy gatunki dziennikarskie: reportaż, komentarz i informację prasową (nazywaną przez autora w niektórych partiach tekstu wiadomościami). W pierwszym etapie badań pogrupował teksty zebrane z gazet i czasopism, uwzględniając ich gatunki. Następnie w celu dokładniejszej weryfikacji poprosił dwóch zawodowych dziennikarzy z różnych miast, żeby na podstawie swojego doświadczenia także dokonali klasyfikacji gatunkowej zebranych tekstów. Do dalszych etapów pracy Pisarek wybrał tylko te, w których obaj dziennikarze byli zgodni. Ponumerował strony i zdania wszystkich tekstów i przy pomocy tablic losowych Velrosego, wybrał określoną liczbę zdań z każdego z gatunków. Łącznie fragmenty poddane analizie obejmowały 127.035 użyć wyrazów. Analiza Pisarka miała na celu znalezienie „słów kluczy” w poszczególnych tekstach, czyli wyrazów, które są typowe dla danego gatunku. Autor chciał także sprawdzić, która z części mowy jest najliczniej reprezentowana w danym gatunku. Wyniki badań pokazały, że w wiadomościach dominują rzeczowniki (45% całego badanego materiału), następne są przymiotniki (13,6%) oraz czasowniki (13%). W komentarzach również dominują rzeczowniki (38,8%), na drugim miejscu są przymiotniki (12,1%), a na trzecim przymiotniki i czasowniki (po 11,6%). W reportażach, tak jak i w innych gatunkach, dominują rzeczowniki (34,4%), na drugim miejscu są czasowniki (13,6%), a na trzecim przymiotniki (11,7%). Autor zamieścił także listę

wyrazów, które występują najczęściej w danych tekstach, po czym na jej podstawie scharakteryzował gatunki. Publikację uzupełnił listami frekwencyjnymi.

Także w 1972 ukazał się pierwszy polski tom frekwencyjny zatytułowany *Słownictwo współczesnej publicystyki polskiej*. Redaktorami pracy byli A. Lewicki, W. Masłowski, J. Sambor i J. Woronczak. W niemalże identycznym składzie osobowym wydany został dwa lata później, czyli w 1974 r., tom *Słownictwo współczesnego języka polskiego*. Zapoczątkował on serię wydawniczą, w której publikowane były, jak piszą autorzy, teksty należące do pięciu głównych stylów polszczyzny pisanej: teksty popularnonaukowe, dramat artystyczny, proza artystyczna, publicystyka oraz wiadomości prasowe. Autorzy zastrzegli we wstępie, że tom wydany w 1972 roku należy traktować jako próbny, a jego treść w nieco zmienionej formie wydana zostanie jako trzeci tom nowej serii.

Pierwszy tom cyklu był analizą tekstów popularnonaukowych. Grupa badawcza postanowiła, że analiza będzie obejmować teksty z lat 1936–1967 oraz że wybrana zostanie z nich próba o łącznej liczbie wyrazów równej 100.000. Materiały tekstowe pozyskano z książek popularnonaukowych, które następnie posegregowano na 4 grupy. Badacze mieli problem z klasyfikacją różnych nauk, bo nie wiedzieli, jaką z nich przyjąć. Ostatecznie podział ustalili, wzorując się na strukturze wydziałów PAN, czyli wyodrębnili nauki: społeczne, biologiczne, matematyczno-fizyczne i techniczne. Z każdej dziedziny wylosowano, wykorzystując tabelę liczb losowych żelaznych, po 2.000 próbek, z których każda była ciągłym tekstem liczącym 50 wyrazów. Materiał opracowano i opublikowano w swoistych gniazdach hasłowych, które wyglądały następująco:

Grafika 3. Przykładowe gniazdo hasłowe

	analiza	1	
6	analiza		
4	analizę		
2	analizie	131	
1	analizie	161	
9	analizy	121	
22			

Źródło: Lewicki, Kurcz, Sambor, Woronczak 1974: 36

W pierwszym wierszu podane jest hasło, opatrzone kwalifikatorem części mowy (1 – rzeczownik). W następnych wierszach umieszczone zostały słowoformy hasła (*analiza, analizę, analizie...*), a po ich lewej stronie liczba wystąpień w tekstach. Trzycyfrowe liczby po prawej stronie to opisy gramatyczne, występują one jeśli są potrzebne, czyli jeśli jakaś

słowoforma ma np. w innym przypadku identyczną formę graficzną (131 – rzeczownik w celowniku l. poj., 161 – rzeczownik w miejscowniku l. poj., 121 – rzeczownik w dopełniaczu l. poj.). Na dole zaś podana jest suma wszystkich słowoform (22). Taki sposób przedstawiania zanalizowanego materiału powtórzony został w następnych tomach serii.

W 1975 roku ukazała się kolejna praca Jadwigi Sambor (*O słownictwie statystycznie rzadkim*), tym razem dotycząca wyłącznie słownictwa statystycznie rzadkiego, ponieważ, jak zaznacza autorka, w tamtym okresie dominowały badania skupiające się wyłącznie na słownictwie, którego frekwencja była wysoka. Pomijano natomiast wyrazy rzadkie, co było błędem, ponieważ te rzadkie wyrazy, z niską frekwencją w tekście dominują z kolei w słownikach. Znaczną część pracy autorka poświęciła derywatom, np. tym, które tworzone są przez dodanie przedrostków *bez-*, *nie-* czy *anty-*. Końcowe fragmenty poświęciła na analizę złożeń.

Badaniami ilościowymi zainteresowała się także Danuta Buttler, która w 1977 roku opublikowała artykuł *Dobór wyrazów do słownika – minimum języka polskiego*. Buttler zaproponowała w tej pracy, żeby podczas tworzenia słownika minimum polszczyzny oprzeć się na wynikach badań frekwencyjnych, czyli na listach. Niestety obecne w tamtych czasach listy frekwencyjne skupiały się tylko na analizie wybranych stylów²⁸, brakowało zbiorczego podsumowania. Brak ten nie zraził autorki, która kontynuowała badania i doszła do wniosku, że tematycznie najbardziej zbliżona do polszczyzny kolokwialnej, którą uwzględniać powinien słownik minimum, jest lista frekwencyjna stworzona na podstawie analizy drobnych wiadomości prasowych. Dlatego to właśnie tę listę powinno wybrać się jako bazową do stworzenia słownika. W dalszych partiach artykułu Buttler podkreśla, że badacz opracowujący taki słownik, który byłby ważną pomocą dla osób tworzących podręczniki dla obcokrajowców, nie może bazować wyłącznie na liście frekwencyjnej, ale musi ją uzupełnić dodatkowymi informacjami. Słownik taki musi być kompromisem między, tym co zdaniem słownikarza jest ważne w polszczyźnie, aby obcokrajowiec mógł poradzić sobie w podstawowych sprawach życia codziennego a tym, co jest w listach frekwencyjnych.

Rok później (w 1978) Bronisław Ročławski w jednym z artykułów przedstawił, jak można wykorzystać komputery do badań językoznawczych. Jest to jeden z pierwszych postulatów, aby do badań nad językiem wykorzystać komputery, które w tamtych czasach często budziły niechęć wśród wielu językoznawców. Ročławski pisze, że głównymi powodami małej liczby prac językoznawczych opracowanych przy wykorzystaniu komputerów są: brak

²⁸ Do tamtej pory wydano tylko 3 tomy serii *Słownictwo współczesnego języka polskiego: teksty popularnonaukowe* (tom 1), *drobne wiadomości prasowe* (tom 2) i *publicystyka* (tom 3).

wiedzy wśród badaczy na temat tego, jak można wykorzystać to narzędzie w badaniach, trudności z dostępem do maszyn oraz brak zrozumienia potrzeb humanistów w środowisku informatyków.

Autor zachęca językoznawców, aby wykorzystywali w swojej pracy komputery. Apeluje także, by przy wykorzystaniu nowoczesnych maszyn i w oparciu o dane ilościowe opracować bardziej aktualną i odpowiadającą rzeczywistości normę językową, ponieważ w tamtych czasach była ona ustalana przez „niewielkie grono badaczy, którzy własne poczucie językowe przypisywali pozostałym (34 milionom) mieszkańcom naszego kraju” [Rocławski 1978: 59]. W dalszej części artykułu autor demonstruje, jak można opracować tekst gwarowy z wykorzystaniem komputera oraz, jak przy pomocy jego autorskiego programu podzielić go na sylaby, by jeszcze efektywniej nad nim pracować. Na końcu tekstu proponuje, żeby zorganizować centralne obserwatorium językowe, które przy pomocy komputerów opracowywałoby tekst polszczyzny mówionej oraz nadzorowało pracę nad tą odmianą języka polskiego w innych ośrodkach, po to, aby prace badawcze nie dublowały się.

Trzeba odnotować, że w tomie, w którym opublikowany został powyżej omówiony artykuł, wydana została także praca Marii Zarębiny *Ustalenie korpusu tekstowego dla badań statystycznych nad słownictwem polszczyzny mówionej* [1978].

Zarówno w pracy Zarębiny, jak i Rocławskiego zaobserwować można, że w tych latach upowszechnia się (wtedy nowy, a obecnie obowiązujący) termin „*korpus tekstów*”. Do połowy lat siedemdziesiątych w literaturze można było spotkać tylko termin „*próby badawcze*”.

M. Zarębina we wspomnianym artykule zastanawia się, jakie formy języka mówionego trzeba by było włączyć do badań, aby były one rzetelne i nie dawały zakłamanych wyników. Stwierdza, że z całą pewnością do tego typu badań wybrać można dialogi, które są najbardziej naturalnymi i często spontanicznymi formami wypowiedzi, włączyć można także polilog oraz w bardzo ograniczonym zakresie monolog, który często jest wtórną realizacją polszczyzny mówionej (np. podczas wykładu, gdy profesor odczytuje jego treść z wcześniej przygotowanych tekstów). Badaczka uważa, że reprezentatywną próbą badawczą byłaby taka, w której udział wzięłoby od 250 do 500 osób. Przy kompletowaniu grupy trzeba by było pominąć dzieci, bo nie mają jeszcze dobrze opanowanej polszczyzny oraz osoby mówiące gwarą, bo nie posługują się wariantem ogólnym języka polskiego. Autorka zaangażowała się w badanie polszczyzny mówionej nie tylko od strony teoretycznej, ale także praktycznej, ponieważ sama zaczęła badać wypowiedzi Polaków, o czym wspomniała w kolejnej publikacji z 1982 roku – *Badania statystyczne nad słownictwem polszczyzny mówionej*.

Artykuł poświęcony został głównie przedstawieniu metodologii pracy z zebraniem materiałem tekstowym. Badaczka wspomina, że po zebraniu materiału reprezentującego polszczyznę mówioną trudno było jej zastosować jakiegokolwiek losowanie tekstów, a tak powinno się czynić przy tego typu badaniach. Ona jednak nie dysponowała tak dużą nadwyżką zebranego materiału, który pozwalałby na losowanie. Postanowiła już podczas gromadzenia tekstów tak dobierać reprezentację, aby z dużym prawdopodobieństwem mogły one posłużyć jej do ostatecznej analizy. Zgromadzony materiał reprezentuje 10 zróżnicowanych społecznie środowisk. Z każdego takiego środowiska zebrano nagrania, które po przetranskrybowaniu na tekst pisany liczyły łącznie 10.000 wyrazów. Czyli, jak łatwo obliczyć, cała analiza opierała się na 100.000 wyrazów.

Zarębina mimo wcześniejszych deklaracji (patrz strona poprzednia), postanowiła ostatecznie włączyć do badań wypowiedzi dzieci w wieku od 7 do 14 lat. Dzieci te stanowiły pierwszą z dziesięciu grup badawczych. Poza tekstami dzieci do badań włączone zostały rozmowy z licealistami, studentami, dziennikarzami, a także dialogi potoczne, wypowiedzi telewizyjne, narady, dyskusje naukowe, wypowiedzi biurowe i rozmowy pozalekcyjne uczniów. Wszystkie one musiały spełniać jeden warunek – być spontaniczne. Gdy nagrania budziły jakieś wątpliwości, co do ich spontaniczności, np. w przypadku nagrań radiowych, gdzie dziennikarz ma możliwość odczytywania tekstu z kartki, to taką próbkę badano, zwracając szczególną uwagę na intonację i porównywano ją z próbką spontaniczną. Materiał do analizy został zebrany w latach 1976–1981. Do jego opracowania użyty został komputer Cyber 72, który niestety obsługiwał tylko wersaliki, wyrazy maksymalnie dwudziestoliterowe i nie posiadał polskich znaków. Problem braku polskich znaków był dość uciążliwy, zwłaszcza że były to badania językowe. Badaczka rozwiązała ten problem poprzez podwojenie niektórych liter, np. wyraz *pokaż* zapisywany był jako POKAZZ, a *póleczka* to POLLECZKA. Zarębina już na tym etapie pracy potwierdziła, że mowa jest znacznie uboższa od tekstu pisanego i nawet najbliższa jej realizacja pisemna, bo ją naśladująca, czyli dramat, ma o około 20% bogatsze słownictwo niż jakakolwiek spontaniczna wypowiedź w języku polskim. Artykuł był kolejnym krokiem zbliżającym badaczkę do całkowitego opracowania polszczyzny mówionej, którą już także zaczęli interesować się inni badacze. Na przykład Halina Zgólkowa w 1983 roku opublikowała książkę *Słownictwo współczesnej polszczyzny mówionej. Lista frekwencyjna i rangowa*.

Praca w dużym stopniu, zwłaszcza pod względem opracowania, nawiązywała do prac z serii *Słownictwo współczesnego języka polskiego* (patrz strona 49–50), czyli także obejmowała próbę wielkości 100.000 wyrazów. „Wyraz” w tej pracy jest rozumiany jako ciąg

znaków od przerwy do przerwy (*wyraz graficzny*). Co ciekawe H. Zgółkowa podczas pracy nad słownikiem nie korzystała z pomocy komputera, wybrała metodę tradycyjną, czyli „obliczenia wykonano przy użyciu ołówka i papieru” [Zgółkowa 1983: 3]. Nagrania wykorzystane do badań, podobnie jak w pracach Zarębiny, musiały być spontaniczne, a te, które budziły wątpliwości były dokładnie weryfikowane, czyli sprawdzano np. czy występuje potok składniowy, czy są anakoluty i przejęzyczenia, czy wypowiedzi charakteryzują się dużą eliptycznością lub czy występuje powtarzalność wyrazów. Zebrany do badań materiał pochodził z lat 1974–1976, a w jego skład weszły wypowiedzi 150 osób, których spontaniczne dialogi nagrane zostały na taśmy magnetofonowe bez ich wiedzy. Dopiero po zakończeniu nagrywania zapytano te osoby, czy zgadzają się na wykorzystanie wypowiedzi do badań. Wśród badanych byli przedstawiciele różnych grup społecznych czy zawodowych, np.: nauczyciele, urzędnicy, studenci czy księża. Po opracowaniu materiału wyniki przedstawione zostały w alfabetycznej liście frekwencyjnej.

Dwa lata później (w 1985 r.) ukazała się duża, bo licząca 239 stron praca M. Zarębiny *Próba statystycznej analizy słownictwa polszczyzny mówionej (synteza danych liczbowych)*. Jest to finałowa i całościowa prezentacja wyników badań, o których fragmentarycznie autorka wspominała już podczas publikacji kolejnych artykułów. Książka w dużej mierze składa się z tabel, w których autorka podaje różne dane statystyczne, np. wykaz najczęstszych wyrazów w polszczyźnie mówionej. Na pierwszym miejscu w ogólnej klasyfikacji jest *być* (3.690 wystąpień²⁹ na 100.000 wyrazów), kolejną pozycję zajmuje *nie* (2.894), a na trzecim miejscu jest *to* (2.334). Autorka tym samym potwierdziła na większej próbie badawczej to, o czym pisała w 1971 roku (patrz strona 47). Dużą część pracy zajmuje także charakterystyka pod względem językowym 10 grup środowiskowych i społecznych, z których pozyskany został materiał badawczy. Książka ta jest zwieńczeniem 14 lat pracy nad analizą ilościową polszczyzny mówionej.

W 1990 roku nastąpił finał prac zespołu opracowującego serię tomów frekwencyjnych *Słownictwo współczesnego języka polskiego*. Pierwszy tom cyklu wydany został w 1972 roku, jednak pierwsze prace nad nim podjęto już w roku 1967. Po 23 latach pracy przy badaniu frekwencji wyrazów w polszczyźnie ukazał się ostatni, zbiorczy tom serii, czyli *Słownik frekwencyjny polszczyzny współczesnej*. Autorzy wykorzystali do jego przygotowania zebrane we wszystkich pięciu tomach materiały, jednak nie ograniczyli się wyłącznie do nich, ponieważ do badań włączyli także wyniki pracy Haliny Zgółkowej, które pierwotnie przedstawione

²⁹ Rozumianych jako suma wyrazów graficznych, a więc suma tekstowych wystąpień wszystkich słowoform.

zostały w książce z 1983 roku. Podstawową wadą słownika było to, że opierał się on materiale zebrany w latach 1963–1967, czyli zgromadzonym przed wydaniem pierwszego tomu serii *Słownictwo współczesnego języka polskiego*. Słownik w momencie publikacji nie był już aktualny. Autorzy zdawali sobie z tego sprawę i już we wstępie zaznaczyli, że nie poleciliby tego słownika np. do tworzenia na jego podstawie jakiegokolwiek pracy dla obcokrajowców, która w jakiś sposób opisywałby życie polityczne w Polsce. Trzeba traktować ten słownik jako pracę, która jest nie tyle analizą współczesnej polszczyzny, co publikacją historyczną. W słowniku tym podano frekwencję wyrazów, zachowując stosowany w poprzednich tomach podział polszczyzny na pięć stylów funkcjonalnych (patrz strona 49). Całość uzupełniono informacjami na temat liczby wystąpień w listach frekwencyjnych polszczyzny mówionej. Podano także dane zbiorcze oraz dane statystyczne, takie jak np. częstość absolutną wyrazów³⁰, częstość względną³¹ i wskaźnik równomierności rozkładu w stylach³². Podsumowując, słownik powstał na podstawie korpusu tekstów, który zbudowany został z 600.000 „wyrazów”³³.

Zofia Kurzowa z kolei zastosowała badania ilościowe w analizie „języka telewizji”. Jednak nie badała materiału sama, tak jak np. Halina Zgółkowska, ale zebrała sześciuosobowy zespół, który podczas prac przygotowawczych do badań zgromadził łącznie prawie 900 godzin nagrań różnych audycji telewizyjnych z I programu telewizji polskiej, które wyemitowane zostały od 1 kwietnia 1978 roku do 31 marca 1979 roku. Owocem ich pracy była opublikowana w 1991 roku monografia *Badania nad językiem telewizji polskiej. Ilościowy opis słownictwa*. Badacze wyróżnili trzy odmiany charakterystyczne dla audycji telewizyjnych: czytaną, monologową i dialogową. Po przetranskrybowaniu tekstu wybrali do analizy łącznie 189.000 słowoform, czyli dla każdej z odmian wylosowano po 900 próbek (po jednej próbce pobierano z każdej godziny materiału). Jedna próbka liczyła siedemdziesiąt wyrazów. Analiza materiału, która uwzględniała najczęściej występujące wyrazy potwierdziła, że odmiany czytana i monologowa najbliższe są stylowi drobnych wiadomości prasowych, natomiast odmiana dialogowa plasuje się między publicystyką a prozą artystyczną. Analiza tematyki słownictwa pokazała natomiast, że w monologu dominuje temat sportu, w dialogu dydaktyka i współdziałania społeczne, a w odmianie czytanej dominują zagadnienia polityczne.

³⁰ „Częstość danego leksemu w całej badanej próbce 500 000 słów” [Lewicki, Kurcz, Sambor i in. 1990: L].

³¹ Powstaje ona przez przemnożenie częstości absolutnej przez wskaźnik równomierności rozkładu. Dzięki temu zabiegowi według badaczy relatywizowane są dane o częstości.

³² „Charakteryzuje rozproszenie jednostki w poszczególnych podzbiorach” [Lewicki, Kurcz, Sambor i in. 1990: L]. Sposób na wyliczenia rozproszenia także wykorzystuję w swojej dysertacji. Dokładnie został on (współczynnik rozproszenia / zmienności) opisany na stronie 212.

³³ Według obecnej terminologii – *wyrazów graficznych*.

W tomie, poza wynikami zbiorczymi, opublikowano także bardziej szczegółowe analizy zebranego materiału. Mirosław Skarżyński zbadał materiał pod kątem występujących w nim form homonimicznych i stwierdził, że dominuje homonimia wewnątrzkatégorialna. Zofia Kubiszyn-Mędrala [1991] ustaliła listę 100 najczęstszych wyrazów w „języku telewizji” i okazało się, że najpopularniejszy wyraz to w – łączna frekwencja 5975, następny w kolejności był czasownik *być* – 5043 wystąpień oraz i – 3746 wystąpień. Natomiast Jolanta Antas [1991] pracę swoją poświęciła analizie części mowy. Po zestawieniu wyników okazało się, że w tekście i słowniku dominują rzeczowniki. Na drugim miejscu (uwzględniając tylko wystąpienia w tekście) znalazły się zaimki, zaś na trzecim przyimki. W słowniku na drugim miejscu uplasowały się przymiotniki, a na trzecim bezokoliczniki.

Dwa lata później (w 1993 roku) ukazał się *Schematyczny indeks a tergo polskich form wyrazowych* Jana Tokarskiego³⁴. W historii powstawania słownika autor ujawnił, że prace nad jego przygotowaniem rozpoczęły się jeszcze przed rokiem 1973, w którym to wydany został *Indeks a tergo do Słownika języka polskiego* pod redakcją Witolda Doroszewskiego. Tokarskiemu więc trzeba przypisać pierwszeństwo pomysłu opracowania takiej pozycji. Jej kontynuator – Z. Saloni wraz z zespołem – pracował nad słownikiem przez ponad 10 lat, a konkretnie od 1980 roku, kiedy to przejął nieopublikowaną autorską wersję słownika. Nowi redaktorzy opracowali wersję komputerową publikacji. Co ciekawe, za podstawę do pracy wybrano *Słownik języka polskiego* Witolda Doroszewskiego, który, jak wiadomo, sam później stworzył własny słownik/indeks tego typu. Praca Tokarskiego i Saloniego zasługuje na odnotowanie, ponieważ zastosowano w niej analizatory morfologiczne³⁵, ważne narzędzie w pracy z korpusami tekstów.

Przełomowy, głównie dla środowiska łódzkiego, był rok 1995, kiedy to nawiązana została współpraca między pracownikami Katedry Języka Angielskiego Uniwersytetu Łódzkiego a Anthonym McEnery – przedstawicielem Uniwersytetu w Lancaster. Zaowocowała ona powstaniem w 1997 roku projektu PELCRA (*Polish and English Language Corpora for Research and Application*), którego głównym celem było zbieranie zarówno polskich, jak i angielskich danych językowych w formie elektronicznej, a także gromadzenie tłumaczeń z polskiego na angielski i angielskiego na polski. Ważnym elementem projektu było także tworzenie nowych narzędzi do pracy z zebraniem korpusem tekstów. Jedną z kluczowych

³⁴ Ukazał się on po śmierci autora, kiedy to pracę nad słownikiem przejął Zygmunt Saloni.

³⁵ Programy analizujące wyrazy graficznie (ich końcową część), a następnie przyporządkowujące je do określonej kategorii morfologicznej.

osób, które uczestniczyły w tym przedsięwzięciu była Barbara Lewandowska-Tomaszczyk, autorka wielu publikacji dotyczących badań korpusowych.

W 2002 roku Marek Świdziński, Magdalena Derwojoweda i Michał Rudolf opublikowali artykuł (*Dehomonizacja i desyntytyzacja w procesie automatycznego przetwarzania wielkich korpusów tekstów polskich*), w którym opisali jedną z trudności, która bardzo często występuje w pracy z dowolnym korpusem tekstów, a mianowicie homonimię. Po wstępnych pracach na materiale ze *Słownika frekwencyjnego współczesnej polszczyzny* badacze zauważyli, że cały zbiór słownictwa, który przedstawiony został w tej pracy zawiera prawie 58% różnego rodzaju form homonimicznych. Autorzy artykułu zaczęli więc pracować nad rozwiązaniem tego problemu. Zamierzali stworzyć program, który potrafiłby dobrze rozpoznawać i interpretować homonimy występujące w polszczyźnie. We wspomnianym artykule nie zaproponowali oni rozwiązania tego problemu, ponieważ ich prace nie dobiegły jeszcze końca. Ograniczyli się tylko do przedstawienia i opisanie nowych terminów, których używają w pracy. Są to: *dehomonizacja* („zinterpretowanie homoforny polegające na przypisaniu jej odpowiednich leksemów”), *desyntytyzacja* („zinterpretowanie homoforny polegające na przypisaniu jej odpowiednich charakterystyk gramatycznych”) oraz *dehomofornizacja* („na oznaczenie każdego rozpoznania analizowanego słowa, tzn. zarówno *dehomonizacji*, jak *desyntytyzacji*” [Świdziński, Derwojoweda, Rudolf 2002: 191]).

Rozwinięcie tematu związanego z metodologią badań korpusowych przedstawił w 2004 roku Michał Rudolf. W monografii [Rudolf 2004] autor pokazał krok po kroku, jak wygląda proces przetwarzania korpusów, jakie można do tego wykorzystać narzędzia oraz jakie problemy są jeszcze do rozwiązania w zakresie językoznawstwa korpusowego, a konkretnie w subdyscyplinie, jaką jest inżynieria korpusowa³⁶ (reprezentantem tej dyscypliny jest autor monografii).

Rok później, czyli w 2005 Barbara Lewandowska-Tomaszczyk publikuje pracę *Podstawy językoznawstwa korpusowego*. „Pierwsze takie przedsięwzięcie na rynku wydawniczym w Polsce” [Lewandowska-Tomaszczyk 2005: 8]. We wstępie autorka zaznacza, że podwaliny do powstania polskiego językoznawstwa korpusowego dały prace Idy Kurcz (patrz strona np. 49), a także opracowania dotyczące różnego rodzaju analizatorów morfologicznych, które zostały wykorzystane np. przy tworzeniu *Słownika a tergo* przez Jana Tokarskiego (patrz strona poprzednia). Autorka podkreśla, że języki fleksyjne stanowią „spore wyzwanie dla automatycznej analizy morfologicznej i składniowej” [Lewandowska-

³⁶ Dziedzina zajmująca się opracowaniem narzędzi do pracy z korpusami tekstowymi.

Tomaszczyk 2005: 17]. Dzieje się tak, ponieważ występuje w nich mnogość form gramatycznych, ale jeszcze większym problemem jest homonimia, zwłaszcza w kręgu kategorii gramatycznych (np. *książki*, to mianownik liczby mnogiej lub dopełniacz liczby pojedynczej). W dalszej części pracy Lewandowska-Tomaszczyk zaznacza, że nie każdy zbiór tekstów może być korpusem. Może nim być tylko taki zbiór, który spełnia pewne kryteria, a są one dosyć płynne. Według badaczy z ośrodka w Lancaster musi mieć on np. skończoną wielkość oraz format elektroniczny, natomiast ośrodek badawczy w Birmingham akcentuje, że wielkość korpusu musi być „znaczna” (trudno jednak ustalić, ile musi zawierać jednostek, bo z roku na rok wartość ta wzrasta) oraz musi być autentyczny. Trudno nie zgodzić się z tymi warunkami, ale jeden z nich budzi pewne wątpliwości, a mianowicie skończoność korpusu. Po przyjęciu tego kryterium, raz stworzonego korpusu nie moglibyśmy rozbudowywać, na czym moim zdaniem ucierpiałaby jego aktualność. Autorka poświęca także fragment pracy różnym rodzajom korpusów. Wyróżnia np. dwa główne typy zbiorów tekstów, a mianowicie takie, które zawierają indeks oraz takie, które są go pozbawione. Indeksowanie polega na tym, że do wyrazów dodawane są specjalne kody, które dostarczają informacji na temat kategorii gramatycznych danych jednostek. Jako przykład podany został fragment angielskiego tekstu, w którym zastosowano indeksowanie: „hospitality_NN is_BEZ an_AT excellent_JJ virtue_NN...”³⁷ [Lewandowska-Tomaszczyk 2005: 33] W kolejnych rozdziałach pracy badaczka skupia się na szczegółowym przedstawieniu informacji o rodzajach korpusów. Pokazuje, czym są korpusy uczniowskie (to zbiór danych językowych, które pozyskane zostały od osób uczących się języka obcego jako drugiego), równoległe (zbiór tekstów wraz z ich przekładami na inny język) czy porównywalne (kolekcja tekstów, które stworzone zostały przez rodzimych użytkowników danego języka w różnych językach obcych). Badaczka przedstawia także, jak działają mechanizmy wyszukiwania haseł.

Rok później (w 2006) powołany został zespół, w którego skład weszły trzy ośrodki akademickie: Instytut Języka Polskiego PAN (reprezentowany przez Rafała Górskiego), Instytut Podstaw Informatyki PAN (Adam Przepiórkowski) i Uniwersytet Łódzki (Barbara Lewandowska-Tomaszczyk), a także wydawnictwo PWN (Miroslaw Bańko i Marek Łaziński). Zespół ten miał stworzyć korpusu tekstów polskich. W wyniku prac powstał „Narodowy Korpus Języka Polskiego, największy, morfologicznie anotowany zbiór danych języka polskiego” [Przepiórkowski, Bańko, Górski, Lewandowska-Tomaszczyk 2012: 9]. Praca nad projektem zakończona została w 2012 roku, a jej podsumowaniem i zwieńczeniem była

³⁷ W przykładowym zdaniu dużymi literami (NN, BEZ, AT, JJ) oznaczone zostały części mowy.

publikacja zawierająca szereg artykułów związanych z powstawaniem korpusu pt. *Narodowy Korpus Języka Polskiego* [2012].

Badania ilościowe przydatne są nie tylko w badaniach zebranego i utworzonego wcześniej korpusu tekstów, ale obecnie są stosowane także w wielu innych dziedzinach nauki, np. w glottodydaktyce polonistycznej przy dostosowywaniu materiałów dla obcokrajowców. O takim zastosowaniu badań ilościowych wspominała już między innymi Danuta Buttler w 1977 roku (patrz strona 50). Przykładem symbiozy między dydaktyką a badaniami ilościowymi może być na przykład wydana w 2011 roku książka Anny Seretny *Kompetencja leksykalna uczących się języka polskiego jako obcego w świetle badań ilościowych*, w której autorka na podstawie różnych list frekwencyjnych proponuje, aby programy nauczania języka polskiego jako obcego były układane z uwzględnieniem częstości użycia danych wyrazów w ogólnej polszczyźnie. To częstość występowania wyrazów powinna być jako pierwsza brana pod uwagę np. przy pisaniu podręczników. A. Seretny zaznacza, że niestety ciągle dobór treści leksykalnych odbywa się intuicyjnie [Seretny 2011].

Obecnie z roku na rok widać coraz większe zainteresowanie badaniami korpusowymi, np. w 2015 roku cały tom „Prac Filologicznych” poświęcony został temu zagadnieniu. Zebrano w nim teksty zarówno teoretyczne, jak i te, które prezentują wyniki różnych analiz. We wstępie redaktorki (Milena Hebal-Jeziarska i Elżbieta Kaczmarska) podkreślają, że ta dziedzina badań staje coraz popularniejsza, jednak moim zdaniem nie staje się, ale już taka jest, świadczy o tym coraz większa liczba publikacji z nią związanych.

4. USTALENIA TERMINOLOGICZNE

Poniższa część pracy ma charakter przeglądu. Zebrano w niej kilka ważnych dla tej pracy terminów. Często są one różnie definiowane, dlatego należało prześledzić rozwój danego pojęcia i wybrać najbardziej przydatne rozumienie. Pisząc tę część autor wzorował się na sposobie przedstawiania terminologii, który w swoich pracach wykorzystali np. Władysław Miodunka [1989] czy Władysław Lubaś [1979]. Na początku segmentu poświęconego danemu terminowi wymieniono szereg jego różnych definicji, następnie wybrano i pogrupowano elementy wspólne, a w ostatniej części podsumowano zebrane informacje i wskazano, jak dany termin będzie rozumiany w tej dysertacji. Autor ograniczył się tylko do nowszych definicji, czyli tych z ostatnich 20 lat (1998–2018)³⁸.

Przegląd terminologiczny otwierają rozważania nad pojęciem *języka*, które jest kluczowe dla prac językoznawczych.

4.1. JĘZYK

Język to termin wieloznaczny, piszą o tym m.in. Daniela Podlaska i Iwona Płóciennik w *Słowniku wiedzy o języku* [2004: 105]. Definiują go jako:

- * kod, tj. system językowy, na który składają się reguły gramatyczne umożliwiające łączenie mniejszych części języka w większe, oraz słownictwo;
- * budowa każdego tekstu, wypowiedzi (język utworu, język pisarza);
- * synonim mówienia.

[Podlaska, Płóciennik 2004: 105].

Dodają też, że „w językoznawstwie używa się tego pojęcia przede wszystkim w znaczeniu pierwszym” [Podlaska, Płóciennik 2004: 105].

Autorki w definicji wymieniają także cechy języka:

Arbitralność – między formą a treścią znaku nie ma związku naturalnego. (...) Arbitralność nie oznacza jednak dowolności – raz ustalony związek między formą a treścią jest związkiem obowiązującym w obrębie danego społeczeństwa;

Semantyczność – znaki językowe odsyłają odbiorcę do jakichś zjawisk otaczającego świata. Treść znaków jest zawsze ogólna – odsyłać może jednocześnie do jednej, konkretnej rzeczy i do całej klasy (...);

³⁸ Wyjątkiem były takie sytuacje, w których autorzy nowszych opracowań nie podawali własnych definicji, tylko przywoływali jedną ze starszych, tę z którą się zgadzali.

Dwuklasowość (dwustopniowość) – język zorganizowany jest na dwóch poziomach: ze składników mniejszych (niemających znaczenia) buduje się struktury większe (znaczące) – wyrazy, zdania;

Autonomiczność – w każdym języku naturalnym można mówić o sprawach, zdarzeniach odległych w czasie i przestrzeni (...);

Kreacyjność – produktywność systemu. Człowiek może mówić o wszystkim, a więc ma zdolność tworzenia i rozumienia coraz to nowych słów i wypowiedzi. Dzięki tej właściwości możliwe jest powstawanie utworów literackich;

Metajęzykowość – w każdym języku naturalnym można mówić o samym języku (...).

[Podlawska, Płóciennik 2004: 105–106]

Wieloznaczność terminu *język* podkreślona została także w definicji Kazimierza Polańskiego, którą opublikowano w *Encyklopedii języka polskiego* [1999] pod redakcją Stanisława Urbańczyka i Mariana Kucały. Autor zdefiniował *język* następująco:

1. system znaków (prymarnie dźwiękowych, wtórnie pisanych i innych) służący do porozumiewania się w obrębie danej społeczności;
2. w semiotyce i teorii informacji każdy system znaków, który służy do porozumiewania się;
3. anatomiczny organ człowieka odgrywający główną rolę w artykulacji mownej.

[Polański 1999: 153–154]

Włodzimierz Gruszczyński i Jerzy Bralczyk w *Słowniku gramatyki języka polskiego* [2002] także nieco inaczej definiują termin, aczkolwiek nadal jest on wieloznaczny:

o języku myśli się najczęściej jako o pewnym, najbardziej rozwiniętym, sposobie porozumiewania się, także jako o środku (narzędziu), służącym do tego porozumiewania się, wreszcie jako o usystematyzowanym zbiorze znaków i reguł ich łączenia, czyli o systemie znaków. W językoznawstwie mówi się o języku na te wszystkie możliwe sposoby, a samo słowo „język” jako termin najczęściej bywa używane w znaczeniu „system znaków”. Język jako system znaków służących komunikowaniu się ludzi, czyli będący narzędziem międzyludzkiej komunikacji.

[Gruszczyński, Bralczyk 2002: 109]

Autorzy powyższej definicji, czynią jednak w swojej pracy pewne zabezpieczenie i dodają, że słowo *język* bywa używane także w węższym, jak i szerszym, niż podane w definicji, znaczeniu [Gruszczyński, Bralczyk 2002: 109].

Z wieloznacznego definiowania *języka* rezygnuje natomiast Andrzej Markowski, który definiuje ten termin następująco:

język, którym się posługujemy, to system umownych znaków – w uproszczeniu można przyjąć, że są nimi wyrazy – i reguł łączenia tych znaków w większe całości.

[Markowski 1999: 1658]

Powyżej zaprezentowałem tylko kilka definicji, które miały na celu pokazanie, że zdefiniowanie terminu, jakim jest *język* nie jest łatwym zadaniem, gdyż jest on wieloznaczny

i posiada nie dwa lub trzy, ale bardzo dużo różnych znaczeń. Zestawiając powyższe definicje, *język* może oznaczać:

1. system znaków i reguł ich łączenia [twierdzą tak Podlawska i Płóciennik, Gruszczyński i Bralczyk oraz Markowski];
2. budowę każdej wypowiedzi lub tekstu [Podlawska i Płóciennik];
3. synonim mówienia [Podlawska i Płóciennik];
4. system znaków służący do porozumiewania się [Polański, Gruszczyński i Bralczyk];
5. organ człowieka [Polański];
6. sposób porozumiewania się [Gruszczyński i Bralczyk];
7. środek do porozumiewania się [Gruszczyński i Bralczyk].

Wszystkie powyższe definicje w jakiś sposób powiązane są z komunikacją. Może najmniej związana jest definicja 5, chociaż z drugiej strony bez języka-organu komunikacja werbalna byłaby znacznie utrudniona. Jednak tę definicję w dalszych rozważaniach pomnę, ponieważ w mojej pracy, gdy będę przytaczał termin *język*, na pewno nie będzie on rozumiany jako organ.

Warto podkreślić, że w rozumieniu potocznym wszystkie definicje *języka* są jak najbardziej poprawne i gdy rozmawiamy z kolegami, to możemy używać *języka* w dowolnym znaczeniu. Natomiast w nauce, a zwłaszcza w językoznawstwie, nie wszystkie znaczenia powinny być stosowane. Moje wątpliwości budzi definicja 2, według której można mówić o „*języku danego pisarza*”, czy też o „*języku utworu*”. Idąc tym tokiem rozumowania, mógłby istnieć „*język sklepikarki*”, „*język Jana Kowalskiego*” czy „*język „Gazety Wyborczej*””. Na stronie 28 przytoczyłem fragment artykułu Karola Dejny, który twierdził, że istnieje tylko jeden *język* – *język polski*. Podzielam jego zdanie. A powyżej przytoczone połączenia wyrazowe z terminem *język*, np. „*język Jana Kowalskiego*”, proponowałbym zamienić np. na *idiolekt Jana Kowalskiego*. Dla innych połączeń wyrazowych są także dostępne lepsze warianty terminologiczne niż *język*.

Nie zgadzam się także z definicją 6 i 7, ponieważ nie każdy sposób oraz środek porozumiewania się można nazwać *językiem*. Jest to tylko jeden ze sposobów i jeden ze środków komunikacji, nie można więc generalizować. Nie mogę zaakceptować także definicji 3, ponieważ skoro mamy termin *mówienie/mowa* i *język*, to nie utożsamiamy ich ze sobą. *Język* jest terminem nadrzędnym wobec *mówienia*. Synonimia terminologiczna nie jest zjawiskiem dobrym w nauce, ponieważ prowadzi do wielu nieporozumień. Uważam także, że sam system

znaków (definicja 4) nie wystarczy, aby istniał *język*. Niezbędne są także reguły nimi zarządzające.

Podsumowując, w tej pracy będę używał terminu *język* w pierwszym znaczeniu, czyli to „system znaków i reguł ich łączenia”. Trzeba przy tym jeszcze raz podkreślić, za Karolem Dejną, że w Polsce mamy tylko jeden *język* – *język polski*.

4.2. SOCJOLEKT

Terminu *socjolekt* używam już w tytule tej pracy, dlatego wymagane jest, abym przedstawił, jak go rozumiem i w jakim znaczeniu będzie on używany w dysertacji.

Najkrótszą znaną mi definicją *socjolektu* jest ta, umieszczona w *Słowniku wiedzy o języku* [Podlaska, Płóciennik 2004]. *Socjolekt* to:

stała socjalna odmiana języka

[Podlaska, Płóciennik 2004: 247]

Nieco bardziej rozbudowana jest definicja zaproponowana przez Aleksandra Wilkońia w *Typologii odmian językowych współczesnej polszczyzny* [2000]. *Socjolekty*

to odmiany językowe związane z grupami społecznymi, takimi jak klasa, warstwa, środowisko i grupy zawodowe, co nie znaczy, aby w konkretnych warunkach duże grupy społeczne: klasa i warstwa, wyodrębniły się w sposób wyrazisty od strony językowej, aby wytworzyły własne socjolekty.

[Wilkoń 2000: 88]

Walery Pisarek w *Encyklopedii języka polskiego* pod redakcją Stanisława Urbańczyka i Mariana Kucalę wyjaśnia, czym jest *socjolekt*, odwołując się do terminu *dialekt*. *Socjolekt* to w odróżnieniu od dialektu (czyli odmiany uwarunkowanej terytorialnie) odmiana języka uwarunkowana społecznie, a więc używana przez ludzi, których łączy więź wspólnego zawodu, uczenia się, losu, zainteresowania, hobby. S[ocjolekt] (uzupełnienie R.M.) bywa też nazywany językiem środowiskowym, gwarą środowiskową, gwarą zawodową, slangiem, żargonem.

[Pisarek 1999: 364]

Najobszerniejszą definicję opublikował Stanisław Grabias we *Współczesnym języku polskim* [2001] pod redakcją Jerzego Bartmińskiego. *Socjolekty* to

odmiany języka narodowego związane z istnieniem trwałych grup społecznych połączonych jakimś rodzajem więzi. Odmiany takie nazywa się rozmaicie: językami grupowymi, językami zawodowymi, gwarami środowiskowymi, slangiem, żargonem itp. W potocznej świadomości języka wymienione wyżej nazwy traktuje się jako synonimy. W lingwistyce znaczenia ich są coraz częściej różnicowane i w związku z tym tworzy się zwyczaj stosowania terminu „socjolekt” w znaczeniu najogólniejszym, obejmującym wszystkie uwarunkowane społecznie odmiany języka.

[Grabias 2001: 234–236]

Natomiast Tomasz Piekot w książce *Język w grupie społecznej. Wprowadzenie do analizy socjolektu* [2008] przywołuje inną definicję S. Grabiasa. *Socjolekty* to „odmiany języka powstałe we wspólnotach ludzkich wyznaczonych pewnym rodzajem więzi społecznej” (Grabias, 1994: 97).

[Piekot 2008: 31]

Autor w dalszej części pracy dodaje, że „szerokie rozumienie tego pojęcia pozwala uniknąć wieloznaczności wywołanej użyciem innych terminów, dlatego też we współczesnych badaniach termin *socjolekt* powinien występować w znaczeniu nadanym mu przez Grabiasa” [Piekot 2008: 31].

Zestawiając ze sobą powyższe definicje, można stwierdzić, że *socjolekt* to:

1. odmiana języka, która wiąże się z istnieniem stałych/trwałych grup społecznych [Podlaska i Płóciennik, Grabias 2001];
2. odmiana języka używana przez ludzi, których łączy jakaś więź [Wilkoń, Pisarek, Grabias 1994].

W pierwszej definicji autorzy podkreślają, że więzi danej grupy społecznej muszą być trwałe, aby można mówić o *socjolekcie*. Na stronie 29, przytoczyłem fragment pracy Ewy Kołodziejek, która opisywała, że czasami konkretne warianty „*gwary marynarzy*” istnieją tylko przez jeden rejs statkiem. Czy więc w takim przypadku można mówić o *socjolekcie*? Załoga tego rejsu stworzyła pewną więź, na tyle mocną, aby wytworzyć jakiś rodzaj, odmiankę polszczyzny, jednak nie była ona zbyt trwała, bo istniała tylko przez jeden rejs. Autorka w 1990 roku nazwała tę odmiankę/te odmianki „*gwarami*”, dzisiaj można by ją/je nazwać *socjolektami*, ponieważ wtedy ten termin nie był jeszcze rozpowszechniony. Moim zdaniem kryterium trwałości nie jest konieczne, aby powstał *socjolekt*, dlatego w tej pracy przyjmuję definicję drugą. *Socjolekt* to: odmiana języka używana przez ludzi, których łączy jakaś więź.

4.3. WYRAZ, LEKSEM I JEDNOSTKA LEKSYKALNA

Dużo miejsca w tej pracy poświęcam na analizę różnych przykładów z *socjolektu polskich graczy wideo*, dlatego zobligowany jestem do przedstawienia, jak w tej pracy rozumiane są pojęcia: *wyraz*, *leksem* czy *jednostka leksykalna*.

Najbardziej lakoniczna definicja wyrazu została umieszczona w *Słowniku wiedzy o języku* [Podlaska, Płóciennik 2004]. Termin *wyraz*, wraz z synonimami podanymi w nawiasach (*słowo*, *leksem*), definiowany jest jako:

najmniejsza jednostka systemu leksykalnego języka

[Podlaska, Płóciennik 2004: 292]

W powyższej definicji nie uwzględniono zróżnicowania znaczeniowego terminu *wyraz*, ograniczono się tylko do krótkiego opisu, natomiast w wielu innych opracowaniach autorzy wprowadzają różne znaczenia tego terminu. Na przykład Włodzimierz Gruszczyński i Jerzy Bralczyk podają dwa znaczenia terminu *wyraz*:

1. podstawowa, względnie samodzielna jednostka tekstu, rozpatrywanego tylko jako forma (bez uwzględniania treści, czyli znaczenia);
2. jednostka zasobu słownikowego danego języka, leksem. Wyraz rozpatrywany zarówno jako jednostka tekstu, jak i języka, odznacza się względną samodzielnością, tzn. ma określoną funkcję składniową, a najczęściej także semantyczną (zaczeniową).

[Gruszczyński, Bralczyk 2002: 310–311]

W osobnym haśle dodatkowo definiują *leksem* jako:

wyraz słownikowy, bywa rozumiany także nieco inaczej – jako część wspólna wszystkich form fleksyjnych, czyli po prostu jako temat fleksyjny, któremu przypisuje się znaczenie słownikowe. W tym sensie mówi się o leksemie KOMET- (o znaczeniu ‘kometa’) (...)

[Gruszczyński, Bralczyk 2002: 311]

Autorzy piszą także, że „w języku potocznym używa się określenia *wyraz* w znaczeniu *wyraz słownikowy*” [Gruszczyński, Bralczyk 2002: 311].

Także Roman Laskowski w definicji przygotowanej do *Encyklopedii języka polskiego* [Urbańczyk, Kucala 1999] podaje dwa znaczenia terminu. Według niego *wyraz*:

jest najmniejszą względnie samodzielną, znaczącą jednostką językową, zdolną do spełnienia jakiejś (jakichś) funkcji syntaktycznej: może on samodzielnie lub w połączeniu z innym wyrazem stanowić człon wypowiedzenia lub wypowiedzenie (...).

Termin „wyraz” używany jest w rzeczywistości w dwu różnych znaczeniach. Oznacza 1. Jednostkę systemu słownikowego języka, 2. Element tekstu. (...) dla uniknięcia tej dwuznaczności wprowadzono rozróżnienie terminologiczne: leksem ‘wyraz jako element systemu słownikowego’ i wyraz tekstowy ‘wyraz jako element tekstu’.

[Laskowski 1999a: 427]

Osobno autor definiuje także *leksem*:

Wyraz jako abstrakcyjna jednostka systemu słownikowego języka. Jest elementem językowym, na który składa się określone znaczenie leksykalne oraz zespół wszystkich funkcji gramatycznych, jakie dany [leksem] (uzupełnienie R.M.) może spełniać, a także zespół form językowych reprezentujących w tekście [leksem] (uzupełnienie R.M.) w jego poszczególnych funkcjach.

[Laskowski 1999b: 205]

W powyższych definicjach autorzy rozpatrywali tylko terminy *leksem* i *wyraz* (lub *wyraz słownikowy*). Dokładniejsze zróżnicowanie znaczeniowe terminu, jakim jest *wyraz*, znaleźć można w rozdziale *Zagadnienia ogólne morfologii*, który opublikowany został w *Gramatyce współczesnego języka polskiego* [1999]. Autorem tego rozdziału jest Roman Laskowski. Autor definiuje: *wyraz*, *wyraz gramatyczny*, *wyraz tekstowy* oraz *leksem*.

Wyraz to najważniejsza obok zdania jednostka językowa, służąca przekazywaniu informacji za pomocą środków językowych. Równocześnie jest to jedna z najtrudniejszych do zdefiniowania jednostek językowych.

[Laskowski 1999c: 33]³⁹

Wyrazy gramatyczne:

są podstawowymi obiektami językowymi, które podlegają regułom składniowym; one też wchodzą w relacje fleksyjne z innymi wyrazami gramatycznymi jako formy fleksyjne leksemu odmiennego

[Laskowski 1999c: 35]⁴⁰

Wyraz tekstowy.

Przez wyraz tekstowy będziemy w dalszym ciągu rozumieć taki najmniejszy segment tekstu, który spełnia co najmniej jeden z poniższych warunków:

- (1) warunek (względnej) samodzielności – dany segment tekstu może wystąpić samodzielnie jako wypowiedzenie;
- (2) warunek przestawności (inwersji) – segment tekstu wykazuje zmienność pozycji syntagmatycznej w stosunku do innych segmentów tekstu – członów wypowiedzenia (możliwe jest odwrócenie porządku segmentów w co najmniej jednym poprawnym wypowiedzeniu);
- (3) warunek ciągłości – dany segment tekstu jest ciągły w sensie niemożliwości rozbicia go na podsegmenty, między które możliwe jest wstawienie segmentu będącego wyrazem tekstowym na mocy warunku (1) lub (2);
- (4) warunek komplementarności – dany segment tekstu pozostaje w dystrybucji uzupełniającej do innego segmentu na mocy warunków (1)–(3) wyrazem tekstowym;
- (5) warunek fleksyjny – segment otwiera tylko jedno miejsce dla ciągu morfemów fleksyjnych.

[Laskowski 1999c: 36]

Wyraz gramatyczny.

Jednostka funkcjonalna (znacząca), definiowalna poprzez swoje znaczenie, zespół przysługujących jej funkcji syntaktycznych i kategorii morfologicznych, reprezentowana w tekście przez pojedynczy wyraz tekstowy (pisze, domów) lub przez ponadjednoelementowy ciąg wyrazów tekstowych (będzie pisać, boję się, będzie się bał).

[Laskowski 1999c: 39]⁴¹

³⁹ Autor w dalszej części pracy podkreśla, że najczęściej spotykaną w literaturze definicją wyrazu jest „najmniejsza względnie samodzielna jednostka znacząca, zdolna do pełnienia określonych funkcji syntaktycznych” [Laskowski 1999c: 33–34]. Laskowski dodaje, że jest to definicja nieprecyzyjna, a w dalszej części pracy przedstawia, jak bardzo wieloznaczny jest ten termin.

⁴⁰ Autor zaznacza, że zakorzeniony w terminologii termin *forma gramatyczna* bywa używany, jako synonim wyrazu *gramatycznego*. Sam także korzysta zamiennie z tych terminów [Laskowski 1999c: 36].

⁴¹ Wyraz gramatyczny reprezentowany może być przez pojedynczy wyraz tekstowy. Taki wyraz gramatyczny Laskowski nazywa syntetycznym wyrazem gramatycznym (syntetyczną formą gramatyczną) (...)

Leksem autor definiuje

jako abstrakcyjną jednostkę słownikową reprezentowaną w tekście przez pewien niepusty zbiór wyrazów tekstowych lub ciągów wyrazów tekstowych

[Laskowski 1999b: 42]⁴²

Definicje terminów, nawiązujące do tych R. Laskowskiego, znaleźć można w *Zarysie gramatyki polskiej* [2003] Alicji Nagórko. Autorka definiuje pojęcia: *wyraz tekstowy*, *gramatyczny* oraz *słownikowy (leksem)*.

Wyrazem tekstowym nazywamy każdy taki odcinek tekstu (segment), który: 1) nie dopuszcza wstawienia czegokolwiek w środek ani 2) przestawienia kolejności swoich składników, 3) otwiera przy sobie miejsce dla jednego tylko ciągu końcówek gramatycznych

[Nagórko 2003: 74]

Wyraz gramatyczny (albo forma) to wyraz tekstowy wraz z funkcją

[Nagórko 2003: 75]

Wyraz słownikowy to zbiór form gramatycznych, które rozpoznajemy w tekście i które w sposób uporządkowany opisują gramatyki (jako paradygmaty)

[Nagórko 2003: 75]

Wyraz słownikowy albo leksem to abstrakcyjna jednostka systemu gramatyczno-semantycznego, reprezentowana w tekście przez różne formy gramatyczne w przypadku leksemów odmiennych.

[Nagórko 2003: 76]

Najdokładniejsze znane mi opracowanie zagadnienia znajduje się w pracy Piotra Żmigrodzkiego *Wprowadzenie do leksykografii polskiej* [2003]. Autor definiuje takie pojęcia, jak: *wyraz graficzny*, *fonologiczny*, *tekstowy*, *gramatyczny* oraz *leksem* i *jednostka leksykalna*:

Wyraz graficzny (grafemiczny, ortograficzny): ciąg liter między dwiema spacjami

Wyraz fonologiczny: rozumiany jako pojedynczy zespół akcentowy

Wyraz tekstowy: to najmniejszy segment tekstu, taki że jego pozycja jest zmienna w stosunku do innych elementów tego tekstu, możliwe jest wstawienie trzeciego wyrazu między dwa sąsiadujące wyrazy tekstowe. Zakres pojęcia wyraz tekstowy w zasadzie pokrywa się z zakresem pojęcia wyraz graficzny, występują jednak drobne różnice. Przykładowo za pojedyncze wyrazy tekstowe należałoby uznać ciągi: po francusku, po dobru (...), oraz nazwy własne typu Gorzów Wielkopolski

Wyraz gramatyczne: ciąg odpowiadający pojedynczej formie gramatycznej (...) pojedynczymi wyrazami tekstowymi [są] (dodane przez R.M.): pali – palił – będzie palił

[Żmigrodzki 2003: 46]

Wyrazy gramatyczne (...) reprezentowane przez ponadjednoelementowy ciąg wyrazów tekstowych (...) nazywane są analitycznymi (lub peryfrastycznymi) wyrazami (formami) gramatycznymi [Laskowski 1999c: 39].⁴² Autor wyróżnia dwa rodzaje leksemów:

1. leksem nieodmienny, to taki leksem: a) któremu odpowiada jeden tylko wyraz gramatyczny (...) lub b) który jest reprezentowany przez dwa (lub więcej) wyrazy tekstowe będące wariantami formalnymi jednego gramatemu, tzn. mające identyczne znaczenie (funkcję; por. przyimek przez i przeze);
2. leksem odmienny to leksem będący niejednoelementową klasą różnych funkcjonalnie wyrazów gramatycznych. Inaczej mówiąc – leksem odmienny jest to zbiór różnych gramatemów o tożsamym znaczeniu leksykalnym, wyrażanych przez różne wyrazy tekstowe (ciągi wyrazów tekstowych). [Laskowski 1999c: 43].

Leksem to wyraz rozumiany jako jednostka systemu językowego, a nie tekstu. Jest to jednostka abstrakcyjna, element językowy, na który składa się określone znaczenie leksykalne oraz zespół określonych funkcji gramatycznych, a także zespół form językowych, z których pośrednictwem może się one przejawiać w tekstach.

[Żmigrodzki 2003: 47]⁴³

Co ciekawe, Żmigrodzki nie tworzy własnej definicji *jednostki leksykalnej*, ale przytacza definicję, która

pochodzi od M. Grochowskiego (1982: 28): Jednostka leksykalna to ciąg elementów diakrytycznych mających znaczenie globalne, czyli ciąg niepodzielny semantycznie na takie podciągi znaczące, które byłyby elementami klas substytucyjnych niezamkniętych.

[Żmigrodzki 2003: 48]

Żmigrodzki natomiast dodaje, że:

jednostka leksykalna to po prostu pojedynczy element systemu słownikowego

[Żmigrodzki 2003: 48]

Podsumowując, *wyraz* definiowany jest jako:

1. synonim słowa [Podlawska i Płóciennik];
2. synonim leksemu [Podlawska i Płóciennik, Gruszczyński i Bralczyk];
3. najmniejsza jednostka systemu leksykalnego [Podlawska i Płóciennik];
4. podstawowa, samodzielna jednostka tekstu (forma). Ma określoną funkcję składniową (czasem też semantyczną) [Gruszczyński i Bralczyk];
5. jednostka zasobu słownikowego. Ma określoną funkcję składniową (czasem też semantyczną) [Gruszczyński i Bralczyk];
6. względnie samodzielna, znacząca jednostka językowa [Gruszczyński i Bralczyk];
7. jednostka systemu słownikowego języka [Gruszczyński i Bralczyk];
8. najważniejsza obok zdania jednostka językowa, służąca do przekazywania informacji za pomocą środków językowych [Laskowski 1999a].

Nie zgadzam się na traktowanie *słowa* (definicja 1) i *leksemu* (definicja 2), jako synonimów *wyrazu*. Warto w tym miejscu przytoczyć wypowiedź Mirosława Bańki:

W użyciu fachowym słowo to po prostu ciąg liter od spacji do spacji. Nad jego znaczeniem się zastanawiamy, gdyż nie jest istotne. Wyraz natomiast to albo słowo (czasem ciąg słów) rozpatrywany w określonym znaczeniu, albo cały zbiór tak rozumianych słów. Na przykład wyrazem nazywa się zarówno słowo kot rozpatrywane w ogólnie znanym sensie jako mianownik liczby pojedynczej, jak i zbiór wyrazów

⁴³ W celu odróżnienia *leksemów* od innych *wyrazów* Żmigrodzki zapisuje je wersalikami, np. PAPIEROS.

typu: kot, kota, kotu itd. Aby uniknąć tej uciążliwej dwuznaczności, językoznawcy zbiory takie, jak przedstawiony wyżej, nazywają leksemami, a ich elementy – formami wyrazowymi.

[Bańko 2003: www.sjp.pwn.pl/poradnia/haslo/wyraz-a-slowo;3506.html]

Gruszczyński i Bralczyk twierdzą (definicja 4 i 6), że *wyraz* to jednostka samodzielna lub względnie samodzielna. Budzi to pewną wątpliwość nie tylko moją, ale i na przykład Witolda Mizerskiego, według którego samodzielne nie są takie wyrazy, jak: *na*, *do*, *się*, i [Mizerski 2000: 31]. W przypadku niektórych z tych wyrazów, można mieć także wątpliwości, co do ich samodzielności znaczeniowej, bo poza kontekstem trudno stwierdzić, co oznaczają.

Co do reszty definicji (3, 5, 7, 8) są one poprawne w potocznym rozumieniu, ale niewystarczające, jako definicje terminów naukowych.

Powyżej zaprezentowane definicje są bardzo różnorodne, badacze definiowali termin *wyraz* z różnego punktu widzenia (uwzględniając jego funkcję w zdaniu, jego samodzielność lub jego umiejscowienie względem innych jednostek budujących wypowiedź). W potocznym rozumieniu może to wystarczyć, natomiast na potrzeby tej pracy, trzeba skorzystać z dokładniejszych definicji *wyrazu*. Poniższy fragment pracy poświęciłem na prezentację rodzajów i typów wyrazów:

Wyraz tekstowy według badaczy to:

1) odcinek tekstu (segment), który nie dopuszcza wstawienia czegokolwiek w środek i przestawienia kolejności swoich składników oraz otwiera przy sobie miejsce dla jednego tylko ciągu końcówek gramatycznych [Nagórko];

2) najmniejszy segment tekstu, który spełnia przynajmniej jeden z warunków wymienionych na stronie 65 [Laskowski 1999c];

3) najmniejszy segment tekstu, ma zmienną pozycję w stosunku do innych elementów tekstu, można wstawić inny wyraz między dwa wyrazy tekstowe [Żmigrodzki].

Najpełniejsza definicja *wyrazu tekstowego*, to ta stworzona przez R. Laskowskiego. Wymienił on aż pięć warunków, z których przynajmniej jeden musi być spełniony, aby mówić o *wyrazie tekstowym*. Definicje P. Żmigrodzkiego i A. Nagórko zawierają się w definicji R. Laskowskiego, dlatego ją też przyjmuję jako obowiązującą w tej pracy.

Wyraz gramatyczny według badaczy to:

1. wyraz tekstowy wraz z funkcją [Nagórko];

2. synonim formy gramatycznej [Nagórko];

3. podstawowe obiekty językowe, które podlegają regułom składniowym [Laskowski 1999a];

4. jednostka funkcjonalna (znacząca), definiowalna przez swoje znaczenie [Laskowski 1999b];

5. wyraz, który odpowiada jednej formie gramatycznej [Żmigrodzki].

Wszystkie definicje oznaczają mniej więcej to samo, dlatego w tej pracy decyduję się na najogólniejszą, czyli na definicję P. Żmigrodzkiego.

Wyraz słownikowy to:

1. zbiór form gramatycznych, które rozpoznajemy w tekstach [Nagórko];

2. synonim leksemu [Nagórko];

3. abstrakcyjna jednostka systemu gramatyczno-semantycznego [Nagórko].

Leksem to:

1. temat fleksyjny [Gruszczyński i Bralczyk];

2. wyraz słownikowy [Gruszczyński i Bralczyk];

3. abstrakcyjna jednostka systemu słownikowego języka [Gruszczyński i Bralczyk, Laskowski 1999a];

4. jednostka systemu językowego, a nie tekstu. Element językowy, na który składa się znaczenie leksykalne oraz funkcje gramatyczne [Żmigrodzki].

Jak widać w powyższych definicjach termin *wyraz słownikowy*, jako osobne hasło, występuje tylko w definicjach opracowanych przez Alicję Nagórko (*wyraz słownikowy*, definicje 1–3), z kolei Gruszczyński i Bralczyk czynią odwrotnie, czyli *leksem* definiują jako *wyraz tekstowy* (*leksem*, definicja 2). Nie podzielam jednak zdania (patrz definicja 3), że *leksem* może być stosowany jako synonim *tematu fleksyjnego*, ponieważ jest on czymś więcej niż tylko tematem fleksyjnym (patrz definicja Żmigrodzkiego). W pozostałych definicjach nie ma istotnych różnic w definiowaniu *wyrazu słownikowego* i *leksemu*. Często, nawet w pracach językoznawczych, są one używane synonimicznie. W tej pracy będę używał wyłącznie terminu *leksem*, ponieważ jest to krótsza od *wyrazu słownikowego* wersja odnosząca się do tego samego pojęcia. Przyjmuję definicję *leksemu* w rozumieniu Żmigrodzkiego.

Wyraz *graficzny*⁴⁴ w pracy będzie rozumiany, tak jak go definiuje Żmigrodzki „ciąg liter między dwiema spacjami” [Żmigrodzki], natomiast wyraz *fonologiczny*, to „pojedynczy zespół akcentowy” [Żmigrodzki].

Zaś *jednostka leksykalna*, to „pojedynczy element systemu słownikowego” [Grochowski]. I tak w tej pracy będzie rozumiana.

4.4. POLE LEKSYKALNE

Najważniejszym terminem w mojej pracy jest *pole leksykalne (wyrazowe)*. Jest to termin kluczowy, ponieważ posługuję się nim w całej części analitycznej oraz we fragmencie części metodologicznej. Poprawne zdefiniowanie go jest więc bardzo ważne, ponieważ bez tego niemożliwe (albo bardzo utrudnione) byłoby zrozumienie moich badań. Według A. Markowskiego *polem leksykalnym* nazwać można „grupy wyrazów powiązanych znaczeniowo, ale niebędących synonimami, lecz tworzących relacje innego typu” [2012: 107].

Markowski podkreśla także, że *pola leksykalne* odnoszą się do jednego fragmentu rzeczywistości, czyli elementem łączącym grupy wyrazów jest określony desygnat.

W przypadku mojej pracy, *polami leksykalnymi* nazywam te grupy wyrazów, które skoncentrowane są wokół danego znaczenia oraz połączone różnego rodzaju relacjami (przedstawiam je poniżej). W obrębie jednego pola relacje te mogą zachodzić pojedynczo lub łącznie.

Relacje między elementami pól wyrazowych:

- 1) Dane znaczenie realizowane jest w różnych częściach mowy np.: *calak* (rzeczownik), *calakować* (czasownik), *calakowy* (przymiotnik) itd.
- 2) Dane znaczenie kodowane jest w różnych leksemach, pozostających względem siebie w relacji ekwiwalencji, np.: *blaszak* i *pecet*, *patch* i *latka*, *skin* i *skórka* itd.
- 3) Dane znaczenie jest urzeczywistniane w różnych wariantach graficznych/ortograficznych, np.: *crap* i *krap*, *cross-controller* i *cross-controler*, *crossover* i *cross-over* itd.

⁴⁴ W dalszej pracy bardzo często posługuję się terminem *wyraz*, który jest rozumiany właśnie jako *wyraz graficzny* według definicji Żmigrodzkiego. Czasami jednak rezygnuję z członu określającego (*graficzny*), aby zwiększyć czytelność wypowiedzi, jednak nadal jest to *wyraz* w rozumieniu *wyraz graficzny*.

Tak rozumiane *pola leksykalne (wyrazowe)*⁴⁵ będę badał w dalszej części pracy. Ponadto badam też liczbę wystąpień elementów tych pól (tzn. leksemów) w konkretnych tekstach (a więc jako wyrazy tekstowe). Leksemy te (o ile są odmienne) realizowane są jako systemowo przewidziane formy fleksyjne. Moje badania operują więc na obu płaszczyznach języka: abstrakcyjnego systemu językowego i jego konkretnej realizacji jako mowy (w rozumieniu F. de Saussure'a).

⁴⁵ Zdarza się, że w pracy pominę człon określający (*leksykalne* lub *wyrazowe*) i posłużę się tylko terminem *pole*. Jeśli tak się stanie, to zaznaczam, że wtedy także odnoszę się do powyższego rozumienia *pola*.

CZEŚĆ
METODOLOGICZNA

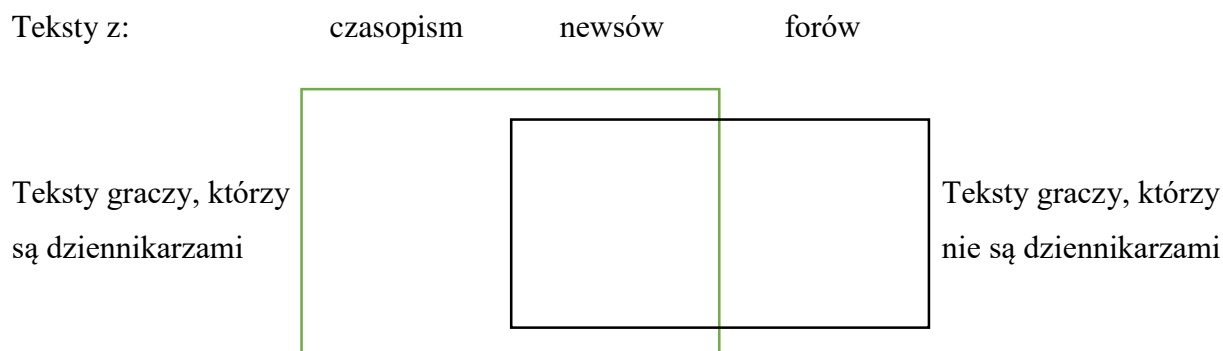
5. CHARAKTERYSTYKA MATERIAŁU

5.1. KRYTERIA POZYSKANIA MATERIAŁU

Przed rozpoczęciem badań nad pisaną odmianą socjolektu graczy wideo zdecydowałem, że najlepiej byłoby włączyć do korpusu teksty napisane zarówno przez *graczy konsolowych*, jak i *komputerowych* a dodatkowo, aby badanie było bogatsze, pozyskać materiał reprezentujący trzy różne sposoby publikowania tekstów. Chciałem dzięki temu zachować balans między tekstami napisanymi przez graczy, którzy są dziennikarzami związanymi zawodowo z jednym z czasopism a tekstami graczy, którzy nie pracują jako dziennikarze. Wybór padł na czasopisma (w nich dominują teksty dziennikarzy), newsy⁴⁶ opublikowane na stronach internetowych (publikują je głównie dziennikarze, ale komentują gracze) oraz wpisy z forów internetowych (dominują wpisy graczy, którzy nie są dziennikarzami).

Było to pierwsze kryterium wyboru materiału do badań, czyli kryterium pochodzenia tekstu. Graficznie kryterium to można przedstawić następująco:

Grafika 4. Kryterium podziału ze względu na pochodzenie tekstu



Źródło: opracowanie własne

Moim zamiarem było także, aby do badań włączony został materiał, który zawiera teksty napisane przez *konsolowców* oraz *komputerowców*. O podziale środowiska na graczy związanych z komputerami lub konsolami pisałem już na stronie 12. Z podziałem tym wiąże się także to, że każda z grup posiada swoje własne czasopismo, stronę internetową oraz forum.

Z *komputerowcami* związane jest czasopismo „CD-Action”, natomiast z *konsolowcami* „PSX Extreme”.

⁴⁶ Czym one są, wyjaśnię w dalszej części pracy.

Pierwsze z czasopism wybranych do badań („CD-Action”), ukazuje się nieprzerwanie od 1996 roku. Jest ono uważane za miesięcznik, jednakże zamiast 12 numerów wydawanych w przeciągu roku ich jest 13. Dodatkowy 13 numer został nazwany numerem specjalnym. Czasopismo skierowane jest głównie do graczy komputerowych, ponieważ teksty w nim publikowane dotyczą wyłącznie gier komputerowych, bądź multiplatformowych, to znaczy takich, które ukazały się na kilka platform, czyli np. nie tylko na komputer, ale także na konsole. Przy czym trzeba dodać, że gra multiplatformowa, której tekst zostanie opublikowany w czasopiśmie, musi zostać wydana także na komputer, nie może to być gra, która jest przeznaczona na wiele różnych platform z pominięciem komputera. Do czasopisma dodawana jest płyta (lub płyty) DVD/CD z grami na komputer oraz innymi dodatkami np. demami czy modami. Średni nakład czasopisma wynosi 127.847 egz., rozpowszechnianie płatne to 77.555 egz.⁴⁷, natomiast czytelnictwo szacuje się na poziomie 368.000 osób⁴⁸.

Drugie czasopismo – „PSX Extreme” ukazuje się od 1997 r. Stworzone zostało przez grupę pasjonatów, którzy postanowili zrobić magazyn wyłącznie dla graczy konsolowych, w opozycji do istniejącego już wtedy czasopisma „CD-Action”. Potwierdzenie zaklasyfikowania czasopisma jako przeznaczonego dla graczy konsolowych znaleźć można także w jego podtytule – największy w Polsce magazyn o konsolach. Trudno ustalić nakład czasopisma, bo takie informacje nie są publicznie udostępniane przez wydawnictwo, jednak na podstawie przesłanek, które dostępne są w Internecie, szacować można, że czasopismo ma sprzedaż na poziomie 30.000 egz., natomiast nakład to 45.000 egz. Dane te pochodzą ze strony www.gieldamedia.pl/archiwum/652-psx-extreme [8.04.2017], a dotyczą przełomu 2009 i 2010 roku.

W przypadku stron internetowych związanych z daną grupą zdecydowałem, że nie włączę wszystkich opublikowanych tam treści, a ograniczę się jedynie do newsów, czyli krótkich informacji na temat najnowszych wydarzeń ze świata, w tym przypadku ze świata gier wideo (są one osobno wydzieloną częścią stron internetowych). Pomiąłem zatem np. artykuły i recenzje tam publikowane, ponieważ te pozyskałem z czasopism drukowanych i nie chciałem powielać materiału. Trzeba podkreślić, że strony internetowe pod względem szybkości publikowania materiałów zdecydowanie przewyższają czasopisma drukowane, efekt tego

⁴⁷ Według danych opublikowanych na stronie: www.reklama.bauer.pl/magazine.php/sid,7/mid,24/title,CDAAction/ (źródło informacji: ZKDP; średni nakład jednorazowy, rozpowszechnianie płatne razem; marzec 2015 - luty 2016).

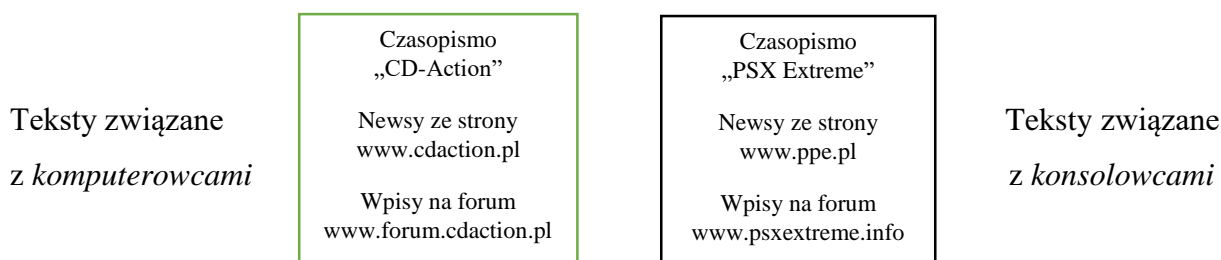
⁴⁸ Według danych opublikowanych na stronie: www.reklama.bauer.pl/magazine.php/sid,7/mid,24/title,CDAAction/ (źródło informacji: Polskie Badania Czytelnictwa; kwiecień 2015 - marzec 2016; CPW; grupa wszyscy 15+; n = 20 056; realizacja Millward Brown).

widać w częstym dodawaniu newsów, które pod względem liczebności dominują na tych stronach. Newsy są dobrym uzupełnieniem materiału pozyskanego z czasopism drukowanych, które, co prawda także je publikują, ale stanowią one szczątkową zawartość każdego z numerów czasopisma drukowanego. Nie stanowią one głównej zawartości czasopism, stanowią natomiast główną zawartość wspomnianych już stron internetowych. Newsy wybrane do pracy pochodzą ze stron: www.ppe.pl i www.cdaction.pl. Pierwsza związana jest z *konsolowcami*, natomiast druga z *komputerowcami*.

Poza tekstami z czasopism i newsów do korpusu włączone zostały także wpisy z forów internetowych – www.psxextreme.info (forum związane z *konsolowcami*), www.forum.cdaction.pl (forum związane z *komputerowcami*).

Tak dobrany zbiór tekstów pozwolił także na jeszcze inny podział materiału, a mianowicie ze względu na powiązanie tekstów z daną grupą środowiska graczy wideo. Innymi słowy kryterium podziału był użytkownik. Graficznie kryterium to można zaprezentować następująco:

Grafika 5. Kryterium podziału ze względu na pochodzenie tekstu



Źródło: opracowanie własne

Reasumując, teksty pozyskałem z 6 źródeł, które przedstawiłem w powyższych grafikach. Wyniki badań ilościowych będę zatem rozpatrywał, oczywiście tam gdzie będzie to możliwe, w odniesieniu do jednej z sześciu/ wszystkich sześciu części korpusu, w odniesieniu do części korpusu związanych z *komputerowcami* lub *konsolowcami* oraz w odniesieniu do części korpusu związanych z czasopismami, newsami lub wpisami na forach.

5.2. ZAKRES POZYSKANEGO DO BADAŃ MATERIAŁU

Kolejnym ważnym pytaniem, na które musiałem znaleźć odpowiedź przed rozpoczęciem badań było to, jak duży materiał wybrać, aby wyniki były reprezentatywne, miarodajne i satysfakcjonujące. Punktem odniesienia do ustalenia ilości materiału, który powinienem włączyć do badań, były inne prace badawcze o nastawieniu ilościowym.

W 1972 roku W. Pisarek stworzył, do badań nad gatunkami dziennikarskimi, korpus zbudowany z 127.035 wyrazów (dokładniej pisałem o tym na stronie 48). Grupa pod redakcją m.in. A Lewickiego, która badała słownictwo współczesnego języka polskiego, zdecydowała się na korpus zbudowany ze 100.000 wyrazów dla jednego ze stylów polszczyzny pisanej. Łącznie zbadali oni 5 stylów, czyli cały korpus zbudowany był z 500.000 wyrazów (więcej na ten temat pisałem na stronach 49–50), natomiast zespół Z. Kurzowej, który analizował język telewizji zdecydował się na 189.000 słowoform.

Narodowy Korpus Języka Polskiego zawiera półtora miliarda słów (www.nkjp.pl), natomiast korpus PWN zbudowany jest z kilku części, których dokładna liczba wyrazów nie jest znana. Na pewno zawiera on w sobie 18.000.000 słów tekstowych pozyskanych z „Rzeczpospolitej”, a poza tym „składa się z 2.070 próbek tekstu wielkości od 1 arkusza do 6 arkuszy o łącznej wielkości 175 MB” (www.sjp.pwn.pl/korpus/struktura). Więc, jeśli korpus tekstowy waży 175 MB, to musi być on olbrzymim zbiorem danych.

Postanowiłem, że mój korpus pod względem liczebności wyrazów będzie plasował się między współczesnymi, miliardowymi korpusami PWN czy też Narodowym Korpusem Języka Polskiego, a małymi, dawnymi korpusami, które wykorzystywane były do badań ponad 40 lat temu.

Początkowo założyłem, że stworzę korpus obejmujący publikacje z 15 ostatnich lat przed momentem rozpoczęcia badań, czyli począwszy od 1 stycznia 2000 roku, a skończywszy na 31 grudnia 2015 roku. Jednak aby badania mogły zostać zrealizowane, potrzebowałem dodatkowego dofinansowania (np. na zakup szybkiego komputera i oprogramowania do niego), wystąpiłem więc o grant Preludium do NCN na realizację mojego projektu. Nie został mi on przyznany, ale otrzymałem recenzje, które uświadomiły mi, że tak dużego materiału nie da się w pojedynkę zbadać. Zdecydowałem więc, że badania obejmą teksty opublikowane w czasopismach począwszy od pierwszego numeru z 2011 roku, do ostatniego numeru 2015, czyli do badań włączyłem ostatnie pięć pełnych roczników czasopism wydanych przed intensywnym rozpoczęciem moich analiz w 2016 r. Reasumując, w skład korpusu wejdą numery czasopisma „PSX Extreme” od 161 do 220 i numery „CD-Action” z przedziału od 186 do 250 z pominięciem 198, 211, 224, 237 i 250. Pomiąłem te numery, ponieważ chciałem zachować równowagę w liczbie czasopism wybranych do badań – po 12 numerów z jednego rocznika danego czasopisma. Pominięcie numerów specjalnych umożliwiło mi to⁴⁹.

⁴⁹ Obecnie poza czasopismami „CD-Action” i „PSX Extreme”, które wydawane są dla graczy, istnieje jeszcze także czasopismo „Pixel”. W swoich badaniach zdecydowałem się jednak pominąć je, ponieważ wydawane jest

Do badań włączyłem więc 60 numerów czasopisma „PSX Extreme” oraz tyle samo numerów czasopisma „CD-Action”, po 6.000 newsów ze stron ppe.pl i cdaction.pl oraz wpisy z forów internetowych.

5.3. SKRÓTOWE NAZWY POSZCZEGÓLNYCH CZĘŚCI KORPUSU

Powyżej wspomniałem, że cały korpus składa się z sześciu części, które badam w różnych konfiguracjach. Aby zwiększyć czytelność pracy i uniknąć żmudnego podawania długich nazw części korpusu (np. część korpusu, która zbudowana jest z tekstów czasopisma „PSX Extreme”), zdecydowałem, że w momencie odwoływania się do nazwy danego korpusu najczęściej posługiwać będę się jego skrótową nazwą.

Skróty te wyglądają następująco:

- psx** – część całego korpusu, która zbudowana jest z tekstów czasopisma „PSX Extreme”,
- cda** – część całego korpusu, która zbudowana jest z tekstów czasopisma „CD-Action”,
- npsx** – część całego korpusu, która zbudowana jest z newsów opublikowanych na stronie www.ppe.pl,
- ncda** – część całego korpusu, która zbudowana jest z newsów opublikowanych na stronie www.cdaction.pl,
- fpsx** – część całego korpusu, która zbudowana jest z wpisów opublikowanych na forum www.psxextreme.info
- fcda** – część całego korpusu, która zbudowana jest z wpisów opublikowanych na forum www.forum.cdaction.pl.

W tym rozdziale pokrótce scharakteryzowałem zebrany do badań materiał, natomiast w następnym rozdziale przedstawię etapy pozyskiwania materiału tekstowego oraz zaprezentuję sposoby pracy z nim.

dopiero od lutego 2015 roku, a moje badania dotyczą także lat wcześniejszych. Gdybym je włączył do badań zaburzone zostałyby ich proporcje.

6. METODOLOGIA POZYSKIWANIA I PRACY Z MATERIAŁEM

ETAP 0. UWAGI TECHNICZNE

W tej części pracy zaprezentuję sposób, w jaki pozyskałem materiał badawczy, następnie przedstawię, jak z niego stworzyłem korpus tekstów, oraz w jaki sposób z tym korpusem pracowałem. Na wstępie chcę zaznaczyć, że w niniejszym tekście dominuje opis pierwszoosobowy, ponieważ jest to swoisty dziennik pracy. Opisałem w nim, jakie czynności po kolei podejmowałem, aby zrealizować cel mojej pracy, a więc poddać analizie ilościowej socjolekt graczy wideo. Przyjęcie takiego sposobu opisu ułatwiło mi wytłumaczenie kolejno podejmowanych czynności i myślę, że ułatwi to czytelnikom odbiór tego tekstu.

Cel, jaki przyświecał mi podczas tworzenia tej części to dążność do opracowania takich metod postępowania z materiałem, aby ułatwiały one pracę przyszłym badaczom i były możliwe do powtórzenia np. przez studentów, magistrantów czy doktorantów, którzy chcą zbadać jakiś element językowy, wykorzystując samodzielnie utworzone korpusy tekstów. Chciałem także przedstawić swoją metodę badawczą po to, aby mogła ona być udoskonalana przez innych badaczy, co być może doprowadzi do wypracowania najlepszego z możliwych sposobów tworzenia i pracy z korpusami tekstów, a przede wszystkim prostego i powszechnie dostępnego.

Przeprowadzając wstępne rozeznanie w temacie tworzenia korpusów, a miało ono miejsce jeszcze przed rozpoczęciem badań, nie znalazłem żadnej metody/instrukcji, która pomogłaby mi w stworzeniu takowego. Nie znalazłem także opisu metod pracy z już utworzonym zbiorem tekstów. W trakcie trwania moich badań, w 2017 roku, udostępnione zostało narzędzie umożliwiające tworzenie korpusów tekstów – *Korpusomat*. Po zapoznaniu się z tym programem i wypróbowaniu go, uznałem, że nie znajdzie on zastosowania w moich badaniach, ponieważ nie obsługiwał dużych plików tekstowych, czyli nie dało się przy jego pomocy stworzyć dużego korpusu tekstów, a taki właśnie był mi potrzebny. W instrukcji do tego programu jego autorzy piszą:

czas przetwarzania przeciętnej wielkości książki o objętości ok. 80-100 tys. słów powinien wynieść około 2-3 minut, choć częściowo zależy to również od aktualnego obciążenia serwera. Obecnie maksymalny czas przetwarzania pliku wynosi 10 minut – zadania dłuższe zakończą się niepowodzeniem...

[www.korpusomat.nlp.ipipan.waw.pl/manual, dostęp: 22.12.2017]

W moim przypadku próba zaimplementowania dłuższego tekstu zakończyła się właśnie niepowodzeniem. Niemniej doceniam to, że taki program istnieje i mam nadzieję, że będzie

dalej rozwijany. W tej pracy jednak nie miałem możliwości z niego skorzystać, ponieważ dysponowałem zbyt dużymi plikami.

Poza *Korpusomatem*, który dostępny jest w Internecie, istnieją inne, bardziej zaawansowane programy do obsługi korpusów. Przykładowo jeden z nich posiada PWN, ale niestety nie jest on powszechnie dostępny. Do innych programów także nie miałem dostępu.

Stałem przed koniecznością samodzielnego opracowania metody pozyskania i pracy z korpusami tekstów. Chciałem, żeby owa metoda nie opierała się na wykorzystaniu drogich, komercyjnych programów, tylko na programach darmowych lub na ich wersjach testowych. Wyjątkiem od tej zasady był program *Word*, który posiada większość studentów, magistrantów czy doktorantów oraz program do rozpoznawania tekstu w grafikach, który po wielu testach na jego bezpłatnych odpowiednikach, okazał się najlepszy. Potrzebny on był w kluczowym momencie tworzenia korpusu tekstów, więc nie mogłem sobie pozwolić na utratę jakości, decydując się na bezpłatne odpowiedniki. Stawiając sobie za cel opracowanie możliwie taniej metody tworzenia korpusów i pracy z nimi chciałem, żeby badania tego typu były częściej podejmowane i żeby pieniądze nie stanowiły przeszkody, a nawet blokady, w ich realizacji.

Opis metody pracy wzbogaciłem wieloma zdjęciami, które stanowią integralną część tekstu głównego i są jego niezbędnym elementem. Dodanie zdjęć pozwoliło mi uniknąć obszernych opisów, w których omawiałbym, gdzie dany element, przycisk czy też zakładka programu się znajduje. Zamiast tego w tekście dodałem odsyłacz do danego zdjęcia, na którym zazaczyłem kolorem omawiany element.

Opisane w tekście czynności postępowania z plikami lub programami, w celu przyspieszenia pracy, były wykonywane łącznie na 3 komputerach typu laptop, z czego na różnych etapach pracy tylko 2 z nich były sprawne. Czynności te z powodzeniem mogłyby być wykonane także na jednym komputerze, tylko zajęłyby one trochę więcej czasu.

Pierwszym komputerem wykorzystanym do badań był *Fujitsu Siemens Amilo Pi 3540*. Był to najstarszy z posiadanych przeze mnie komputerów, który pod koniec 2016 roku przestał działać i zastąpiony został przez *Lenovo Yoga 300*. Ostatnim komputerem, na którym prowadziłem badania był, zakupiony dzięki dotacji uniwersyteckiej dla młodych naukowców, *Lenovo V310*. Specyfikacja komputerów wygląda następująco:

1 komputer	2 komputer	3 komputer
Fujitsu Siemens Amilo Pi 3540	Lenovo Yoga 300	Lenovo V310
Procesor: Intel Core 2 Duo P7350	Procesor: Intel Pentium N3700	Procesor: Intel Core i5-7200U
Pamięć: 4GB (DDR2)	Pamięć: 4 GB (SO-DIMM DDR3)	Pamięć: 8 GB (SO-DIMM DDR4)
Dysk: 500 GB SATA 5400 obr.	Dysk: 500 GB SATA 5400 obr.	Dysk: 500 GB SATA 5400 obr.,
Grafika: GeForce 9300M GS	Grafika: Intel HD Graphics	Grafika: Intel HD Graphics 620
System operacyjny: Windows 7	System operacyjny: Windows 10	System operacyjny: Windows 10

Tabela 1. Zestawienie komputerów, które wykorzystane zostały do przeprowadzanie moich badań

Poza wymienionymi powyżej laptopami, sporadycznie korzystałem także z komputerów w Zakładzie Lingwistyki Stosowanej i Kulturowej Uniwersytetu Łódzkiego, czyli w jednostce, w której moja praca powstała.

Poniżej opiszę sposób pozyskiwania materiałów z edytowalnych z czasopism drukowanych oraz ze stron internetowych.

ETAP 1. POZYSKIWANIE EDYTOWALNYCH MATERIAŁÓW TEKSTOWYCH Z CZASOPISM DRUKOWANYCH

Do stworzenia korpusu tekstów z czasopism, na którym mógłbym pracować, niezbędne okazało się posiadanie tekstów w wersji elektronicznej⁵⁰. Rozwiązaniem problemu z posiadaniem takich wersji, byłoby zakupienie czasopism przygotowanych z myślą o dystrybucji cyfrowej. Nie było to jednak możliwe, ponieważ „CD-Action” nie posiada takowej wersji⁵¹. Natomiast czasopismo „PSX Extreme” posiada swoją wersję elektroniczną dopiero od numeru 167. Ostatecznie podstawą badań stały się wersje drukowane czasopism, które prenumeruję niemalże od momentu powstania. Zebrane numery czasopism wykorzystane do badań zaprezentowałem na zdjęciu 1.



Zdjęcie 1. Kolekcja czasopism

Zdjęcie 2. Kolekcja czasopism (bez grzbietów)

PRACE PRZYGOTOWAWCZE

Zebrawszy wszystkie numery czasopism, w celu ułatwienia sobie dalszej pracy, udałem się do drukarni i poprosiłem o odcięcie grzbietów przy pomocy gilotyny. W efekcie uzyskałem luźne kartki (patrz zdj. 2), z którymi na dalszym etapie tworzenia korpusu mogłem efektywniej

⁵⁰ Badania na tekstach drukowanych z pewnością nie zmieściłyby się w przedziale czterech lat studiów doktoranckich.

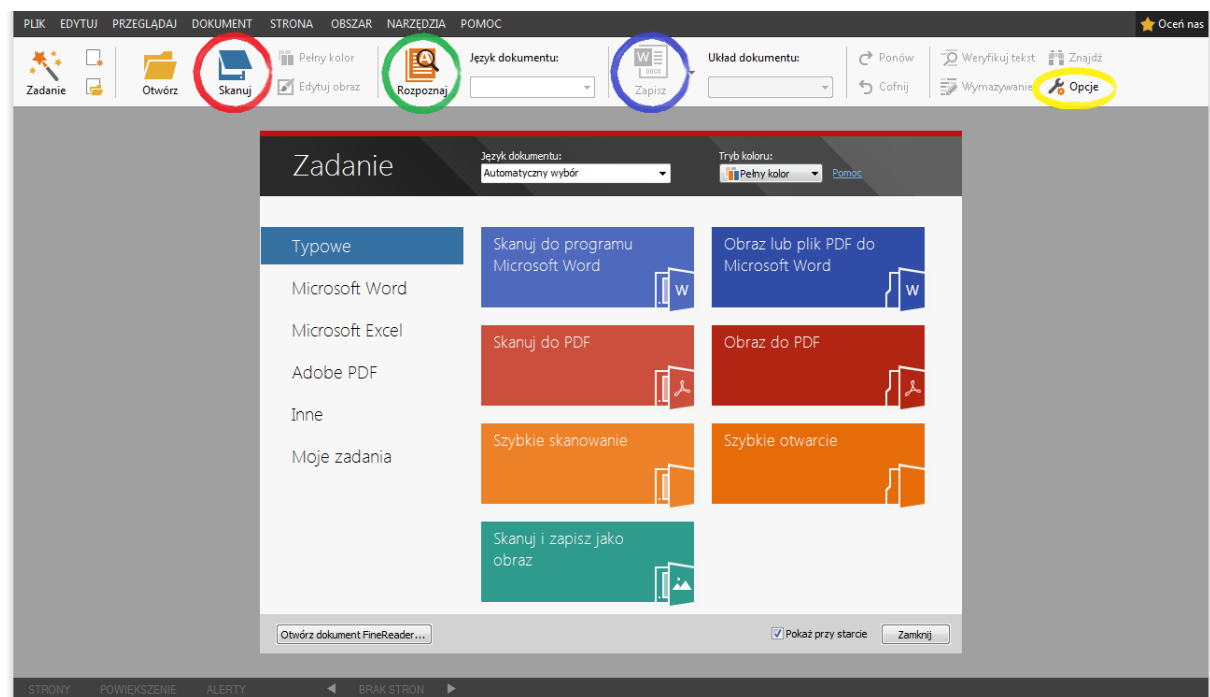
⁵¹ Powodem jest zapewne to, że do czasopisma dodawana jest płyta z obszerną zawartością danych, która skutecznie uniemożliwia wydawanie czasopisma w wersji elektronicznej lub na tyle je utrudnia, że wydawcy zdecydowali się nie wprowadzać takiej formy sprzedaży.

pracować. Niestety jest to metoda bardzo inwazyjna, ponieważ niszczy się czasopisma, ale zważywszy na to, ile czasu można dzięki temu zaoszczędzić, zdecydowałem się na ten krok.

PRACA Z PROGRAMEM ABBYY FINEREADER 12 PROFESSIONAL EDITION

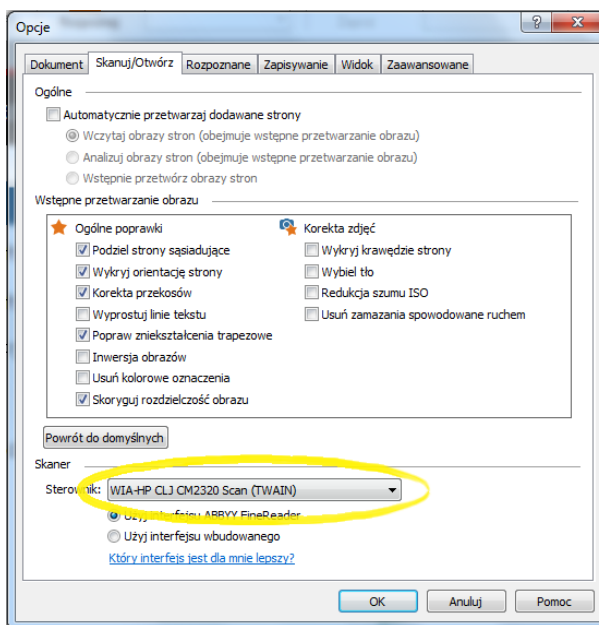
Kolejnym etapem było zeskanowanie wszystkich stron czasopism. I dzięki temu, że posiadałem luźne kartki, mogłem skorzystać z automatycznego podajnika do skanera w moim urządzeniu wielofunkcyjnym (*HP color laserjet CM2320fxi MFP*). Wadą tego urządzenia jest to, że nie posiada ono funkcji *dupleks*, czyli automatycznego obracania kartek podczas skanowania. Przez to mogłem skanować tylko jedną stronę danego numeru czasopisma, natomiast żeby zeskanować drugą, musiałem obracać cały plik kartek i skanować je ponownie.

Do skanowania i jednocześnie do dalszej pracy z pozyskanymi obrazami wykorzystałem program *ABBYY FineReader 12 Professional Edition*, który jest kompleksowym narzędziem do rozpoznawania znaków (liter, cyfr, znaków interpunkcyjnych...) we wszelkiego rodzaju grafikach.

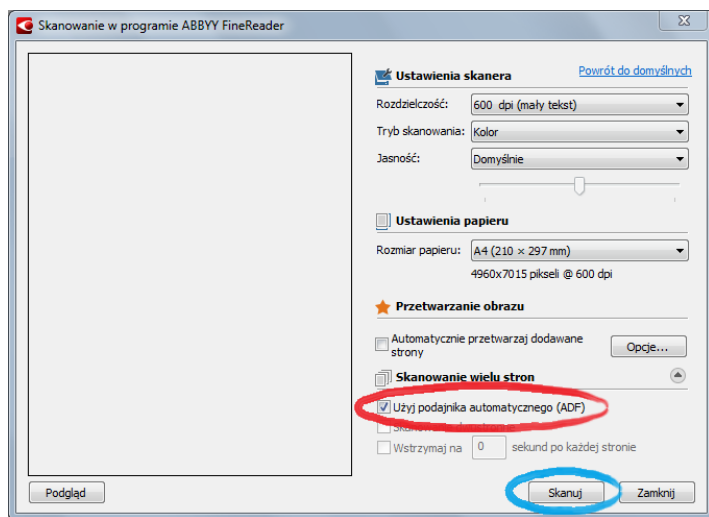


Zdjęcie 3. Główne okno programu ABBYY FineReader 12 Professional Edition

Po podłączeniu urządzenia wielofunkcyjnego do komputera i przygotowaniu czasopism uruchomiłem program *ABBYY* (zdz. 3, główne okno programu) i pierwsze co zrobiłem, to wszedłem w *opcje* (zdz. 3, żółta elipsa) oraz ustawiłem sterownik skanera na *WIA HP CLJ CM2320 Scan (TWAIN)* (zdz. 4, żółta elipsa), pomimo że pierwotnie ustawiony był *HP CLJ CM2320 Scan (WIA)*⁵². Zmiana ta wpłynęła na proces skanowania w taki sposób, że nie musiałem, po zeskanowaniu każdej strony ponownie naciskać przycisku *skanuj*, aby pozyskać kolejny obraz. Dzięki tej zmianie program czynił to automatycznie, aż do momentu, gdy zeskanowane zostały wszystkie strony umieszczone w podajniku. Gdy już wszystko ustawiłem, uruchomiłem skanowanie, naciskając w głównym oknie programu przycisk *skanuj* (zdz. 3, czerwona elipsa). Następnie zazaczyłem opcję *użyj podajnika automatycznego*, aby kartki pobierane były nie ze skanera płaskiego, ale z podajnika (zdz. 5, czerwona elipsa).



Zdjęcie 4. Opcje programu ABBYY FineReader 12...



Zdjęcie 5. Skanowanie przy pomocy ABBYY FineReader 12...

Właściwy proces skanowania rozpocząłem klikając *skanuj* (zdz. 5, niebieska elipsa). Po jednokrotnym zeskanowaniu pliku kartek uzyskałem wyłącznie obrazy stron nieparzystych (np. 3, 5, 7), musiałem więc obrócić plik z kartkami i ponownie uruchomić skanowanie, aby pozyskać obrazy stron parzystych (2, 4, 6). Jest to jedna z konsekwencji braku posiadania *duplexu*.

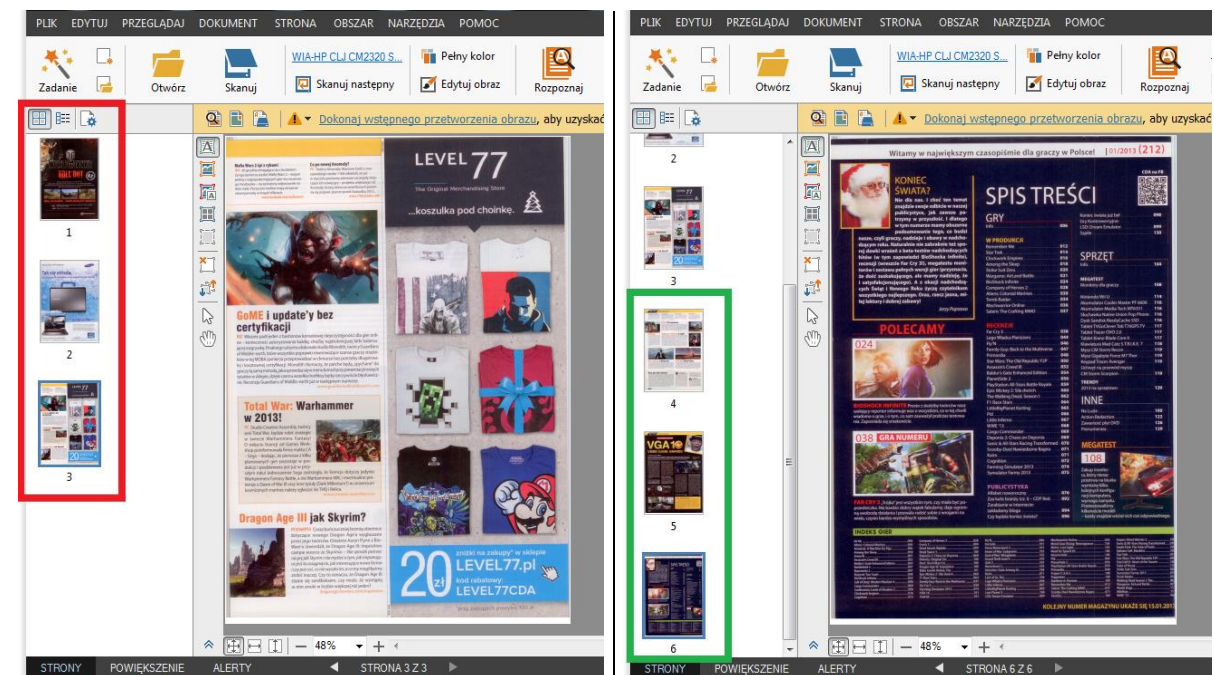
Po zeskanowaniu czasopisma powyżej opisaną metodą, obrazy stron zostały poprzestawiane, mówiąc dokładnie nie były poukładane według kolejności, jaka była

⁵² Trzeba podkreślić, że skanery w różnych urządzeniach wielofunkcyjnych mogą posiadać różne sterowniki, dlatego przed rozpoczęciem skanowania całych czasopism trzeba wykonać próbę na kilku stronach i wybrać najlepszy sterownik.

w czasopiśmie. Aby lepiej to zobrazować, stworzyłem tabelę, gdzie w pierwszym wierszu ułożyłem strony, w takiej kolejności, jak w czasopiśmie drukowanym, a w drugim wierszu tak, jak ułożone zostały one przez program po zakończeniu dwukrotnego skanowania (patrz tabela 2). Poniższa tabela służy tylko zobrazowaniu zaistniałego problemu, dlatego ograniczyłem się do zaprezentowania 3 karetek, a 6 stron, z czasopisma „CD-Action”.

KOLEJNOŚĆ STRON W CZASOPIŚMIE	 s. 3 ⁵³	 s. 4	 s. 5	 s. 6	 s. 7	 s. 8
KOLEJNOŚĆ STRON PO ZESKANOWANIU	 s. 3	 s. 5	 s. 7	 s. 8	 s. 6	 s. 4

Tabela 2. Wybrane strony z czasopisma „CD-Action”



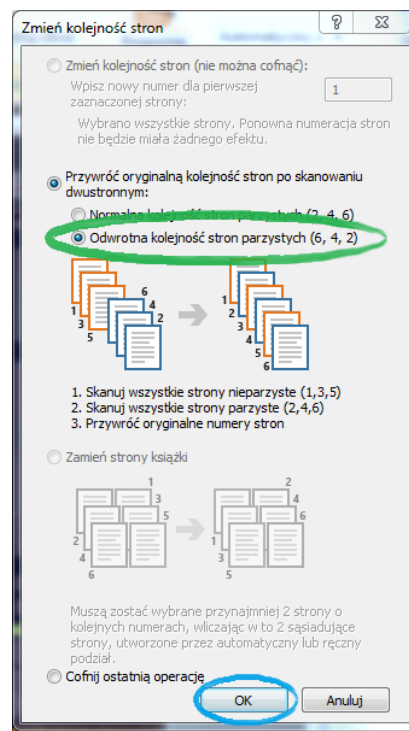
Zdjęcie 6. Efekt skanowania stron 3, 5 i 7

Zdjęcie 7. Efekt skanowania stron 8, 6 i 4

⁵³ W tabeli brakuje pierwszej i drugiej strony, ponieważ numeracja właściwej części czasopisma zaczyna się od strony trzeciej. Pierwsze dwie strony to okładka.

Następnym etapem pracy było przywrócenie właściwej kolejności stron. Sześć zeskanowanych stron, które służą jako przykład do prezentacji sposobu zamiany ich kolejności, znajdowało się w głównym oknie programu, po jego lewej stronie (patrz zdj. 6, czerwona ramka i zdj. 7, zielona ramka). Kliknąłem na zeskanowaną stronę, która przez program została oznaczona numerem jeden, a następnie trzymając klawisz *shift*, nacisnąłem na ostatni z pozyskanych skanów, ten oznaczony numerem sześć. W ten sposób wszystkie zeskanowane strony zostały zaznaczone. Następnie najechałem kursorem na jedną z grafik i kliknąłem prawym przyciskiem myszy, dzięki czemu uruchomione zostało *menu podręczne* programu, z którego wybrałem opcję *zmień kolejność stron*. Otworzyło się nowe okno (patrz zdj. 8), w którym zazaczyłem pole *odwrotna kolejność stron parzystych* (zdj. 8, zielona elipsa) i potwierdziłem wybór przyciskiem *OK* (zdj. 8, niebieska elipsa). W ten sposób przywrócona została oryginalna, taka jak w czasopiśmie drukowanym, kolejność.

Kolejnym krokiem było uruchomienie głównej funkcji programu *ABBYY FineReader*, czyli rozpoznawania tekstu z pozyskanych w wyniku skanowania grafik. Kliknąłem przycisk *rozpoznawaj* (zdj. 3, zielona elipsa) i po kilkudziesięciu minutach, w zależności od szybkości używanego komputera, program rozpoznał i wygenerował tekst edytowalny. Po zakończeniu rozpoznawania tekstu, kliknąłem przycisk *zapisz* (zdj. 3, niebieska elipsa) i po kolejnych kilkunastu minutach, program stworzył plik wordowski (*docx*). Powyższe czynności powtórzyłem jeszcze 119 razy i zakończyłem pracę z tym programem.



Zdjęcie 8. Zmiana kolejności stron w programie ABBYY...



Nowy Hitman z Monteralu

PC X360 PS3 Co rok pror... Hitman. Tore Blystad, jeden z ojców sukcesu Hitmana: Rozgrzeszenia, zdradził, że kolejna część cyklu jest już w produkcji. Po raz pierwszy w historii odpowiada za nią jednak nie IO Interactive, a inne studio – Square Enix Montreal. Blystad nie ukrywa, że inspiracją dla nich stanowiła seria Call of Duty, obrabiana na przemian przez Treyarcha i Infinity Ward. Jego zdaniem zyskuje na tym także sama gra, bo zespoły inspirują się nawzajem, dzięki czemu formuła się nie wypala. Nic, czekamy na Hitmana: Zadośćuczynienie.

www.hitman.com



Arma III i Václav Klaus

BIZ Greckiej tragedii Bohemii ciąg dalszy. Blisko trzy miesiące po aresztowaniu Ivana Buchty i Martina Pezlara przez policję wojskową pod zarzutem szpiegostwa (robili fotki bazy wojskowej na Lemnos) w sprawę włączył się prezydent Czech Václav Klaus. Głowa państwa zwróciła się listownie do prezydenta Grecji Karolosza Papouliasa z prośbą, by zechciał „szczególnie uważnie przyjrzeć się temu niefortunnemu zdarzeniu przy uwzględnieniu znakomych stosunków między narodami, tak by nie rzuciło niepotrzebnie cienia na wzajemne relacje”. Czy pomoże – zobaczymy.

www.arma3.com

Next-gen dużo droższy?

BIZ Temperatura spekulacji wokół nowych konsol rośnie odwrotnie proporcjonalnie do tej za oknem. Tym razem nastroje podsyłał Tim Sweeney z Epic Games, który w trakcie Montreal International Game Summit stwierdził, że koszt produkcji gier na kolejną generację sprzętu może „wzrosnąć trzy-, cztero-, a nawet pięciokrotnie”. Myśl absolutnie przerażająca: budżety sięgające setek milionów dolarów w praktyce oznaczają przecież krach naszej branży. Na szczęście ta ponura przepowiednia pochodzi od szefa technologii Epic, autentycznego przybysza z matplanety, który ekstrapolował koszty kompletnego tytułu na podstawie nakładów na demo Samaritan z 2011 roku. A przecież wszyscy Ziemiańskie wiedzą, że po to jedni robią drogie badania, by inni mogli tanio korzystać z ich efektów. Innymi słowy: Epic Games pakuje miliony w opracowanie Unreala 4 właśnie dlatego, by jego licencjodawcy mogli oszczędzić miliony na tworzeniu własnych engine'ów. Skorzystaliśmy jednak z okazji, by zapytać polskich twórców gier, co sądzą o nadchodzących budżetach.



Michał Nowakowski

(CD Projekt RED): Według mnie wzrost kosztów nie jest absolutnym pewnikiem. Nie wydaje mi się, żeby nagle wraz z nową generacją konsol trzeba było powiększać zespoły o rzesze ludzi, kupować masę nowych urządzeń do studia czy wydawać gigantyczne pieniądze na oprogramowanie. Powiększenie budżetów będzie bardziej związane z faktem, że projekty mogą być większe, bardziej epickie. Pewien wzrost kosztów może też leżeć po stronie roboczogodzin poświęconych na „nauczenie się” nowego sprzętu.



Paweł Zawodny

(Techland): Przeżyliśmy do tej pory kilka generacji konsol, a nowe urządzenia były szybsze, miały więcej pamięci i oferowały wyższą rozdzielczość, co w sumie skutkowało coraz lepszą jakością gier, a zwłaszcza grafiki. Szczerze mówiąc, nie zauważyłem, by wzrost tych moż-

tu. Nie sądzę jednak, by miało to sprawić, że gry będą dwa razy droższe. Poza tym każdy zespół będzie się uczył raz i nie będzie ponosił podobnych nakładów przy kolejnych przedsięwzięciach.

Zdjęcie 9. Fragment strony czasopisma „PSX Extreme” (skan wersji drukowanej)



Nowy Hitman z Monteralu

PC X360 PS3 Co rok pror... Hitman. Tore Blystad, jeden z ojców sukcesu Hitmana: Rozgrzeszenia, zdradził, że kolejna część cyklu jest już w produkcji. Po raz pierwszy w historii odpowiada za nią jednak nie IO Interactive, a inne studio - Square Enix Montreal. Blystad nie ukrywa, że inspiracją dla nich stanowiła seria Call of Duty, obrabiana na przemian przez Treyarcha i Infinity Ward. Jego zdaniem zyskuje na tym także sama gra, bo zespoły inspirują się nawzajem, dzięki czemu formuła się nie wypala. Nic, czekamy na Hitmana: Zadośćuczynienie.

www.hitman.com



Arma III i Václav Klaus

BIZ Greckiej tragedii Bohemii ciąg dalszy. Blisko trzy miesiące po aresztowaniu Ivana Buchty i Martina Pezlara przez policję wojskową pod zarzutem szpiegostwa (robili fotki bazy wojskowej na Lemnos) w sprawę włączył się prezydent Czech Václav Klaus. Głowa państwa zwróciła się listownie do prezydenta Grecji Karolosza Papouliasa z prośbą, by zechciał „szczególnie uważnie przyjrzeć się temu niefortunnemu zdarzeniu przy uwzględnieniu znakomych stosunków między narodami, tak by nie rzuciło niepotrzebnie cienia na wzajemne relacje”. Czy pomoże - zobaczymy.

www.arma3.com

Next-gen dużo droższy?

BIZ Temperatura spekulacji wokół nowych konsol rośnie odwrotnie proporcjonalnie do tej za oknem. Tym razem nastroje podsyłał Tim Sweeney z Epic Games, który w trakcie Montreal International Game Summit stwierdził, że koszt produkcji gier na kolejną generację sprzętu może „wzrosnąć trzy-, cztero-, a nawet pięciokrotnie”. Myśl absolutnie przerażająca: budżety sięgające setek milionów dolarów w praktyce oznaczają przecież krach naszej branży. Na szczęście ta ponura przepowiednia pochodzi od szefa technologii Epic, autentycznego przybysza z matplanety, który ekstrapolował koszty kompletnego tytułu na podstawie nakładów na demo Samaritan z 2011 roku. A przecież wszyscy Ziemiańskie wiedzą, że po to jedni robią drogie badania, by inni mogli tanio korzystać z ich efektów. Innymi słowy: Epic Games pakuje miliony w opracowanie Unreala 4 właśnie dlatego, by jego licencjodawcy mogli oszczędzić miliony na tworzeniu własnych engine'ów. Skorzystaliśmy jednak z okazji, by zapytać polskich twórców gier, co sądzą o nadchodzących budżetach.



Michał Nowakowski

(CD Projekt RED): Według mnie wzrost kosztów nie jest absolutnym pewnikiem. Nie wydaje mi się, żeby nagle wraz z nową generacją konsol trzeba było powiększać zespoły o rzesze ludzi, kupować masę nowych urządzeń do studia czy wydawać gigantyczne pieniądze na oprogramowanie. Powiększenie budżetów będzie bardziej związane z faktem, że projekty mogą być większe, bardziej epickie. Pewien wzrost kosztów może też leżeć po stronie roboczogodzin poświęconych na „nauczenie się” nowego sprzętu.

tu. Nie sądzę jednak, by miało to sprawić, że gry będą dwa razy droższe. Poza tym każdy zespół będzie się uczył raz i nie będzie ponosił podobnych nakładów przy kolejnych przedsięwzięciach.



Paweł Zawodny

(Techland): Przeżyliśmy do tej pory kilka generacji konsol, a nowe urządzenia były szybsze, miały więcej pamięci i oferowały wyższą rozdzielczość, co w sumie skutkowało coraz lepszą jakością gier, a zwłaszcza grafiki. Szczerze mówiąc, nie zauważyłem, by wzrost tych moż-

Zdjęcie 10. Fragment strony czasopisma „PSX Extreme” (po użyciu programu ABBYY FineReader 12...)

PSX EXTREME

Nazwa	Rozmiar	Nazwa	Rozmiar	Nazwa	Rozmiar	Nazwa	Rozmiar	Nazwa	Rozmiar
psx_161.docx	17 866 KB	1 kwartał 2011.docx	50 505 KB	1 kwartał 2011.zip	6 346 KB	1 kwartał 2011.docx	50 505 KB	1 kwartał 2011.zip	6 346 KB
psx_162.docx	14 888 KB	2 kwartał 2011.docx	49 756 KB	2 kwartał 2011.zip	6 157 KB	2 kwartał 2011.docx	49 756 KB	2 kwartał 2011.zip	6 157 KB
psx_163.docx	12 292 KB	3 kwartał 2011.docx	40 779 KB	3 kwartał 2011.zip	5 057 KB	3 kwartał 2011.docx	40 779 KB	3 kwartał 2011.zip	5 057 KB
psx_164.docx	17 815 KB	4 kwartał 2011.docx	35 268 KB	4 kwartał 2011.zip	4 773 KB	4 kwartał 2011.docx	35 268 KB	4 kwartał 2011.zip	4 773 KB
psx_165.docx	12 849 KB	11 kwartał 2012.docx	44 514 KB	11 kwartał 2012.zip	6 622 KB	11 kwartał 2012.docx	44 514 KB	11 kwartał 2012.zip	6 622 KB
psx_166.docx	11 241 KB	22 kwartał 2012.docx	52 674 KB	22 kwartał 2012.zip	6 551 KB	22 kwartał 2012.docx	52 674 KB	22 kwartał 2012.zip	6 551 KB
psx_167.docx	12 521 KB	33 kwartał 2012.docx	50 096 KB	33 kwartał 2012.zip	7 010 KB	33 kwartał 2012.docx	50 096 KB	33 kwartał 2012.zip	7 010 KB
psx_168.docx	13 804 KB	44 kwartał 2012.docx	50 449 KB	44 kwartał 2012.zip	6 862 KB	44 kwartał 2012.docx	50 449 KB	44 kwartał 2012.zip	6 862 KB
psx_169.docx	10 321 KB	111 kwartał 2013.docx	34 625 KB	111 kwartał 2013.zip	6 450 KB	111 kwartał 2013.docx	34 625 KB	111 kwartał 2013.zip	6 450 KB
psx_170.docx	12 239 KB	222 kwartał 2013.docx	44 880 KB	222 kwartał 2013.zip	6 067 KB	222 kwartał 2013.docx	44 880 KB	222 kwartał 2013.zip	6 067 KB
psx_171.docx	9 001 KB	333 kwartał 2013.docx	32 326 KB	333 kwartał 2013.zip	5 795 KB	333 kwartał 2013.docx	32 326 KB	333 kwartał 2013.zip	5 795 KB
psx_172.docx	10 045 KB	444 kwartał 2013.docx	42 258 KB	444 kwartał 2013.zip	6 047 KB	444 kwartał 2013.docx	42 258 KB	444 kwartał 2013.zip	6 047 KB
psx_173.docx	16 823 KB	1111 kwartał 2014.docx	31 738 KB	1111 kwartał 2014.zip	6 099 KB	1111 kwartał 2014.docx	31 738 KB	1111 kwartał 2014.zip	6 099 KB
psx_174.docx	12 199 KB	2222 kwartał 2014.docx	37 006 KB	2222 kwartał 2014.zip	6 467 KB	2222 kwartał 2014.docx	37 006 KB	2222 kwartał 2014.zip	6 467 KB
psx_175.docx	9 973 KB	3333 kwartał 2014.docx	33 950 KB	3333 kwartał 2014.zip	5 741 KB	3333 kwartał 2014.docx	33 950 KB	3333 kwartał 2014.zip	5 741 KB
psx_176.docx	18 498 KB	4444 kwartał 2014.docx	38 709 KB	4444 kwartał 2014.zip	5 891 KB	4444 kwartał 2014.docx	38 709 KB	4444 kwartał 2014.zip	5 891 KB
psx_177.docx	12 959 KB	11111 kwartał 2015.docx	32 880 KB	11111 kwartał 2015.zip	6 563 KB	11111 kwartał 2015.docx	32 880 KB	11111 kwartał 2015.zip	6 563 KB
psx_178.docx	15 633 KB	22222 kwartał 2015.docx	34 248 KB	22222 kwartał 2015.zip	6 085 KB	22222 kwartał 2015.docx	34 248 KB	22222 kwartał 2015.zip	6 085 KB
psx_179.docx	15 896 KB	33333 kwartał 2015.docx	46 782 KB	33333 kwartał 2015.zip	5 823 KB	33333 kwartał 2015.docx	46 782 KB	33333 kwartał 2015.zip	5 823 KB
psx_180.docx	13 400 KB	44444 kwartał 2015.docx	39 257 KB	44444 kwartał 2015.zip	5 880 KB	44444 kwartał 2015.docx	39 257 KB	44444 kwartał 2015.zip	5 880 KB
psx_181.docx	14 868 KB								
psx_182.docx	14 612 KB								
psx_183.docx	13 324 KB								
psx_184.docx	16 437 KB								
psx_185.docx	9 131 KB								
psx_186.docx	10 230 KB								
psx_187.docx	10 007 KB								
psx_188.docx	11 856 KB								
psx_189.docx	14 819 KB								
psx_190.docx	13 211 KB								
psx_191.docx	10 864 KB								
psx_192.docx	5 698 KB								
psx_193.docx	11 178 KB								
psx_194.docx	13 780 KB								
psx_195.docx	12 591 KB								
psx_196.docx	10 959 KB								
psx_197.docx	9 131 KB								
psx_198.docx	8 414 KB								
psx_199.docx	9 204 KB								
psx_200.docx	11 225 KB								
psx_201.docx	12 641 KB								
psx_202.docx	7 846 KB								
psx_203.docx	9 278 KB								
psx_204.docx	11 982 KB								
psx_205.docx	10 452 KB								
psx_206.docx	12 686 KB								
psx_208.docx	10 405 KB								
psx_209.docx	8 293 KB								
psx_210.docx	9 124 KB								
psx_211.docx	10 054 KB								
psx_212.docx	10 269 KB								
psx_213.docx	11 632 KB								
psx_214.docx	9 034 KB								
psx_215.docx	16 046 KB								
psx_216.docx	12 322 KB								
psx_217.docx	13 537 KB								
psx_218.docx	10 738 KB								
psx_219.docx	12 377 KB								
psx_220.docx	11 290 KB								

CD-ACTION

Nazwa	Rozmiar	Nazwa	Rozmiar	Nazwa	Rozmiar	Nazwa	Rozmiar	Nazwa	Rozmiar
1 kwartał 2011.docx	82 958 KB	1 kwartał 2011.zip	7 020 KB	1 kwartał 2011.docx	82 958 KB	1 kwartał 2011.docx	82 958 KB	1 kwartał 2011.zip	7 020 KB
2 kwartał 2011.docx	85 030 KB	2 kwartał 2011.zip	8 371 KB	2 kwartał 2011.docx	85 030 KB	2 kwartał 2011.docx	85 030 KB	2 kwartał 2011.zip	8 371 KB
3 kwartał 2011.docx	77 680 KB	3 kwartał 2011.zip	6 929 KB	3 kwartał 2011.docx	77 680 KB	3 kwartał 2011.docx	77 680 KB	3 kwartał 2011.zip	6 929 KB
4 kwartał 2011.docx	71 766 KB	4 kwartał 2011.zip	8 628 KB	4 kwartał 2011.docx	71 766 KB	4 kwartał 2011.docx	71 766 KB	4 kwartał 2011.zip	8 628 KB
11 kwartał 2012.docx	85 648 KB	11 kwartał 2012.zip	8 605 KB	11 kwartał 2012.docx	85 648 KB	11 kwartał 2012.docx	85 648 KB	11 kwartał 2012.zip	8 605 KB
22 kwartał 2012.docx	91 031 KB	22 kwartał 2012.zip	8 893 KB	22 kwartał 2012.docx	91 031 KB	22 kwartał 2012.docx	91 031 KB	22 kwartał 2012.zip	8 893 KB
33 kwartał 2012.docx	93 305 KB	33 kwartał 2012.zip	8 683 KB	33 kwartał 2012.docx	93 305 KB	33 kwartał 2012.docx	93 305 KB	33 kwartał 2012.zip	8 683 KB
44 kwartał 2012.docx	83 441 KB	44 kwartał 2012.zip	8 990 KB	44 kwartał 2012.docx	83 441 KB	44 kwartał 2012.docx	83 441 KB	44 kwartał 2012.zip	8 990 KB
111 kwartał 2013.docx	81 411 KB	111 kwartał 2013.zip	8 543 KB	111 kwartał 2013.docx	81 411 KB	111 kwartał 2013.docx	81 411 KB	111 kwartał 2013.zip	8 543 KB
222 kwartał 2013.docx	91 585 KB	222 kwartał 2013.zip	8 407 KB	222 kwartał 2013.docx	91 585 KB	222 kwartał 2013.docx	91 585 KB	222 kwartał 2013.zip	8 407 KB
333 kwartał 2013.docx	86 260 KB	333 kwartał 2013.zip	8 240 KB	333 kwartał 2013.docx	86 260 KB	333 kwartał 2013.docx	86 260 KB	333 kwartał 2013.zip	8 240 KB
444 kwartał 2013.docx	85 010 KB	444 kwartał 2013.zip	8 531 KB	444 kwartał 2013.docx	85 010 KB	444 kwartał 2013.docx	85 010 KB	444 kwartał 2013.zip	8 531 KB
1111 kwartał 2014.docx	83 156 KB	1111 kwartał 2014.zip	8 324 KB	1111 kwartał 2014.docx	83 156 KB	1111 kwartał 2014.docx	83 156 KB	1111 kwartał 2014.zip	8 324 KB
2222 kwartał 2014.docx	87 433 KB	2222 kwartał 2014.zip	8 547 KB	2222 kwartał 2014.docx	87 433 KB	2222 kwartał 2014.docx	87 433 KB	2222 kwartał 2014.zip	8 547 KB
3333 kwartał 2014.docx	88 232 KB	3333 kwartał 2014.zip	8 844 KB	3333 kwartał 2014.docx	88 232 KB	3333 kwartał 2014.docx	88 232 KB	3333 kwartał 2014.zip	8 844 KB
4444 kwartał 2014.docx	74 756 KB	4444 kwartał 2014.zip	8 521 KB	4444 kwartał 2014.docx	74 756 KB	4444 kwartał 2014.docx	74 756 KB	4444 kwartał 2014.zip	8 521 KB
11111 kwartał 2015.docx	83 080 KB	11111 kwartał 2015.zip	9 129 KB	11111 kwartał 2015.docx	83 080 KB	11111 kwartał 2015.docx	83 080 KB	11111 kwartał 2015.zip	9 129 KB
22222 kwartał 2015.docx	80 441 KB	22222 kwartał 2015.zip	9 701 KB	22222 kwartał 2015.docx	80 441 KB	22222 kwartał 2015.docx	80 441 KB	22222 kwartał 2015.zip	9 701 KB
33333 kwartał 2015.docx	79 919 KB	33333 kwartał 2015.zip	6 587 KB	33333 kwartał 2015.docx	79 919 KB	33333 kwartał 2015.docx	79 919 KB	33333 kwartał 2015.zip	6 587 KB
44444 kwartał 2015.docx	80 217 KB	44444 kwartał 2015.zip	8 599 KB	44444 kwartał 2015.docx	80 217 KB	44444 kwartał 2015.docx	80 217 KB	44444 kwartał 2015.zip	8 599 KB

Tabela 3. Korpus tekstów z czasopism „PSX Extreme” i „CD-Action”

Trzeba zaznaczyć, że główną funkcją programu *ABBYY* jest generowanie tekstu z pozyskiwanych grafik, dlatego szata graficzna czy rodzaj czcionek w końcowych plikach wordowskich, znacznie różni się od tych w drukowanych czasopismach. Dla przykładu dodałem skan fragmentu jednej strony z czasopisma drukowanego (zdj. 9) oraz ten sam fragment po obróbce przez program *ABBYY* (zdj. 10). Porównując oba zdjęcia, zauważyć można, że czasami program błędnie odczytuje/interpretuje pewne znaki. Przykładem tego może być złe odczytanie liczby rzymskiej *III* jako *S* (*ArmA III* -> *ArmAS*). Czasami program może też pomylić podobne do siebie litery np. *q* i *g*. Niemniej jednak nie ma obecnie na rynku programu idealnego, który wyciągnie tekst z grafiki bez ani jednego błędu. Inne tego typu programy zdecydowanie gorzej radzą sobie z rozpoznawaniem polskich tekstów. Szacuję, że program *ABBYY FineReader 12* dobrze interpretuje 95% przypadków wszystkich zeskanowanych znaków.

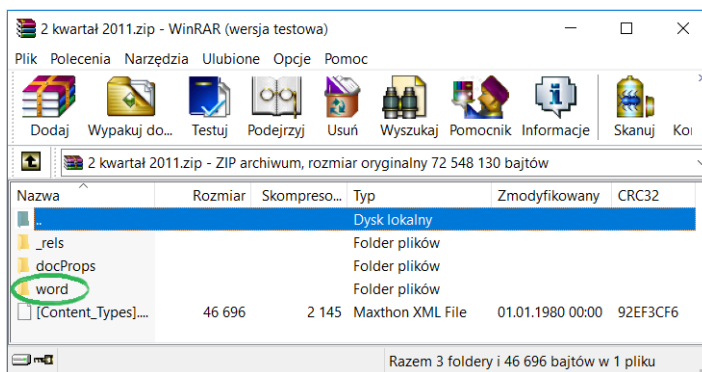
GRUPOWANIE PLIKÓW MIESIĘCZNYCH W KWARTAŁY

W efekcie wszystkich powyżej opisanych czynności, uzyskałem 120 plików *docx*, czyli 60 z czasopisma „*PSX Extreme*” i 60 z „*CD-Action*” (patrz tabela 3, pomarańczowa ramka). Na tym etapie pracy jeden plik odpowiadał jednemu numerowi czasopisma. Aby zwiększyć efektywność pracy postanowiłem pogrupować wszystkie pliki (numery czasopism) w kwartały, a następnie połączyć ich treści (teksty) w większe pliki, z których każdy miał zawierać jeden kwartał danego czasopisma. Łączenie plików wykonałem ręcznie, ponieważ nie było dużo czynności do zrobienia. Otwierałem po trzy pliki np. *psx 161* (plik ze styczniowym numerem „*PSX Extreme*” z 2011 roku), *psx 162* (numer lutowy) i *psx 163* (numer marcowy), a następnie zawartość plików *psx 162* i *psx 163* skopiowałem i wkleiłem do *psx 161*. Powstały w ten sposób plik (łącznie z 3 numerów czasopisma) zapisywałem jako *1 kwartał 2011*. Z pozostałymi plikami czyniłem identycznie, tylko zmieniałem ich nazwy, aby opisywały, który kwartał czasopisma jest w nich zawarty. Ze 120 plików utworzyłem 40 (patrz tabela 3, czerwone ramki). Aby zachować porządek pracy, który pozwolił uniknąć błędów podczas dalszych etapów pracy, wprowadziłem pewien system nazewnictwa. Zdecydowałem, że pliki kwartalne z roku 2011 roku będą nazywały się *1 kwartał 2011*, *2 kwartał 2011*, *3 kwartał 2011* i *4 kwartał 2011*, z kolei nazwy plików zawierających numery czasopisma z następnego roku (2012) zaczynają się od podwojonych cyfr na początku nazwy, czyli *11 kwartał 2012*, następnie *22 kwartał 2012* aż do *44 kwartał 2012*. Zwielokrotnienie cyfr zapobiegło przemieszczeniu się plików, ponieważ gdybym na przykład pierwszy plik z 2012 roku nazwał *1 kwartał 2012*, to

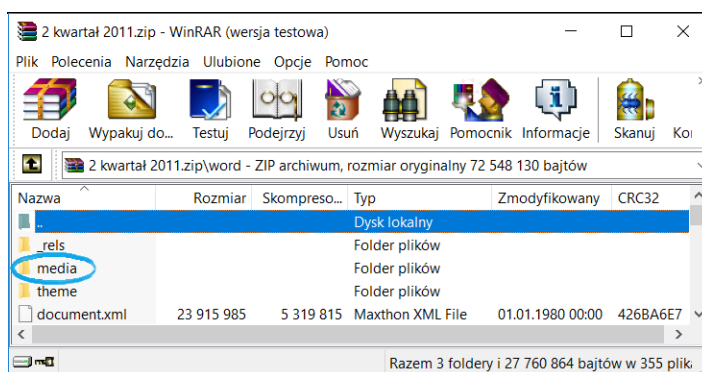
system automatycznie ulokowałby go po pliku *1 kwartał 2011*, bo w nazwach obu plików na początku jest cyfra *1*, a tego chciałem uniknąć.

ZMNIEJSZENIE WAGI/ROZMIARU PLIKÓW

Kolejną przeszkodą do rozwiązania była zbyt duża waga plików⁵⁴, zawierających numery z danego kwartału. Pliki pozyskane z czasopisma „PSX Extreme” ważyły od około 31 do 52 MB, zaś w przypadku „CD-Action” waga wahała się od 71 do 93 MB (tabela 3, czarne ramki). Zważywszy na to, że były to pliki tekstowe, to ważyły bardzo dużo. Moim kolejnym celem było zmniejszenie ich objętości po to, aby komputer mógł szybciej odczytywać dane zawarte w plikach, ponieważ z dużymi plikami pracował bardzo wolno.

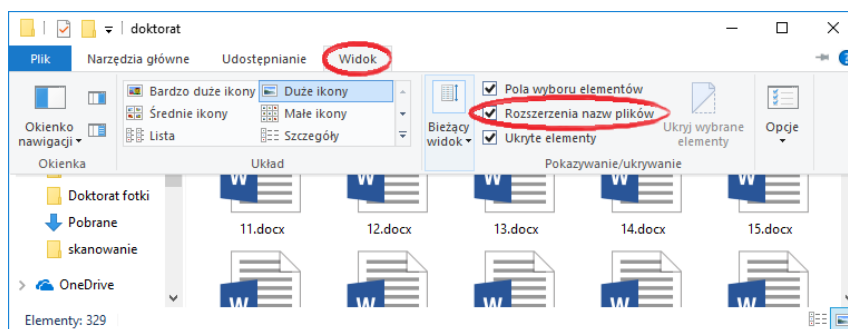


Zdjęcie 12. Struktura pliku wordowskiego



Zdjęcie 13. Struktura folderu word w pliku wordowskim

Rozwiązanie tego problemu znalazłem w likwidacji grafik z plików wordowskich. Jednak obrazów było w nich tak dużo, że pojedyncze ich usuwanie



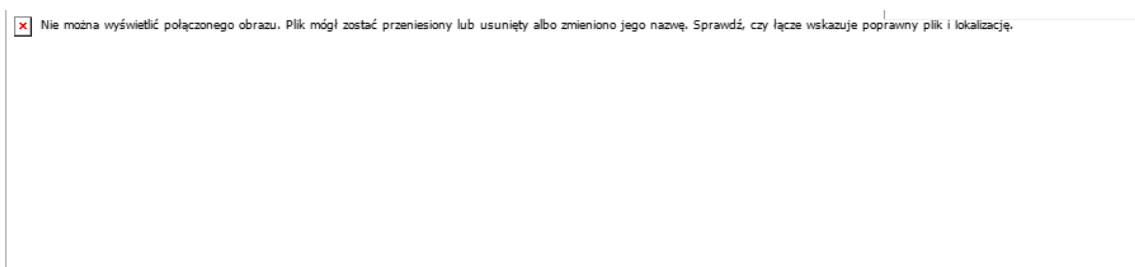
Zdjęcie 11. Rozszerzenia plików

nie było możliwe do wykonania w krótkim czasie, a nie chciałem go zbyt dużo przeznaczać na tę czynność. W hurtowym usunięciu grafik pomocna okazała się wiedza na temat budowy/struktury plików *docx*. Nie są one jednolitymi plikami, ale pakietami danych, a wszystkie pakiety można rozłożyć na mniejsze pliczki. Włączyłem opcję pokazywania

⁵⁴ Używam tutaj określenia z socjolektu informatycznego, które moim zdaniem lepiej oddaje sens opisywanego problemu. *Wagę* w języku oficjalnym należałoby zastąpić *wielkością* lub *objętością*.

rozszerzeń plików, która w większości systemów jest domyślnie ukryta. W systemie operacyjnym, na którym pracowałem (Windows 10) otworzyłem dowolne okno z plikami/folderami, a następnie z belki na górze wybrałem menu *widok* i zazaczyłem opcję *rozszerzenia nazw plików* (patrz zdj. 11, czerwone elipsy). Od tego momentu po nazwie plików wyświetlały się także ich rozszerzenia (np. *docx*).

Następnie zmieniłem rozszerzenia wszystkich kwartalnych plików wordowskich z *docx* na *zip*. Uczyniłem to poprzez edycję ich nazw, np. nazwę pliku *1 kwartał 2011.docx* zamieniłem na *1 kwartał 2011.zip*. Rezultaty tej czynności zaprezentowałem w tabeli 2 (zielone ramki). Później przy pomocy programu WinRAR⁵⁵ otworzyłem każdy z plików. Strukturę takiego pliku pokazałem na zdjęciu 12. Odszukałem w nim folder *word* (zdj. 12, zielona elipsa), otworzyłem go i skasowałem katalog *media* (zdj. 13, niebieska elipsa), w którym umieszczone były wszystkie grafiki. Po usunięciu ze wszystkich plików tego folderu, diametralnie zmniejszyła się ich waga (patrz tabela 3, niebieskie ramki). Ostatnim etapem prac z tymi plikami było ponowne zmodyfikowanie ich rozszerzeń, czyli tym razem zamieniłem rozszerzenia z *zip* na *docx* (tabela 3, żółte ramki).



Nowy Hitman z Monteralu

PC X360 PS3 Co rok ~~pror...~~ ~~Hitman~~, ~~Tore Blystad~~, jeden z ojców sukcesu ~~Hitmana~~. Rozgrzeszenia, zdradził, że kolejna część cyklu jest już w produkcji. Po raz pierwszy w historii odpowiada za nią jednak nie IO Interactive, a inne studio - ~~Square Enix~~ Montreal. ~~Blystad~~ nie ukrywa, że inspiracją dla nich stanowiła seria *Call of Duty*, obrabiana na przemian przez ~~Treyarcha~~ i ~~Infinity Ward~~. Jego zdaniem zyskuje na tym także sama gra, bo zespoły inspirowują się nawzajem, dzięki czemu formuła się nie wypala. Nic, czekamy na ~~Hitmana~~. Zadośćuczynienie.

www.hitman.com

ArmAS i Vaclav Klaus

BIZ Greckiej tragedii Bohemii ciąg dalszy. Blisko trzy miesiące po aresztowaniu Ivana Buchty i Martina ~~Pezara~~ przez policję wojskową pod zarzutem szpiegostwa (robili fotki bazy wojskowej na Lemnos) w sprawę włączył się prezydent Czech Vaclav Klaus. Głowa państwa zwróciła się listownie do prezydenta Grecji ~~Karolos~~ ~~Papou~~ ~~ujliasa~~ z prośbą, by ~~zechciał~~ ~~szczególnie~~ uważnie przyjrzeć się temu niefortunnnemu zdarzeniu przy uwzględnieniu znakomych stosunków między narodami, tak by nie rzuciło niepotrzebnie cienia na wzajemne relacje". Czy pomoże - zobaczymy.

www.arma3.com

Next-gen dużo droższy?

BIZ Temperatura spekulacji wokół nowych konsol rośnie odwrotnie proporcjonalnie do tej za oknem. Tym razem nastroje podsylił Tim ~~Swanney~~ z ~~Epix~~ Games, który w trakcie Montreal International Game Summit stwierdził, że koszt produkcji gier na kolejną generację sprzętu ~~może wzrosnąć~~ trzy-, cztero-, a nawet pięciokrotnie". Myśl absolutnie przerażająca: budżety sięgające setek milionów dolarów w praktyce oznaczają przecież krach naszej branży. Na szczęście ta ponura przepowiednia pochodzi od szefa technologii ~~Epixa~~, autentycznego przybysza z ~~matplanety~~, który ekstrapolował koszty kompletnego tytułu na podstawie nakładów na demo ~~Samaritan~~ z 2011 roku. A przecież wszyscy Ziemianie wiedzą, że po to jedni robią drogie badania, by inni mogli tanio korzystać z ich efektów. Innymi słowy: ~~Epix~~ Games pakuje miliony w opracowanie ~~Unreal~~ 4 właśnie dlatego, by jego licencjodawcy mogli oszczędzić miliony na tworzeniu własnych ~~engine'ow~~. Skorzystaliśmy jednak z okazji, by zapytać polskich twórców gier, co sądzą o nadchodzących budżetach.

Michał Nowakowski (CD Projekt RED): Według mnie wzrost kosztów nie jest absolutnym pewnikiem. Nie wydaje mi się, żeby nagle wraz z nową generacją konsol trzeba było powiększać zespoły o rzesze ludzi, kupować masę nowych urządzeń do studia czy wydawać gigantyczne pieniądze na oprogramowanie. Powiększenie budżetów będzie bardziej związane z faktem, że projekty mogą być większe, bardziej epickie. Pewien wzrost kosztów może też leżeć po stronie roboczo godzin poświęconych ~~na~~ ~~nauczenie~~ się nowego ~~sortu~~.

tu. Nie sądzę jednak, by miało to sprawić, że gry będą dwa razy droższe. Poza tym każdy zespół będzie się uczył raz i nie będzie ponosił podobnych nakładów przy kolejnych przedsięwzięciach.

Paweł Zawodny (Techland): Przeżyliśmy do tej pory kilka generacji konsol, a nowe urządzenia były szybsze, miały więcej pamięci i oferowały wyższą rozdzielczość, co w sumie skutkowało coraz lepszą jakością gier, a zwłaszcza grafiki. Szczepnie mówiąc, nie zauważyłem, by wzrost tych ~~roz-~~

Zdjęcie 14. Fragment strony czasopisma PSX Extreme (po usunięciu grafik)

⁵⁵ Program ten jest na licencji shareware, czyli rozpowszechniany jest bezpłatnie, ale żeby w pełni wykorzystywać jego możliwości lub pominąć ograniczenia trzeba za niego zapłacić. W przypadku programu WinRAR ograniczeniem jest okienko wyświetlane po uruchomieniu program, że jest to wersja testowa.

Rezultatem skasowania grafik było to, że w miejscach ich pierwotnego osadzenia, wyświetlane były teraz puste ramki (patrz zdj. 14). Dzięki ich skasowaniu praca z plikami mogła przebiegać dużo efektywniej.

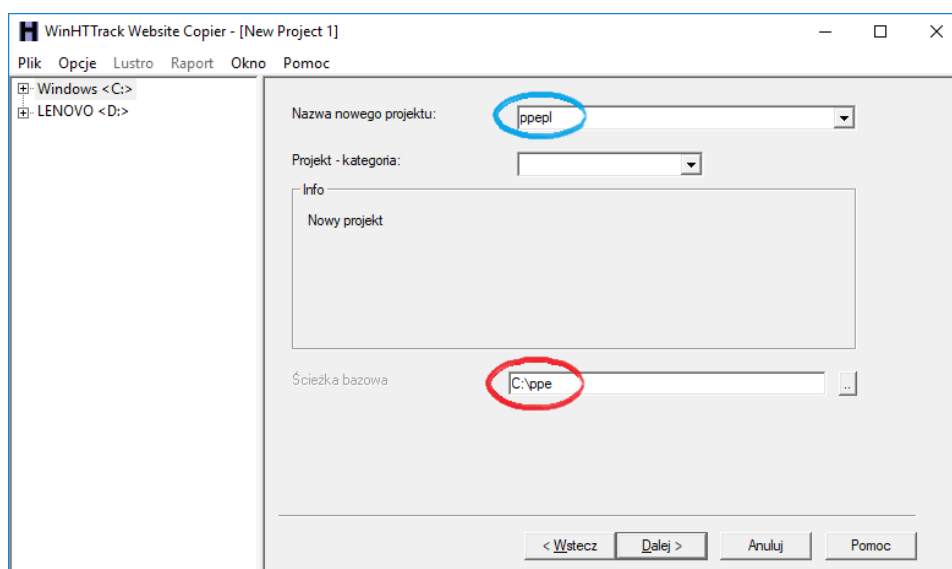
Tak przygotowany materiał jest pierwszym elementem budującym korpus tekstów do moich badań. Kolejnym elementem są teksty wyciągnięte ze stron internetowych. Proces ich pozyskiwania opisałem poniżej.

ETAP 2. POZYSKIWANIE EDYTOWALNYCH MATERIAŁÓW TEKSTOWYCH Z INTERNETU

Aby móc pracować z tekstami opublikowanymi w Internecie miałem do wyboru dwie drogi postępowania: albo każdy news pobierać ręcznie, albo znaleźć sposób na automatyczne ich pobranie. Wybrałem drugą opcję i skorzystałem z programu, który pobiera strony internetowe w całości, bądź wybrane ich elementy. Program ten nazywa się *HTTrack Website Copier*⁵⁶ i jest on zaawansowanym narzędziem posiadającym wiele możliwości konfiguracji.

PRACA Z PROGRAMEM HTTRACK WEBSITE COPIER

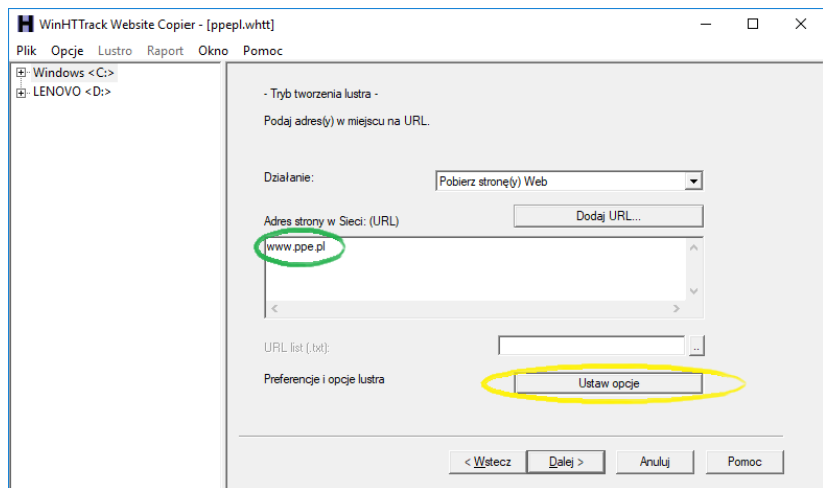
Po uruchomieniu programu, a następnie kliknięciu przycisku *dalej*, wpisałem nazwę projektu – *ppepl* (zdj. 15, niebieska elipsa) i podałem ścieżkę bazową, czyli wskazałem miejsce, do którego ma zostać pobrana strona internetowa (zdj. 15, czerwona elipsa). Kliknąłem *dalej* i



Zdjęcie 15. Początkowe okno programu *HTTrack Website Copier*

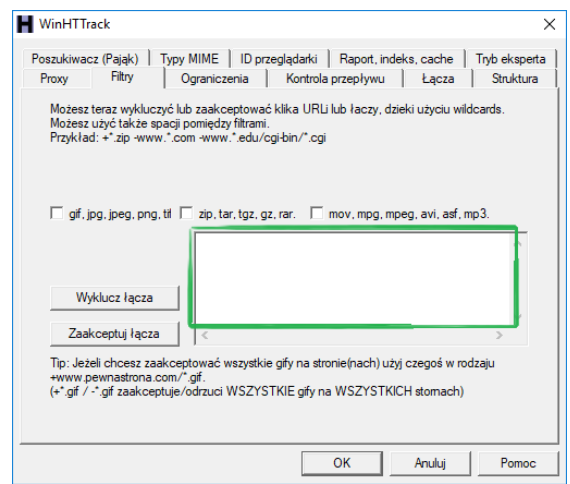
⁵⁶ Program *HTTrack Website Copier* dostępny jest na licencji shareware.

w kolejnym oknie wpisałem adres strony www, którą chciałem pobrać (zdj. 16, zielona elipsa), później wszedłem w menu *ustaw opcje* (zdj. 16, żółta elipsa), ponieważ program, aby dobrze pracował, musiał zostać odpowiednio skonfigurowany.

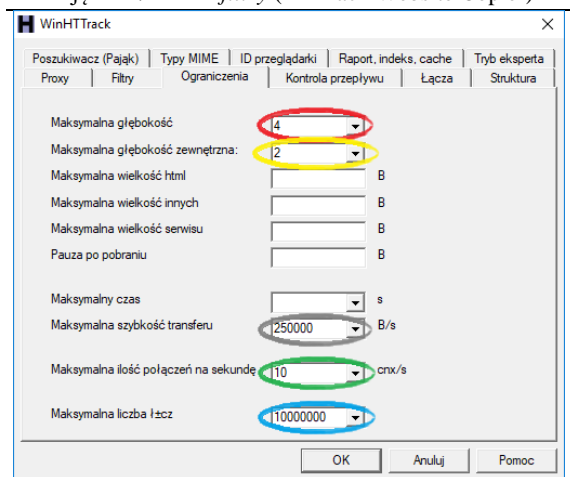


Zdjęcie 16. Główne okno programu HTTrack Website Copier

Konfiguracja, którą tutaj zaprezentowałem, sprawdziła się w przypadku mojego komputera, jednak gdybym pracował na innym, to przed przystąpieniem do pobrania dużej strony, sprawdziłbym, metodą prób i błędów, inne konfiguracje, aby osiągnąć jak najlepsze efekty pracy programu. użytą przeze mnie konfigurację prezentuję na zdjęciach 17 i 18. W zakładce *filtry* usunąłem z ramki domyślnie ustawiony zapis „+*.png +*.gif +*.jpg +*.jpeg +*.css +*.js -ad.doubleclick.net/* mime:application/foobar” (patrz zdj. 17, zielona ramka), ponieważ nie chciałem, żeby pobierane były także grafiki (*png, gif, jpg czy jpeg*), pliki ze stylami (*css*) czy skrypty java (*js*). Chciałem, żeby program pobrał tylko pliki *html*, w których zawarta jest główna treść strony, w tym między innymi teksty *newsów* i wpisy na forach internetowych, które następnie miałem wykorzystać do stworzenia korpusu.



Zdjęcie 17. Menu *filtry* (HTTrack Website Copier)



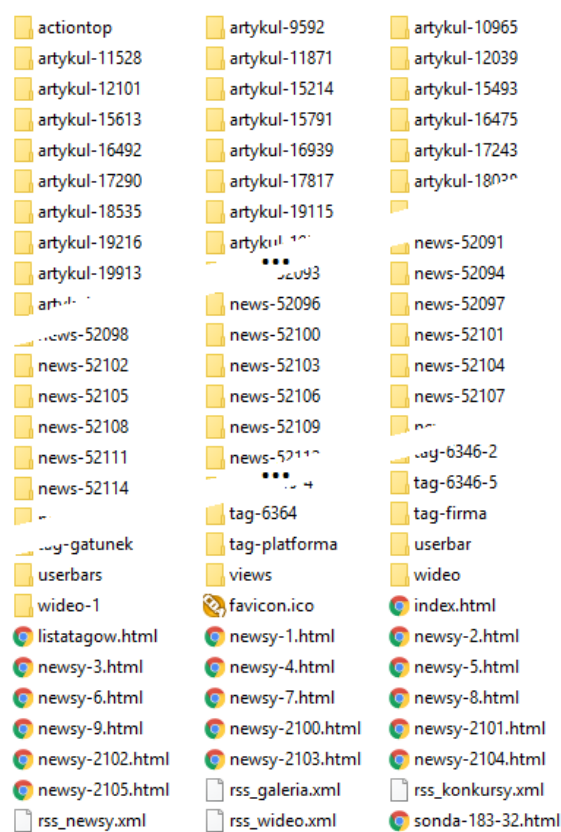
Zdjęcie 18. Menu *ograniczenia* (HTTrack Website Copier)

W zakładce *ograniczenia* ustawiłem *maksymalną głębokość* na 4 (zdj. 18, czerwona elipsa). Ustawiając ten współczynnik wybrałem, że poza główną stroną, program będzie pobierał także podstrony, do których odnośniki znajdują się na pierwszej stronie. Będą one pobierane aż do 3 poziomu podstron, ponieważ za pierwszy poziom uważana jest strona

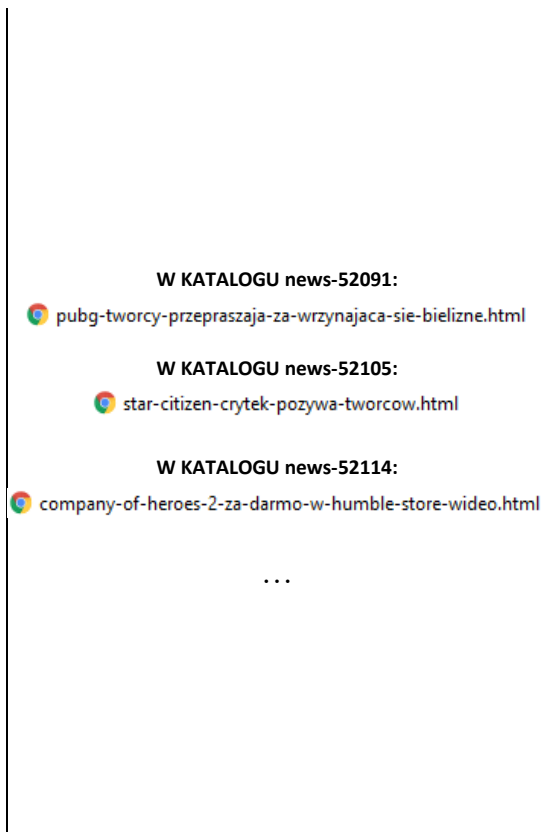
główna. *Maksymalną głębokość zewnętrzną* ustawiłem na 2 (zdj. 18, żółta elipsa). Oznacza to, że jeżeli na stronie głównej pojawi się jakiś odnośnik do innej strony, to program także ją pobierze. Jednak nie pobierze jej podstron, a wyłącznie stronę główną, natomiast gdybym ustawił 3, to pobrane zostałyby także jej podstrony. Ustawiłem także wartości *maksymalnej szybkości transferu* (zdj. 18, szara elipsa), *ilości połączeń na sekundę* (zdj. 18, zielona elipsa) oraz *liczby łącz* (zdj. 18, niebieska elipsa), ponieważ chciałem, aby program pobrał stronę możliwie szybko. Po skonfigurowaniu wszystkiego kliknąłem *OK*, następnie *dalej i zakończ*, tym samym rozpoczynając pobieranie strony.

Po kilku dniach, gdy program zakończył pracę, uruchomiłem go ponownie, ponieważ podczas pierwszego pobierania strony niektóre pliki zostają pominięte, dlatego aby zminimalizować ich liczbę warto program uruchomić ponownie. Po zakończeniu pracy programu i pobraniu w taki sam sposób strony *cdaction.pl*, uzyskałem kopie stron internetowych zapisane na dysku komputera, w identycznej postaci, jaka jest na serwerach.

		<p>W KATALOGU 371: dice-nie-chce-kasy-za-mapy.html</p> <p>W KATALOGU 455: ekranizacja-gears-of-war-zagrozona.html</p> <p>W KATALOGU 548: devil-may-cry-5.html</p> <p>...</p>
<p>...</p> <p>Zdjęcie 19. Struktura strony www.ppe.pl</p>	<p>...</p> <p>Zdjęcie 20. Foldery z katalogu <i>newsy</i> (ppe.pl)</p>	<p>Zdjęcie 21. Newsy ze strony www.ppe.pl</p>



Zdjęcie 22. Struktura strony www.cdaction.pl



Zdjęcie 23. Newsy ze strony www.cdaction.pl

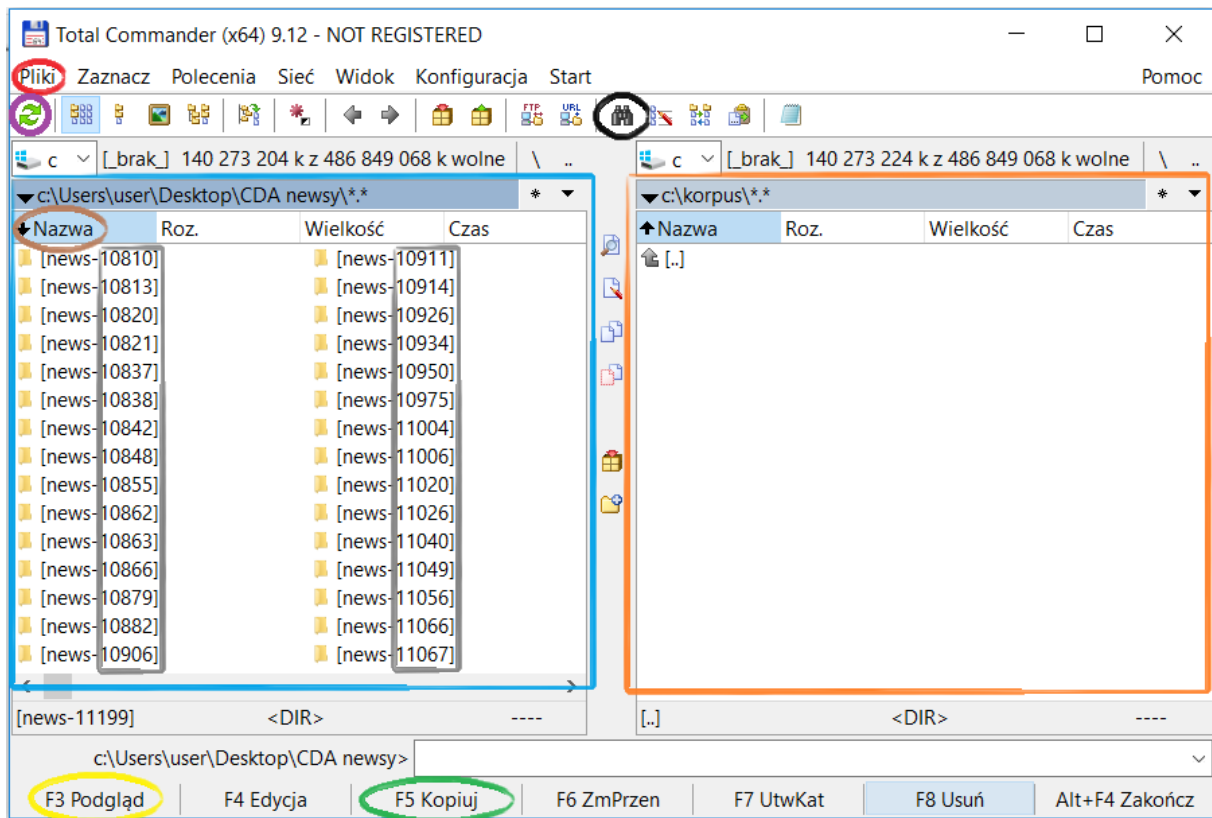
Musiałem poznać strukturę stron, czyli zobaczyć, otwierając poszczególne foldery, gdzie ulokowane zostały pliki z tekstami *newsów*. Nie było to trudne zadanie, ponieważ te strony mają bardzo przejrzystą strukturę (patrz zdj. od 19 do 23).

W przypadku *ppe.pl*, *newsy* zostały umieszczone w folderze o identycznej nazwie – *newsy* (zdj. 19), a dodatkowo każdy z plików znajdował się w osobnym folderze, którego nazwa była kolejną liczbą porządkową (zdj. 20). Z kolei *cdaction.pl* posiada mniej uporządkowaną strukturę, a każdy z *newsów* został umieszczony tylko w osobnym folderze, którego nazwa składała się z członu *news* oraz kolejnej liczby porządkowej (zdj. 22).

PORZĄDKOWANIE ZGROMADZONEGO MATERIAŁU

Mimo iż struktura stron była klarowna, to idealną sytuacją dla mnie do dalszej pracy byłoby, gdyby wszystkie pliki nie były w osobnych folderach, ale w jednym i nie posiadałyby nazw identycznych, jak tytuły *newsów*, które są w nich zawarte (zdj. 21 i 23), tylko były ponumerowane od najmniejszej (najstarszy *news*) do największej (*najnowszy*) liczby porządkowej. Musiałem sobie taką sytuację stworzyć. Reasumując, problem stanowiło to, że *newsy* były umieszczone w osobnych folderach, które co prawda były ponumerowane, ale

właściwe pliki stron (te z rozszerzeniem *html*) już takiej numeracji nie posiadały. Gdybym teraz wyciągnął (skopiował lub wyciął) wszystkie pliki z folderów i umieścił w jednym, to system zapewne poustawiałby je alfabetycznie, czyli nie zachowałyby się kolejność od najstarszego *news*a do najnowszego. Z kolei ręczna numeracja kilkunastu tysięcy plików stwarzała dużą szansę pomylenia się w ich numerowaniu. Rozwiązaniem problemu była hurtowa zmiana nazw plików, a taką umożliwił program *Total Commander 9.0a*⁵⁷.



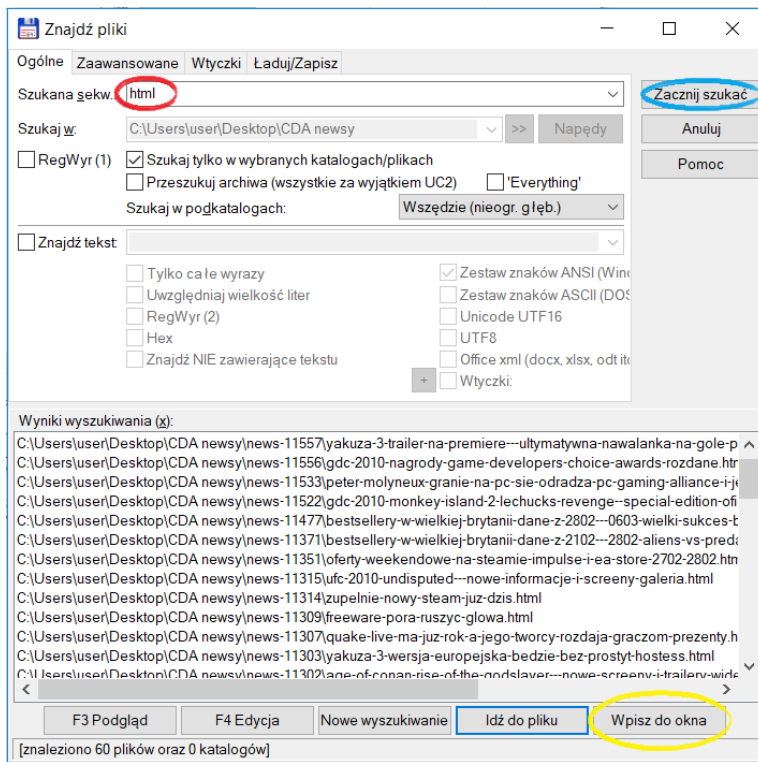
Zdjęcie 24. Główne okno programu Total Commander 9.12

Po uruchomieniu programu, w oknie po lewej stronie (zdz. 24, niebieska ramka) wskazałem mu, gdzie na moim dysku twardym ulokowany był folder z *newsami* (c:\Users\user\Desktop\CDA newsy\), natomiast w prawym oknie (zdz. 24, pomarańczowa ramka) wybrałem lokalizację pustego folderu, do którego, pod koniec pracy z programem, skopiowane miały zostać finalne wersje plików.

Zazaczyłem wszystkie foldery w oknie po lewej stronie, czyli kliknąłem na pierwszy z nich, a następnie trzymając klawisz *shift*, kliknąłem na ostatni. Wszystkie nazwy katalogów z koloru czarnego zmieniły się na czerwony, potwierdziło to zaznaczenie wszystkich folderów. Ważne, aby foldery były ułożone począwszy od *news*a z najmniejszą liczbą porządkową na końcu, a skończywszy na tym z największą liczbą (zdz. 24, szare ramki). Początkowo były one

⁵⁷ Program *Total Commander 9.12* dostępny jest na licencji shareware.

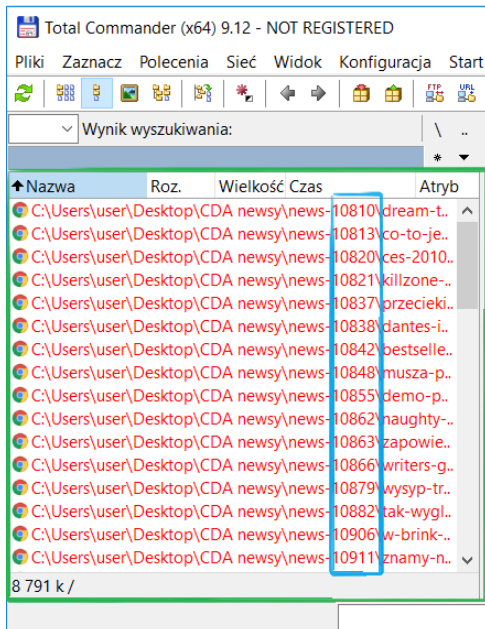
pomieszczone, ale po kliknięciu na przycisk *nazwa* (zdj. 24, brązowa elipsa), zostały ułożone. Gdy już pliki zostały poukładane i zaznaczone na czerwono, wybrałem polecenie *szukaj* (zdj. 24, czarna elipsa). W okienku, które zostało otworzone (zdj. 25), w polu wyszukiwania



Zdjęcie 25. Wyszukiwarka programu Total Commander 9.12

wpisałem *html* (zdj. 25, czerwona elipsa). Był to element wspólny dla wszystkich plików z *newsami*, a tylko te chciałem wyszukać. Kliknąłem *zaczynj szukać* (zdj. 25, niebieska elipsa), a po wyszukaniu plików nacisnąłem *wpisz do okna* (zdj. 25, żółta elipsa).

Wszystkie pliki *html* umieszczone zostały w głównym oknie programu po jego lewej stronie (zdj. 26, zielona ramka). Następnie zazaczyłem je i sprawdziłem ponownie, czy są



Zdjęcie 26. Pliki html w Total Commanderze 9.12

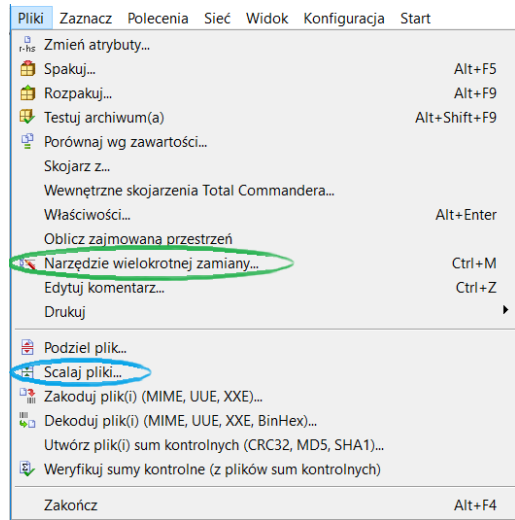
poukładane od najmniejszej do największej wartości liczby porządkowej katalogu (patrz zdj. 26, niebieska ramka). Z menu *plik* (zdj. 24, czerwona elipsa) wybrałem *narzędzie wielokrotnej zamiany...* (zdj. 27, zielona elipsa).

W nowo otwartym oknie (zdj. 28) ponownie upewniłem się, czy pliki są odpowiednio ułożone (zdj. 28, brązowa ramka), a następnie ustawiłem, poprzez naciśnięcie przycisku *licznik* (zdj. 28, czerwona elipsa), żeby pliki z *newsami* (*html*) nosiły nazwy, które są kolejnymi liczbami porządkowymi. Sprawdziłem czy w okienku *matryca zmiany: nazwa pliku* znajduje się

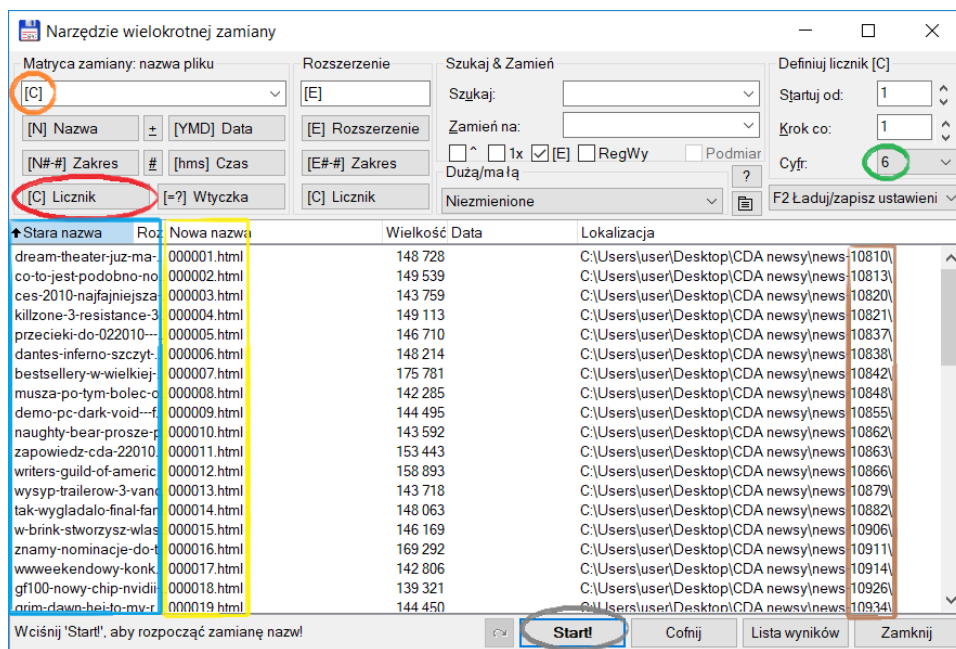
tylko litera C (patrz zdj. 28, pomarańczowa elipsa), a później dla bezpieczeństwa, aby pliki nie zostały poprzestawiane przez system operacyjny wybrałem, żeby nowa nazwa plików składała się z 6 cyfr (zdj. 28, zielona elipsa). Dzięki temu nazwy plików będą się zaczynały nie od 1, ale od 000001. Dużym ułatwieniem dla badacza jest to, że program pokazuje jednocześnie obecne

nazwy plików (zdz. 28, niebieska ramka) oraz te, które będą po zmianie (zdz. 28, żółta ramka). Po sprawdzeniu ich, kliknąłem przycisk *start* (zdz. 28, szara elipsa), a po zakończeniu procesu zmiany nazw plików, nacisnąłem przycisk *odśwież* (zdz. 24, purpurowa elipsa).

Program automatycznie wrócił do folderu, w którym znajdowały się katalogi z plikami *html* (*newsami*), czyli do stanu, który prezentuje zdjęcie 24, z tą różnicą, że teraz w folderach były wyłącznie pliki z nowymi nazwami. Ponownie zazaczyłem wszystkie foldery i wybrałem opcję *szukać*. W okienku wyszukiwania jeszcze raz wpisałem frazę *html*, a znalezione pliki, poprzez naciśnięcie przycisku *wpisz do okna*, umieszczone zostały w głównym oknie programu. Wszystkie je zazaczyłem i kliknąłem przycisk *kopiuj* (zdz. 24, zielona elipsa), po czym pliki z lewego okna programu (zdz. 24, niebieska ramka) zostały skopiowane do prawego (zdz. 24, pomarańczowa ramka). Chwilowo zakończyłem pracę z programem *Total Commander*.



Zdjęcie 27. Menu pliki (Total Commander 9.12)



Zdjęcie 28. Narzędzie wielokrotnej zmiany programu Total Commander 9.12

Z poziomu systemu operacyjnego wszedłem do katalogu, w którym umieszczone zostały pliki po procesie zmiany nazw. Otwierając poszczególne pliki przy pomocy

przeglądarki internetowej (korzystałem z *Google Chrome*), szukałem *newsów*, które były pierwszymi rozpoczynającymi dany kwartał. Szukałem więc np. pliku z 1 stycznia 2011 roku, 1 kwietnia 2014 roku czy 1 października 2015 roku. Gdy znalazłem już takie pliki, to zmodyfikowałem ich nazwy w taki sposób, że po myślniku dopisywałem, który kwartał dany plik rozpoczyna, np. pierwszy *news* z 2011 roku otrzymał nazwę „xxxxxx⁵⁸ – 2011 1 kwartał”, a rozpoczynający ostatni kwartał 2015 roku – „xxxxxx – 2015 4 kwartał”. Po znalezieniu wszystkich plików, które rozpoczynały dane kwartały, pogrupowałem je w katalogi. W jednym folderze znajdowały się *newsy* wyłącznie z jednego kwartału danego roku. Aby zachować proporcje badanych tekstów zdecydowałem, że z każdego kwartału wylosuję (przy pomocy generatorów liczb losowych dostępnych w Internecie) 300 *newsów*. Podsumowując, łącznie do badań włączyłem 6.000 *newsów* ze strony *ppe.pl* i tyle samo z *cdaction.pl*. Przez losowanie plików i usuwanie tych, które do badań nie zostały włączone, jednolity ciąg nazewniczy przestał istnieć. Nie posiadałem już ponumerowanych np. od 000001 do 011100 plików, tylko wybiórcze numery z tego przedziału. Musiałem to z powrotem uporządkować, ponieważ nie mógłbym wykonać kolejnych czynności zmierzających do stworzenia korpusu tekstów.

Dzięki wykorzystaniu powyżej przedstawionej metody zmiany nazw plików, uzyskałem w każdym folderze zawierającym jeden kwartał *newsów*, pliki ponumerowane od 001 do 300.

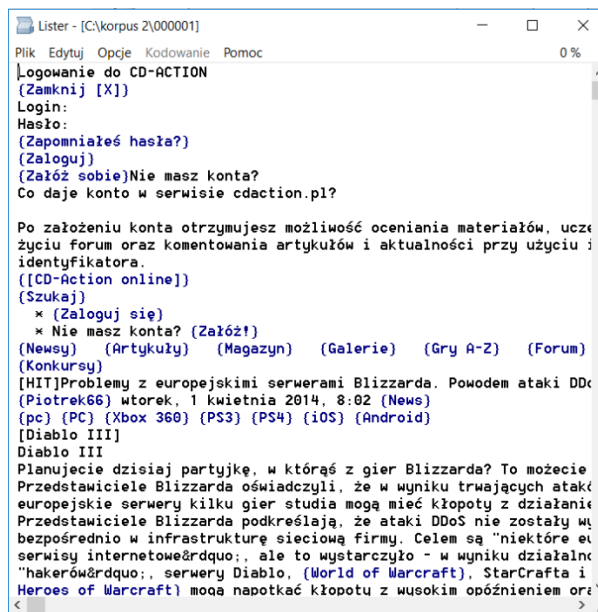
ŁĄCZENIE MAŁYCH PLIKÓW W WIĘKSZE – TWORZENIE WERSJI FINALNYCH

Ponownie uruchomiłem program *Total Commander* i w oknie po lewej stronie wskazałem ścieżkę dostępu do folderu z plikami podzielonymi na kwartały. Następnie wszedłem do pierwszego katalogu (1 kwartał 2011) i zazaczyłem pierwszy plik (001.html). Z menu *plik* (zdj. 24, czerwona elipsa) wybrałem *scalaj pliki...* (zdj. 27, niebieska elipsa). Mogłem je scalić, ponieważ posiadałem pliki, których nazwy były ciągiem następujących po sobie liczb, gdyby one nazywały się inaczej lub ciągi liczb były poprzerywane, do scalenia by nie doszło.

Po zakończeniu procesu zamiast 300 plików z *newsami* uzyskałem jeden duży zawierający je wszystkie. Później proces scalania powtórzyłem z resztą plików w pozostałych folderach kwartalnych. W rezultacie uzyskałem 40 dużych plików, a każdy z nich zawierał *newsy* z jednego kwartału.

⁵⁸ xxxxxx symbolizuje dowolną liczbę porządkową.

Ostatnim etapem pracy z *newsami* było wybranie dużego, powstałego w wyniku scalenia, pliku i kliknięcie przycisku *podgląd* (zdj. 24, żółta elipsa). Plik otworzył się w nowym oknie – *listerze* (zdj. 29), skąd mogłem skopiować tekst w nim wyświetlany i wkleić do nowo utworzonego pliku Word. Identycznie postąpiłem w przypadku pozostałych 39 dużych plików. Dzięki wszystkim powyżej opisanym czynnościom uzyskałem 40 plików *docx*, a każdy z nich zawierał treści 300 newsów z 1 kwartału.



Zdjęcie 29. Lister programu Total Commander 9.12

Podobnie postępowałem z plikami pozyskanymi z forów internetowych, tylko z tą różnicą, że nie dzieliłem plików na kwartały, ponieważ nie było to możliwe. Forum jest inaczej skonstruowane niż strona internetowa, nie ma tam klarownego podziału na wpisy z uwzględnieniem daty ich powstania. Na forum każdy wątek może być edytowany w dowolnym momencie, dlatego nie da się wyróżnić wpisów przykładowo z 1, 2 czy 3 kwartału 2013 roku. Pliki pozyskane z forum scalałem, wyciągałem z nich tekst i skopiowałem do plików wordowskich.

ETAP 3. PRZYGOTOWANIE MATERIAŁU LEKSYKALNEGO DO PRACY Z KORPUSAMI TEKSTOWYMI

Po stworzeniu korpusu tekstów, ale przed przystąpieniem do badań ilościowych, ponownie zapoznałem się ze zgromadzonym materiałem. Przeczytałem artykuły w czasopiśmie, które wchodzi w skład korpusu, zapoznałem się z zebranymi newsami internetowymi oraz wpisami na forach internetowych. Zrobiłem to w celu wynotowania jednostek leksykalnych charakterystycznych dla środowiska graczy wideo. Ale czym kierowałem się decydując, czy dana jednostka jest charakterystyczna dla tego środowiska, czy też nie jest?

W rozstrzygnięciu tego problemu kierowałem się intuicją badawczą i jednocześnie intuicją gracza. Zakładam, że żeby badać dany socjolekt trzeba być częścią społeczności, która go używa.

Podczas ponownego zapoznawania się z materiałem wynotowałem łącznie 1.732 jednostki, które następnie podzieliłem na 4 grupy tematyczne: *branża gier wideo* (440 jednostek), *środowisko graczy* (386 jednostek), *sprzęt graczy* (546 jednostek) i *gra* (360 jednostek).

Podział ten powstał w wyniku obserwacji środowiska, którego w jakimś stopniu jestem częścią oraz dzięki ogólnej znajomości leksyki występującej w tekstach tej grupy społecznej.



Grafika 6. Podział środowiska graczy

Jak widać na załączonej powyżej grafice (6), w centrum podziału znajduje się gracz, jest on jednocześnie punktem odniesienia do wytyczenia kręgów tematycznych. Kręgiem najbardziej oddalonym od gracza jest *branża gier wideo*, natomiast najbliższym *gra*. Po między są jeszcze *środowisko graczy* i *sprzęt gracza*. Podziału na kręgi dokonałem także z uwzględnieniem pewnych kryteriów, które pozwoliły na umieszczenie danej jednostki w konkretnym zbiorze.

W kręgu tematycznym – *branża gier wideo*, umieściłem podstawowe formy wyrazowe, które odnoszą się do rodzaju umieszczonych w grze płatności (np. *mikrotransakcje*⁵⁹), programów/aplikacji tworzonych przez firmy deweloperskie dla graczy (np. *patch*⁶⁰), nazwy

⁵⁹ Gry oferowane darmo często posiadają w sobie system *mikrotransakcji*. Oznacza to, że aby zdobyć jakieś ułatwienie w grze, to trzeba zapłacić realnymi pieniędzmi.

⁶⁰ Plik z programem, który po zainstalowaniu poprawia błędy w wydanej już grze/programie.

rodzajów gier, ale tylko tych, które są zależne od firmy (np. *remaster*⁶¹) oraz elementów zaimplementowanych w grach, które poprzedzają właściwą rozgrywkę (np. *multi*⁶²). W kręgu tym są więc te formy wyrazowe, które odnoszą się do czegoś, na co wpływ miały/mają firmy produkujące gry.

Drugi krąg, bliższy graczowi, stanowi środowisko graczy. Umieściłem w nim wszelkie nazwy graczy (np. *niedzielnik* ‘osoba grająca rzadko’), wyrazy służące ocenie danych produkcji (np. *crap* ‘gra słaba’), nazwy czynność podejmowanych przez graczy w grze, ale także poza nią – w środowisku graczy (np. *farmic*⁶³), a także nazwy klasyfikujące gry do danego gatunku, które nie są odgórnie zaplanowane przez firmy, ale są inicjatywą oddolną – graczy (np. *przygodówka* ‘gra przygodowa’). Z analizy czasopism wynika, że klasyfikacje te są bardzo płynne i gracze potrafią różnie klasyfikować dane gry. Brak konsekwencji widać, gdy porówna się recenzje tych samych gier w różnych czasopismach.

Trzeci krąg to sprzęt gracza. Tutaj umieściłem wszystkie formy związane z elementami materialnymi, które niezbędne są graczom do grania (np. komputer). Przy czym trzeba podkreślić, że musi to być sprzęt ogólnodostępny, a nie przeznaczony np. dla zamkniętego grona developerów produkujących gry.

Ostatni krąg stanowi natomiast gra. Umieściłem w nim wszystkie formy podstawowe, które związane są ze światem gry, z tym jednak zastrzeżeniem, że nie są to formy odnoszące się do elementów gry, które poprzedzają rozgrywkę, jak np. wybór z menu startowego poziomu trudności, ale tylko te, które dotyczą już jej części właściwej (np. *HP* ‘punkty życia’).

Następnie z wykorzystaniem programu losującego wybrałem 15% z każdej z grup, ponieważ zakładam, że procesy lub tendencje, które pokazałyby się w 100% zebranego słownictwa graczy (1.732 jednostki) ujawnią się również w 15% tego słownictwa. Z drugiej strony, przebadanie całego materiału wymagałoby co najmniej większego zespołu badawczego. Historia badań ilościowych dostarcza przykładów właśnie takiego postępowania badawczego, że wybiera się próbki ze zgromadzonego materiału. Procedury te zostały zreferowane w rozdziale trzecim części teoretycznej.

Wśród wytypowanych wyrazów znalazły się jednostki leksykalne charakterystyczne tylko dla środowiska graczy wideo (np. *kieszonsolka* ‘przenośna konsola’), jednostki

⁶¹ Ulepszona wersja wydanej już wcześniej gry. Twórcy coraz częściej decydują się na wydanie gry, która miała swoją premierę kilka lub kilkanaście lat wcześniej, w nowszej, poprawionej wersji, która dostosowana jest do współczesnych growych realiów.

⁶² Wielosobowy tryb gry. Często jest on możliwy do wyboru z ekranu startowego gry, ale czasami jest trzonem gry, która posiada tylko grę w trybie *multi*. Przeciwnieństwem jest gra w trybie jednoosobowym (*singiel*).

⁶³ Usilnie powtarzać daną czynność w grze, żeby uzyskać nagrodę, najczęściej złoto lub punkty doświadczenia.

charakterystyczne dla wielu socjolektów, ale szczególnie często używane w środowisku graczy wideo (np. *sequel*) oraz jednostki przynależące do języka ogólnego, ale ważne dla środowiska graczy i częściej niż w języku ogólnym przez nie używane (np. *przejsiówka*). W przypadku pierwszej grupy do badań włączone zostały wyrazy z różną frekwencją, czyli również takie, które występują rzadko w badanych tekstach.

Ostatnie zdanie wymaga dodatkowego wyjaśnienia. Moim zdaniem ograniczenie się wyłącznie do jednostek charakterystycznych dla danego socjolektu i do tego jeszcze z wysoką frekwencją bardzo zubaża tego typu prace, ponieważ nie poznamy całego spektrum zabiegów językowych występujących w danych środowiskach. Jednostki rzadkie danych socjolektów także mogą być podstawą interesujących badań, które przynoszą wiele ciekawych odpowiedzi na temat badanej odmiany języka polskiego (pisała o tym J. Sambor w pracy *O słownictwie statystycznie rzadkim*. Książce tej poświęciłem fragment na stronie 50). Zdaję sobie sprawę z tego, że koncentrowanie się na słownictwie najczęstszym jest atrakcyjne naukowo. Podzielałm jednak pogląd tych badaczy, którzy nie tracą z pola widzenia również słownictwa rzadkiego, a nawet wyjątkowego. Aby nie ulec pokusie dużych liczb, zdecydowałem się na losowy wybór jednostek, które będą analizowane w pracy.

W efekcie wcześniejszego podziału na kręgi i późniejszego losowego wyboru z nich po 15%, lista wybranych wyjściowych jednostek leksykalnych wyglądała następująco:

Tabela 4. Lista jednostek wybranych do dalszych badań

Branża gier wideo	Środowisko graczy	Sprzęt gracza	Gra
1. Budżetowiec	1. Action RPG	1. Blaszak	1. Cel-shading
2. Cross-Buy	2. Ban	2. Bumper	2. Checkpoint
3. Cross-Chat	3. Bijatyka	3. Chlebak	3. Cooldown
4. Cross-Controler	4. Calak	4. Czarnała	4. Creditsy
5. Cross-gen	5. Casual	5. DC	5. Cutsцена
6. Crossover	6. Celowniczek	6. Dokanałówki	6. Deathmatch
7. Cross-over	7. Cheat	7. D-pad	7. Dungeon
8. Cross-platformowy	8. Craft	8. Dreamcast	8. Feature
9. Cross-Play	9. Crap	9. Drim	9. Ficzer
10. Cross-Save	10. Erpeg	10. Eksbokslan	10. Framerate
11. Current-gen	11. Farmić	11. Gacek	11. Gameplay
12. Demo	12. Gaming	12. GameCube	12. Hapek
13. Devkit	13. Geek	13. GC	13. Healer
14. DLC	14. Grind	14. Gładzik	14. HP
15. Egranizacja	15. Grywalny	15. Grzybek	15. Klatkować
16. Ekskluzyw	16. Hardkor	16. Handheld	16. Loading
17. Emulator	17. Hejt	17. Hardware	17. Marian
18. Ex	18. Jednostrzałowiec	18. Headset	18. Minimapa
19. Exclusive	19. JRPG	19. Kartridż	19. Multimedia
20. Flashówka	20. Każual	20. Kieszonsolka	20. NPC
21. Free2play	21. Krap	21. Kinect	21. Odblokowajka
22. Free-to-play	22. Liniowy	22. Kława	22. One-liner
23. Goty	23. Live-stream	23. Kławiatura	23. Otekstutowany
24. High-end	24. Masterowy	24. Komp	24. Pauza
25. Indie	25. Metagra	25. Komputer	25. Piaskownica
26. Indyk	26. Must-have	26. Konsola	26. Przedłużajka
27. Konwersja	27. MMO RPG	27. Kontroler	27. Przeszkadzajka
28. Lokalizacja	28. MMORPG	28. Kostka	28. QTE
29. Łatka	29. Naparzanka	29. Krzyżak	29. Quest
30. Mikrotransakcja	30. Niedzielnia	30. LPM	30. Randomizowacja

31. Multi	31. Niegrywalny	31. Makaron	31. Real-time
32. Multiplatformowy	32. Pełniak	32. Matryca	32. Respawn
33. Multiplayer	33. Płatyna	33. Multikart	33. Rozdzielczość
34. Multitasking	34. Pre-order	34. Multitap	34. Sandboks
35. Next-gen	35. Przetwaniówka	35. Multitouch	35. Sandbox
36. Offroad	36. Przygodówka	36. Myziadelko	36. Save
37. Off-road	37. Rogalik	37. NGC	37. Sejw
38. Ofrold	38. Roguelike	38. Nocniczek	38. Shader
39. Pacz	39. RPG	39. Notebook	39. Side-quest
40. Past-gen	40. RPG akcji	40. Ośmiobitowy	40. Skin
41. Patch	41. Samochodówka	41. Pad	41. Skórka
42. Pay-to-win	42. Samograj	42. Padlet	42. Skrypt
43. Peesplusowy	43. Scrafcić	43. PC	43. Soundtrack
44. Podstawa	44. Shmup	44. PC-t	44. Spawn
45. Pre-alfa	45. Shooter	45. Pecet	45. Sprajt
46. Pre-alpha	46. Stream	46. Peescztery	46. Sprite
47. Przeglądarka	47. Strzelanina	47. Peesczwórka	47. Sub-boss
48. PvE	48. Subskrypcja	48. Peesdwójka	48. Sub-quest
49. PvP	49. Ścigałka	49. Peesiks	49. Tank
50. Redzi	50. Tekstówka	50. Peeska	50. Tekstura
51. Reedycja	51. Ultrahardkor	51. Peesowy	51. Wiesiek
52. Remake	52. Unbox	52. Peespek	52. Zasejwować
53. Remaster	53. Używka	53. Peestrójka	53. Znajdzka
54. Rimejk	54. Wymasterować	54. Piec	54. Zrespawnować
55. Sequel	55. Zapychacz	55. PlayStation	
56. Sieciówka	56. Zbanować	56. Plej	
57. Singiel	57. Zblurowany	57. PPM	
58. Software	58. Żywotność	58. Przejściówka	
59. Spaczować		59. Przenośka	
60. Spatchować		60. PSP	
61. Trailer		61. PSP 2	
62. Turówka		62. PSP go	
63. Twarzowanie		63. QWERTY	
64. Wtyczka		64. Router	
65. Zegraniżować		65. RR0D	
66. Zremasterować		66. Smartfon	
		67. Stacjonarka	
		68. Stodwudziestośmiobitowiec	
		69. Stylus	
		70. Szarak	
		71. Szesnastobitowy	
		72. Tablet	
		73. Tableto-pad	
		74. Three-sixty	
		75. Touchpad	
		76. Trigger	
		77. WASD	
		78. WSAD	
		79. Xboks	
		80. Xbox	
		81. Xkloc	
		82. YLOD	

Jaki widać, w powyższej tabeli znalazły się nie tylko jednostki odnoszące się do różnych desygnatów, ale także te, które różnią się między sobą sposobem zapisu (np. *crap* i *krap*). Nie mogłem na tym etapie pracy połączyć ich i potraktować jako jedną jednostkę, ponieważ w dalszej części pracy będę wyszukiwał je po unikalnym układzie znaków, dlatego zachowałem taki podział.

Jak już wcześniej wspominałem, do badań wybrałem taki sam procent form podstawowych z każdego z czterech kręgów tematycznych, które powstały ze wszystkich wynotowanych do badań jednostek. Warto zaznaczyć, że na tym etapie pracy nie wiedziałem, jakie inne jednostki leksykalne (pochodne od tych powyższych) uda mi się pozyskać w wyniku badań. Te w tabeli powyżej stanowiły punkt wyjścia.

ETAP 4. PRACA Z KORPUSAMI TEKSTÓW

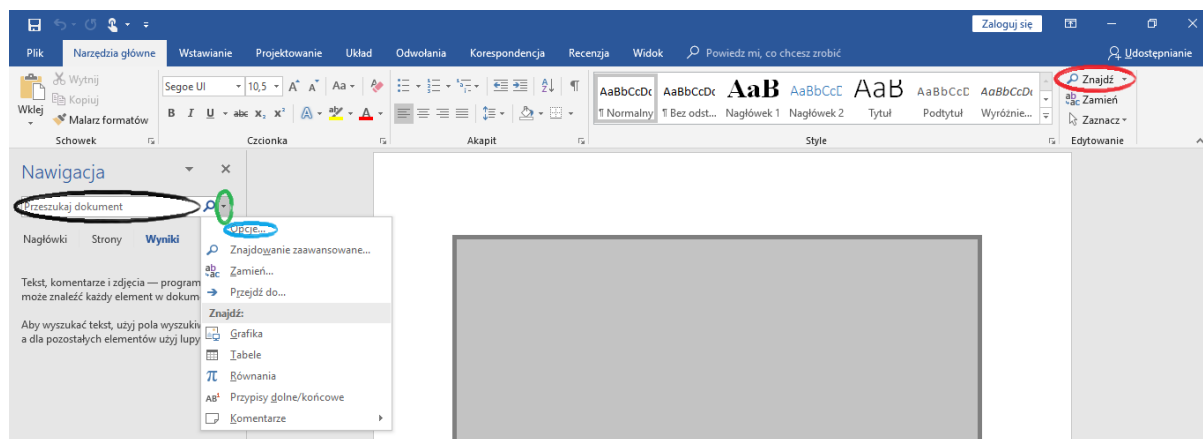
Po uzyskaniu edytowalnych plików wordowskich i stworzeniu listy jednostek do dalszej pracy, rozpocząłem część właściwą badań, a mianowicie wyszukiwanie w korpusach tekstów wystąpień wszystkich wyrazów tekstowych:

– będących reprezentantami wybranych wcześniej leksemów związanych ze środowiskiem graczy, chodzi tu przede wszystkim o formy fleksyjne oraz warianty graficzne/ortograficzne

– będących reprezentantami rodzin wyrazów, a więc form pochodnych słowotwórczo od wyjściowych leksemów

Do ich wyszukiwania wykorzystałem wbudowaną w program Word wyszukiwarkę wyrazów.

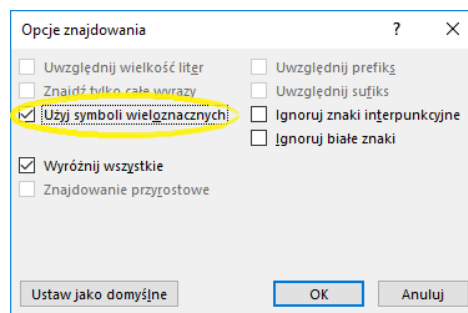
Aby ją włączyć, wybrałem z menu *Narzędzia główne*, w zakładce *edytowanie*, funkcję *znajdź* (zdj. 30, czerwona elipsa). W moich badaniach nie mogłem jednak skorzystać z niej, bez odpowiedniego skonfigurowania, ponieważ bez tego nie wyszukiwałaby mi ona tego, co było zgodne z celem badań. Aby dostosować ją do pracy z korpusem tekstów, rozwinąłem w okienku *nawigacja* menu *więcej możliwości wyszukania* (uczyniłem to, klikając małą strzałeczkę w prawej części okienka *nawigacja*, patrz zdj. 30, zielona elipsa), a następnie wszedłem w *opcje* wyszukiwarki (zdj. 30, niebieska elipsa).



Zdjęcie 30. Wyszukiwarka programu Word

W nowo otwartym okienku zazaczyłem opcję *Użyj symboli wieloznacznych* (zdjęcie 31, żółta elipsa). W dalszej części pracy wyjaśnię, w jakim celu to zrobiłem.

Mając tak przygotowaną wyszukiwarke programu Word, mogłem rozpocząć wyszukanie interesujących mnie wyrazów.



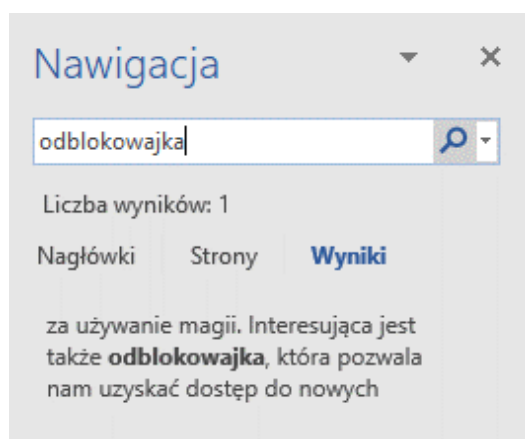
Zdjęcie 31. Konfiguracja wyszukiwarki

SPOSÓB WYSZUKIWANIA WYRAZÓW

Aby możliwie dobrze opisać wykorzystany w tej pracy sposób wyszukiwania wyrazów, czyli z zastosowaniem wyszukiwarki wbudowanej w program Word z włączoną opcją wykorzystania symboli wieloznacznych, zdecydowałem, że w tej części dysertacji wykorzystam przykłady autentycznych wyrazów ze środowiska graczy po to, aby zobrazować metodę pracy. Towarzyszące im dane liczbowe nie są jednak właściwe. Te zaprezentowane zostaną w części analitycznej.

W tym podrozdziale aby pokazać, dlaczego konieczne w moich badaniach było włączenie opcji korzystania z symboli wieloznacznych, posłużę się jako przykładem wyrazem *odblokowajka* ‘element do odblokowania w grze’. Wybrałem go, ponieważ nie posiada on rozbudowanej rodziny wyrazów oraz jest zbudowany z więcej niż 4 znaków/liter.

Gdybym do okienka wyszukiwarki (zdz. 30, czarna elipsa) wpisał wyraz *odblokowajka*, to program wyszukałoby mi wszystkie miejsca w tekście, w których użyto jego mianownikowej formy oraz, jeżeli wyniki wyszukiwania nie przekraczałyby 100 form wyrazowych, wyświetliłby mi fragment tekstu poprzedzający ten wyraz oraz fragment następujący po nim (patrz zdj. 32). Gdybym poszukiwał jedynie form



Zdjęcie 32. Wyszukiwanie wyrazu

mianownikowych tego wyrazu, to takie rozwiązanie by mi wystarczyło. Jest ono jednak niewystarczające w przypadku moich badań, ponieważ poza formami mianownikowymi, chciałbym znaleźć w korpusach tekstów wszystkie formy deklinacyjne tego wyrazu, czyli:

Tabela 5. Deklinowanie wyrazu *odblokowajka*

	Liczba pojedyncza	Liczba mnoga
MIANOWNIK	odblokowaj ka	odblokowaj ki
DOPEŁNIACZ	odblokowaj ki	odblokowaj ek
CELOWNIK	odblokowaj ce	odblokowaj kom
BIERNIK	odblokowaj kę	odblokowaj ki
NARZĘDNIK	odblokowaj ką	odblokowaj kami
MIEJSCOWNIK	odblokowaj ce	odblokowaj kach
WOŁACZ	odblokowaj ko	odblokowaj ki

OPRACOWANIE FORMUŁ WYSZUKIWANIA

Rozwiązaniem powyższego problemu, byłoby wpisywanie pojedynczych form deklinacyjnych do okienka wyszukiwarki, następnie zapisywanie/wynotowywanie liczby wystąpień danej formy, aż wreszcie zsumowanie wszystkich zebranych liczb. W ten sposób miałbym sumę wystąpień danego wyrazu łącznie ze wszystkimi jego formami. Rozwiązanie to sprawdziłoby się, ale jest ono czasochłonne. To właśnie w celu usprawnienia pracy z wyszukiwarką, załączona została opcja *korzystania z symboli wieloznacznych*.

W celu wyszukania wszystkich form, które mają po ciągu znaków **odblokowaj** literę **k**, czyli:

odblokowajka
odblokowajki (dop., l. poj.)
odblokowajkę
odblokowajką
odblokowajko
odblokowajki (mian., l. mn.)
odblokowajkom
odblokowajki (bier., l. mn.)
odblokowajkami
odblokowajkach
odblokowajki (woł., l. mn.)

wykorzystałem formułę `<odblokowajk*>`.

Formy, które po ciągu znaków **odblokowaj** mają literę **c**, czyli:

odblokowajce (cel., l.poj.)

odblokowajce (miej., l.poj.)

wyszukałem wykorzystując formułę **<odblokowajc*>**.

Chcąc wyszukać powyższe formy, mógłbym do okienka wyszukiwarki wpisać także po prostu wyraz *odblokowajce*, ponieważ, pomimo że są to różne przypadki gramatyczne, to posiadają one tylko jedną wersję graficzną dla dwóch form deklinacyjnych.

Natomiast, aby wyszukać wyrazy, które po ciągu znaków **odblokowaj** mają literę **e**, czyli w tym przypadku tylko jedna forma – **odblokowajek**, wpisałem **<odblokowaje*>**.

Podobnie jak powyżej, tutaj także mógłbym wpisać po prostu *odblokowajek*, pomijając różne dodatkowe symbole wpisane do okienka wyszukiwania.

Układając trzy powyższe formuły (zaznaczone żółtymi kolorami), wykorzystałem symbole wieloznaczne. Nawiasy kątowe są interpretowane przez program Word jako początek (<) i koniec wyrazu (>), natomiast gwiazdka (*) służy do uzupełnienia fragmentu wyrazu (np. **odblokowajk**) dowolnym ciągiem znaków (np. **om**)⁶⁴. Dzięki zastosowaniu gwiazdki program sam uzupełni ciągi literowe (*odblokowajk*, *odblokowajc* i *odblokowaje*) odpowiednimi literami, które w tabeli na stronie poprzedniej, zaznaczone zostały kolorem zielonym. Po zsumowaniu wyników wyszukiwania przy pomocy trzech powyższych formuł, otrzymałem liczbę wystąpień wszystkich form deklinacyjnych wyrazu *odblokowajka*.

POŁĄCZENIE FORMUŁ WYSZUKIWANIA

Rozwiązanie powyższe jest lepsze niż wyszukiwanie pojedynczo każdej z form deklinacyjnych wyrazu, ale w przypadku tej metody pod koniec badania także trzeba było wszystkie wyniki zsumować. Aby zlikwidować konieczność wykonywania dodatkowych działań matematycznych, a w rezultacie, po zakończeniu przez program wyszukiwania uzyskać już ostateczny wynik liczby wystąpień wszystkich form deklinacyjnych wyrazu *odblokowajka*, postanowiłem połączyć trzy powyższe formuły wyszukiwania (te oznaczone żółtą ramką), w jedną obszerniejszą – **<odblokowaj[ekc]*>**. Zastosowanie nawiasów kwadratowych ([])

⁶⁴ Jeśli wyszukiwana sekwencja kończy się gwiazdką np. <kot*>, to wyszukane zostaną wszystkie formy wyrazu kot, które rozpoczynają się ciągiem kot, np. kota, kotem..., ale także sam wyraz kot. Czyli gwiazdka nie oznacza, że musi tam znajdować się jakikolwiek znak lub litera.

spowodowało, że teraz program mógł wybrać, spośród umieszczonych między nawiasami liter, którą z nich uzupełnić ciąg podstawowy. Innymi słowy, po ciągu liter *odblokowaj* program sam uzupełnia ten ciąg znaków o literę *e*, *k* lub *c*, a następnie całość jest uzupełniana dowolnymi literami – dzięki zastosowaniu symbolu gwiazdka (*). W tym przypadku dzięki dodaniu gwiazdki, ciąg zostanie uzupełniony o takie zakończenia, jak: *a*, *i*, *e*, *q*, *o*, *om*, *ami*, *ach* lub *k* (patrz tabela na stronie 105, litery zaznaczone na zielono).

Nowa formuła (ta w zielonej ramce) wystarczyłaby do wyszukania wszystkich form deklinacyjnych wyrazu *odblokowajka*, ale niestety tylko tych, które występują w środku zdania lub na jego końcu, pominięte zostałyby natomiast te wyrazy, które rozpoczynają zdania. Dzieje się tak, ponieważ program otrzymuje ode mnie informację, że ma szukać wyłącznie wyrazy, które zaczynają się małą literą *o*, pomija więc te, które mają na początku duże *O*. Rozwiązaniem tego problemu jest dodanie do formuły wyszukiwania kolejnej pary nawiasów kwadratowych, które zawierają między sobą wielką i małą literę, wówczas formuła wyszukiwania wygląda następująco: <[Oo]dblokowaj[ekc]*>.

W ten sposób zakończyłem opis procesu powstawania formuły, dzięki której możliwe jest wyszukiwanie wszystkich form fleksyjnych wyrazu *odblokowajka*.

Gdy stworzyłem już formułę wyszukiwania, to otwierałem dany plik korpusu tekstów, uruchamiałem wyszukiwarkę i wpisywałem do niej: <[Oo]dblokowaj[ekc]*>. Wynik wyszukiwania, czyli liczbę wystąpień danego wyrazu w pliku tekstowym korpusu, wpisywałem do odpowiednio przygotowanej fiszki (tabelki), która wygląda następująco:

Tabela 6. Przykładowa fiszka (*odblokowajka*)

<[Oo]dblokowaj[ekc]*>																				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
psx	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0
cda	0	0	0	7	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0
npsx	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	8	8	0	0	0
ncda	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0
fpsx	0	7	0	0	0	0	6	0	0	0										
fcda	0	7	0	0	0	0	0	0	0	0	0	0	0							

W główce tabeli wpisywałem formułę wyszukiwania, która posłużyła mi do znalezienia danych wyrazów, w tym przypadku <[Oo]dblokowaj[ekc]*>. W pierwszej kolumnie umieściłem skróty, które reprezentują dany korpus (ich rozwinięcie podałem na stronie 77).

W drugim wierszu umieściłem liczby, które reprezentują dany kwartał. Cyfra 1 odnosi się do 1 kwartału 2011 roku, natomiast liczba 20 do 4 kwartału 2015 r. Warto tutaj zaznaczyć, że podział na kwartały dotyczył tylko czasopism i newsów, nie dotyczył natomiast forów internetowych (wyjaśnienie tej sytuacji dodałem na stronie 98).

W pozostałych komórkach umieszczałem sumy wystąpień poszczególnych wyrazów w danych korpusach.

Pozyskiwanie danych liczbowych było najbardziej czasochłonną częścią moich badań, ponieważ każdą jedną formułę wyszukiwania musiałem umieścić w aż 103 plikach⁶⁵ korpusu. Stworzenie fiszek zakończyło kolejny etap moich badań.

Zanim jednak przejdę do kolejnego etapu, chciałbym wspomnieć o kilku problemach powstałych podczas wyszukiwania danych wyrazów oraz pokazać, jak sobie z nimi poradziłem.

PROBLEM PIERWSZY – ZNAKI DIAKRYTYCZNE

Powyżej na przykładzie *odblokowajki* przedstawiałem sposób działania wyszukiwania przy pomocy symboli wieloznacznych. Przykład wykorzystany w tamtym podrozdziale był prosty, ponieważ nie zawierał znaków diakrytycznych. Nieco inaczej wygląda sposób pracy i tworzenia formuł wyszukiwania w przypadku, gdy trzeba znaleźć jakiś wyraz zawierających owe diakryty.

Jako przykład w tej części pracy posłuży mi wyraz *kartridż* ‘nośnik danych. Kość pamięci z danymi (np. gry) zamknięta w plastikowej obudowie przypominającej opakowanie do kasy magnetofonowej’.

Na wstępie pewna uwaga. Gdybym badał występowanie wyrazu *kartridż* w tekstach drukowanych czasopism, które przeszły dwie korekty (w tym jedną po składzie), to formuła wyszukiwania tego wyrazu (ze wszystkimi jego formami deklinacyjnymi) wyglądałaby tak – <[Kk]artridż*>. Jednak w skład mojego badania wchodzi nie drukowane czasopisma, ale ich elektroniczne wersje, pozyskane dzięki skanowaniu i rozpoznaniu przez program komputerowy (dokładnie temat omówiłem na w części *Pozyskiwanie edytowalnych materiałów tekstowych z czasopism drukowanych*), który czasami popełnia błędy w interpretacji znaków i zamiast

⁶⁵ Na tę liczbę składa się: 20 plików korpusu cda (1 plik = 1 kwartał), 20 plików neda, 20 plików psx, 20 plików npsx oraz 10 plików fpsx i 13 plików feda. W przypadku fpsx i feda podział na pliki został dokonany, aby komputery mogły je obsługiwać – jednego dużego pliku by się nie udało otworzyć. W tym przypadku jeden plik nie odpowiada jednemu kwartałowi, bo taki podział nie został dokonany.

ż może zinterpretować go jako z. Poza czasopismami w wersjach elektronicznych do moich badań włączyłem także newsy opublikowane na stronach www oraz wpisy na forach internetowych.

Newsy nie przechodzą korekty, dlatego zdarzają się tam wyrazy, które powinny zawierać znaki diakrytyczne, ale są ich pozbawione. Natomiast z całą pewnością takich wyrazów bez diakrytów można się spodziewać we wpisach na forach internetowych, na których ludzie często po prostu nie używają znaków diakrytycznych lub szerzej tzw. znaków polskich, ponieważ dla nich ważne jest szybkie przekazanie informacji, a nie dbałość o poprawność językową.

Reasumując, należało opracować sposób pozyskiwania wyrazów z ewentualnymi błędami/wariantami⁶⁶, gdyż odnosiły się do tego samego desygnatu. Pominięcie ich znacznie zniekształciłoby wyniki badań ilościowych, w których celem było zbadanie nie liczby poprawnie użytych form wyrazowych, ale właśnie liczby wystąpień wyrazów odnoszących się do poszczególnych desygnatów.

Ostatecznie zdecydowałem, że rozwiązaniem problemu będzie dodanie wariantowości [ż i z]. Gdy szukałem np. wyrazu *kartridż*, zamiast formy wyszukiwania – <[Kk]artridż*>, używałem <[Kk]artrid[żz]*> – formuły z dodatkową parą nawiasów kwadratowych, między którymi umieściłem literę z diakrytem – ż oraz bez diakrytu – z. W ten sposób program wyszukiwał wszystkie formy fleksyjne wyrazu *kartridż* oraz *kartridz*.

Wariantywność dodawałem także w wyrazach, w których występowały takie litery jak:

ą	ć	ę	ł	ń	ó	ś	ż	ź
---	---	---	---	---	---	---	---	---

W sytuacjach, w których wystąpiła któraś z podanych w powyższej tabeli litera, to umieszczałem ją w nawiasie kwadratowym i obok dopisywałem znak bez diakrytu, czyli dodawałem następujące warianty:

[ąa]	[ćc]	[ęe]	[łl]	[ńn]	[óo]	[śs]	[żz]	[źz]
------	------	------	------	------	------	------	------	------

Zdarzało się także, że w niektórych krojach pisma bardzo podobne do siebie są litery q i g, np. q i g (Verdana) czy q i g (Arial). Gdy w szukanym wyrazie wystąpiła jedna z tych liter, to także obejmowałem ją nawiasem kwadratowym i dodawałem drugą ([g lub q]).

⁶⁶ Moja praca dotyczy socjolektu graczy wideo, który jest tworem niestabilizowanym, dlatego też nie posiada jakiegokolwiek normy. W przypadku braku normy trudno mówić o błędach językowych, dlatego w powyższym zdaniu zastosowałem dwa określenia – błędy i warianty.

Do tej pory zaprezentowałem sposób wyszukiwania oraz to, jak rozwiązać problem ze znakami diakrytycznymi na przykładzie rzeczowników *odblokowajka* i *kartridż*.

W badanych przeze mnie korpusach rzadko zdarzało się, że dana rodzina wyrazów była jednoelementowa i składała się tylko z rzeczownika, częstszą sytuacją było to, że zbiór wyrazów zawierał co najmniej dwa elementy, np. rzeczownik i utworzony od niego przymiotnik. W tym momencie stanąłem przed problemem metodologicznym - czy poprzestać tylko na wybranych, na podstawie pierwszej intuicji, leksemach, czy też pozwolić na wciągnięcie przez program wyszukujący całych rodzin wyrazów. Ostatecznie zdecydowałem się włączyć do badań nie tylko wszystkie formy fleksyjne pierwotnie wybranych leksemów, ale również całe rodziny wyrazów spokrewnionych (wraz z ich formami fleksyjnymi). Uzasadniam to tym, że pierwotne dla socjolektu jest samo pojęcie, a więc sfera myśli, natomiast zakodowanie pojęcia w formie części mowy (rzeczownik, czasownik czy przymiotnik) jest już bardziej sprawą składni, a więc gramatyki. Stąd poszerzenie pola ekscerpcji.

W tej części pokażę zatem, jak postępowanie w przypadku, gdy wyniki liczbowe zostały pozyskane dla wyrazu, który wchodził w skład większej rodziny. W tym podrozdziale zagadnienie omówię na przykładzie rodziny wyrazów – *stream* ‘przesyłanie danych, np. transmitowanie rozgrywki danej gry przez internet’.

Pierwszą czynnością, gdy już wiedziałem, że mam do czynienia z rodziną wyrazów było sprawdzenie, jak bardzo jest ona rozbudowana. Uruchomiłem wyszukiwarkę i skorzystałem ze zwykłej formuły wyszukiwania, stworzonej tak jak np. przy wyrazie *odblokowajka*, czyli z wykorzystaniem nawiasów kwadratowych, kątowych oraz gwiazdki. W ten sposób powstała formuła <[Ss]tream*>. W wynikach wyszukiwania zobaczyłem, że poza formami fleksyjnymi wyrazu podstawowego *stream*, wyszukało mi wiele różnych form następujących wyrazów:

streaming (rzeczownik odrzeczownikowy),
streamingowy (przymiotnik),
streamingujący (imiesłów przymiotnikowy czynny),
streamuję (czasownik – forma osobowa),
streamując (imiesłów przysłówkowy współczesny),
streamujący (imiesłów przymiotnikowy czynny),
streamowy (przymiotnik),

streamować (czasownik – bezokolicznik),
streamowany (imiesłów przymiotnikowy bierny),
streamowanie (rzeczownik odczasownikowy),
streamer (rzeczownik – subiekt rodzaju męskiego),
streamerka (rzeczownik – subiekt rodzaju żeńskiego),
streamerski (przymiotnik odrzeczownikowy),
live-stream (rzeczownik)⁶⁷,
live-streamowy (przymiotnik),
live-streamowany (imiesłów przymiotnikowy bierny).

W badaniach zależało mi na tym, aby wyodrębnić różne wyrazy pochodne słotwórczo. Musiałem więc znaleźć sposób, jak je wyodrębnić. Udało się to dzięki zastosowaniu różnych formuł z wykorzystaniem znaków wieloznacznych, a następnie ułożeniu zależności między wynikami uzyskanymi podczas wyszukiwania, a także wykonaniu odpowiednich działań matematycznych.

Po zorientowaniu się, z jak dużą rodziną wyrazową mam do czynienia, ponownie do wyszukiwarki wpisałem formułę <[Ss]tream*>, a następnie uzupełniłem wcześniej przygotowaną fiszkę o wyniki wyszukiwania. Dane, które wpisałem do fiszki zawierały liczbę wszystkich wyrazów, które rozpoczynały się od ciągiem liter *Stream* lub *stream*, czyli miałem w niej łączną liczbę wystąpień wyrazów, których listę umieściłem powyżej (np. *streamerski*, *live-streamowy* czy *streaming*).

Kolejnym krokiem było ułożenie formuł wyszukiwania dla reszty wyrazów z listy. Na przykład dla rzeczownika odrzeczownikowego *streaming* ułożyłem formułę <[Ss]treaming*>, dzięki której program wyszukał mi wszystkie formy deklinacyjne wyrazy *streaming*, czyli:

Tabela 7. Deklinowanie wyrazu *streaming*

	L. POJEDYNCZA	L. MNOGA
MIANOWNIK	STREAMING	STREAMINGI
DOPEŁNIACZ	STREAMINGU	STREAMINGÓW
CELOWNIK	STREAMINGOWI	STREAMINGOM
BIERNIK	STREAMING	STREAMINGI

⁶⁷ Dlaczego wyszukane zostały także wyrazy rozpoczynające się ciągiem znaków *live* wyjaśnię w dalszej części pracy.

NARZĘDNIK	STREAMINGIEM	STREAMINGAMI
MIEJSCOWNIK	STREAMINGU	STREAMINGACH
WOŁACZ	STREAMINGU	STREAMINGI

a także formy przymiotnikowe:

Tabela 8. Deklinowanie wyrazu *streamingowy* (ujęcie 1)

	L. POJEDYNCZA	L. MNOGA
MIANOWNIK	STREAMINGOWY/OWA/OWE	STREAMINGOWI/OWE
DOPEŁNIACZ	STREAMINGOWEGO/OWEJ	STREAMINGOWYCH
CELOWNIK	STREAMINGOWEMU/OWEJ	STREAMINGOWYM
BIERNIK	STREAMINGOWY/OWE/OWA/OWEGO	STREAMINGOWYCH/OWE
NARZĘDNIK	STREAMINGOWYM/OWĄ	STREAMINGOWYMI
MIEJSCOWNIK	STREAMINGOWYM/OWEJ	STREAMINGOWYCH
WOŁACZ	STREAMINGOWY/OWA/OWE	STREAMINGOWI/OWE

oraz imiesłowy przymiotnikowe czynne:

Tabela 9. Deklinowanie wyrazu *streamingujący* (ujęcie 1)

	L. POJEDYNCZA	L. MNOGA
MIANOWNIK	STREAMINGUJĄCY/UJĄCA/UJĄCE	STREAMINGUJĄCY/UJĄCE
DOPEŁNIACZ	STREAMINGUJĄCEGO/UJĄCEJ	STREAMINGUJĄCYCH
CELOWNIK	STREAMINGUJĄCEMU/UJĄCEJ	STREAMINGUJĄCYM
BIERNIK	STREAMINGUJĄCEGO/UJĄCĄ/UJĄCE	STREAMINGUJĄCYCH/UJĄCE
NARZĘDNIK	STREAMINGUJĄCYM/UJĄCĄ	STREAMINGUJĄCYMI
MIEJSCOWNIK	STREAMINGUJĄCYM/UJĄCEJ/UJĄCEMU	STREAMINGUJĄCYCH
WOŁACZ	STREAMINGUJĄCY/UJĄCA/UJĄCE	STREAMINGUJĄCY/UJĄCE

Po zakończeniu wyszukiwania, podobnie jak w przypadku wyrazu *stream*, stworzyłem nową fiszkę, w której wpisałem formułę wyszukiwania – <[Ss]streaming*>, po czym uzupełniłem komórki tabeli o wyniki wyszukiwania. W ten sposób zebrałem dane na temat liczby wystąpień wszystkich wyrazów, które rozpoczynają się ciągiem znaków *streaming*,

a więc zarówno formy rzeczownikowe, przymiotnikowe, jak i imiesłowowe, które umieszczone zostały w trzech powyższych tabelach.

Następnie przystąpiłem do wyszukiwania form fleksyjnych przymiotnika *streamingowy*, które wyszukałem przy użyciu formuły <[Ss]treamingow*>, w ten sposób program wyszukał takie formy wyrazowe, jak:

Tabela 10. Deklinowanie wyrazu *streamingowy* (ujęcie 2)

	L. POJEDYNCZA	L. MNOGA
MIANOWNIK	STREAMINGOWY/A/E	STREAMINGOWI/E
DOPEŁNIACZ	STREAMINGOWEGO/EJ	STREAMINGOWYCH
CELOWNIK	STREAMINGOWEMU/EJ	STREAMINGOWYM
BIERNIK	STREAMINGOWY/E/A/EGO	STREAMINGOWYCH/E
NARZĘDNIK	STREAMINGOWYM/A	STREAMINGOWYMI
MIEJSCOWNIK	STREAMINGOWYM/EJ	STREAMINGOWYCH
WOŁACZ	STREAMINGOWY/A/E	STREAMINGOWI/E

Wyniki wyszukiwania także zapisałem w osobnej fiszce i przeszedłem do wyszukiwania form wyrazu *streamingujący* przy pomocy formuły <[Ss]treaminguj[aa]c*>. Program wyszukał wszystkie formy deklinacyjne danego imiesłowu⁶⁸:

Tabela 11. Deklinowanie wyrazu *streamingujący* (ujęcie 2)

	L. POJEDYNCZA	L. MNOGA
MIANOWNIK	STREAMINGUJĄCY/A/E	STREAMINGUJĄCY/E
DOPEŁNIACZ	STREAMINGUJĄCEGO/EJ	STREAMINGUJĄCYCH
CELOWNIK	STREAMINGUJĄCEMU/EJ	STREAMINGUJĄCYM
BIERNIK	STREAMINGUJĄCEGO/A/E	STREAMINGUJĄCYCH/E
NARZĘDNIK	STREAMINGUJĄCYM/A	STREAMINGUJĄCYMI
MIEJSCOWNIK	STREAMINGUJĄCYM/EJ/EMU	STREAMINGUJĄCYCH
WOŁACZ	STREAMINGUJĄCY/A/E	STREAMINGUJĄCY/E

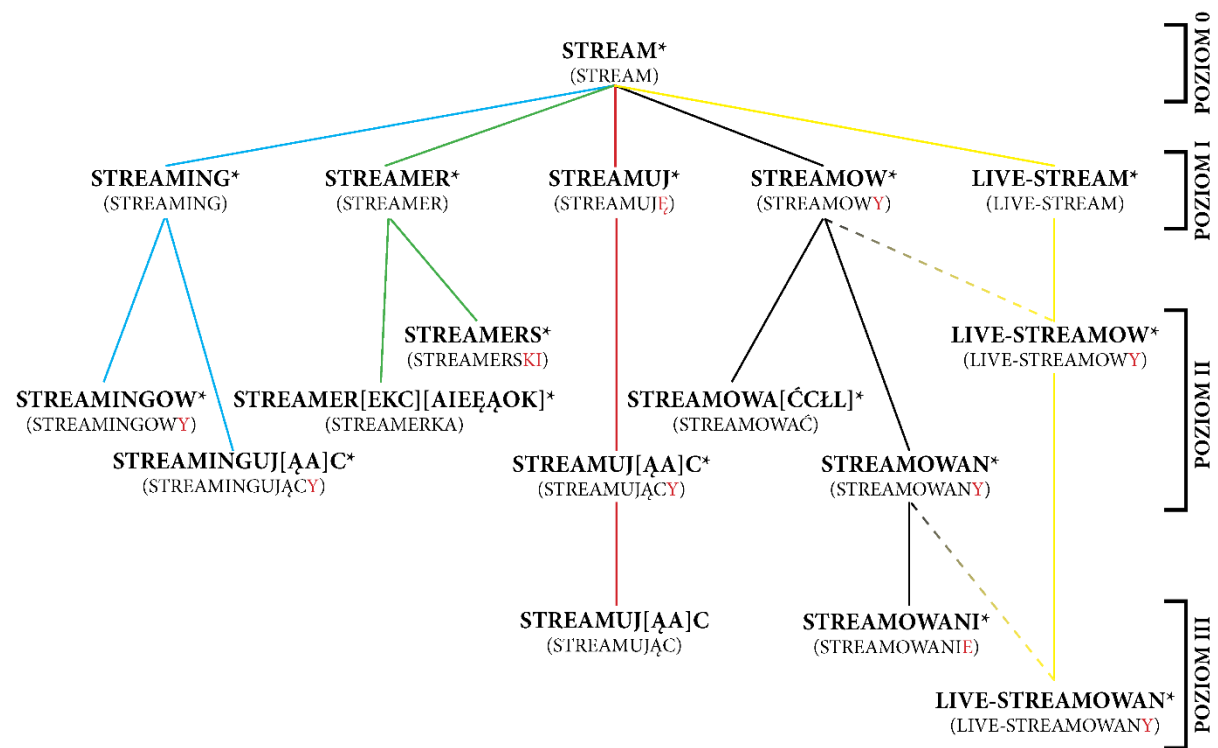
Wyniki powyższego wyszukiwania również zapisałem w nowej fiszce.

⁶⁸ Wyszukane zostały także formy, które zamiast ą w swojej budowie mają a. Nie umieszczałem ich w tabeli, ponieważ tylko dodatkowo utrudniłyby zrozumienie omawianych przykładów.

Reasumując, po wpisaniu do okienka wyszukiwarki formuły <[Ss]treaming*> program wyszukał mi formy deklinacyjne rzeczownika *streaming*, przymiotnika *streamingowy* oraz imiesłowu *streamingujący*. Po wpisaniu <[Ss]treamingow*> wyszukane zostały wyłącznie formy deklinacyjne przymiotnika *streamingowy*, natomiast po użyciu formuły <[Ss]treaminguj[ąa]c*> program wyszukał wyłącznie formy deklinacyjne imiesłowu *streamingujący*.

Aby uzyskać rzeczywiste dane na temat liczby wystąpień form deklinacyjnych rzeczownika *streaming*, musiałem odjąć od wyników wyszukiwania uzyskanych przy pomocy formuły <[Ss]treaming*>, te wyniki, które uzyskałem przy pomocy formuł <[Ss]treamingow*> oraz <[Ss]treaminguj[ąa]c*>.

Natomiast, aby uzyskać realną liczbę wystąpień form fleksyjnych wyrazu *stream* musiałem od wyników wyszukiwania przy pomocy formuły <[Ss]tream*> odjąć między innymi wyniki wyszukiwania uzyskane przy pomocy polecenia <[Ss]treaming*>. Napisałem *między innymi*, ponieważ, aby otrzymać rzeczywisty wynik liczebności form fleksyjnych *stream* w korpusie, trzeba odjąć także inne wyniki wyszukiwań, wśród których *streaming* jest tylko jednym z nich. Podrodzina *streaming* jest jedną z gałęzi dużego drzewa rodziny wyrazów *stream*. Cała rodzina wyrazowa wygląda następująco:



Grafika 7. Rodzina wyrazów *stream*

Na powyższej grafice zobrazowałem strukturę rodziny wyrazów *stream*. Głównymi elementami grafiki są formuły wyszukiwania poszczególnych wyrazów. Dla ułatwienia w nawiasach podałem przykładowe wyrazy, które wyszukane zostaną po wpisaniu danego polecenia do wyszukiwarki. Wszystkie formuły podzielone zostały na poziomy. Te, które umieszczone zostały na poziomie 1 są pod względem graficznym najbliższe formie z poziomu 0, czyli wyrazowi *stream*, natomiast te z poziomu 3, są graficznie najdalsze formie zerowej.

Formy wyrazowe są ze sobą połączone, tworząc swoiste gałęzie. Między połączeniami a poziomami istnieje dosyć ważna zależność. Zależność tę wyjaśnię na przykładzie gałęzi niebieskiej – *streaming*. Składa się ona z 2 poziomów. Swój początek ma na poziomie 2, na którym umieszczone zostały formuły wyszukiujące takie wyrazy, jak *streamingujący* i *streamingowy*. Formuły te (z poziomu 2) łączą się z formułą umieszczoną na poziomie pierwszym – *streaming*, a ta z kolei łączy się z formą zerową – *stream*. Zależność, o której wspominałem, jest taka, że wyniki wyszukiwania uzyskane osobno dla form fleksyjnych wyrazu *streamingujący* i osobno dla form *streamingowy* (obie formy umieszczone są na poziomie 2) są odejmowane od wyników uzyskanych dla wyrazu *streaming*, który jest na poziomie 1. Natomiast wyniki uzyskane dla *streaming* (poziom 1) odejmowane są od *stream* (poziom 0). Odejmowanie między poziomami można wykonać tylko wtedy, kiedy istnieje połączenie danych formuł wyszukiwania.

Mówiąc krótko, poziomy są tak ułożone, że od poziomu 0 odejmujemy wyniki wyszukiwania uzyskane przy pomocy formuł z poziomu 1, od wyników z poziomu 1 odejmujemy wyniki uzyskane przy pomocy formuł z poziomu 2, od wyników z poziomu 2 odejmujemy wyniki uzyskane przy pomocy formuł z poziomu 3, jednak aby wykonać odejmowanie musi istnieć połączenie między formułami wyszukiwania poszczególnych wyrazów. Dlatego nie można np. odjąć wyników uzyskanych przy pomocy formuły <[Ss]treamers*> (poziom 2) od wyników pozyskanych przy pomocy formuły <[Ss]treaming*> (poziom 1), ponieważ nie ma między nimi połączenia.

Na powyższej grafice każda z gałęzi drzewa wyrazowego *stream* została wyróżniona innym kolorem, aby łatwiej było wyjaśnić zależności między formami wyszukiwania wyrazów.

Gałąź niebieska została dokładnie omówiona w powyższej części tego podrozdziału, dlatego przejdę do omówienia gałęzi zielonej.

Wyjaśnienia wymaga, dlaczego w przypadku wyszukiwania form fleksyjnych wyrazu *streamerka* zastosowałem dwie pary nawiasów kwadratowych, zamiast, jak w innych przykładach, maksymalnie jednego, nie licząc oczywiście nawiasu początkowego, który łączy wielką i małą literę. Zdecydowałem się na dodanie nawiasów, ponieważ po wpisaniu frazy,

z pominięciem liter zawartych w ostatnim nawiasie kwadratowym, czyli <Streamer[ekc]*>, program wyszukałby mi takie formy jak:

streamerka
streamerki
streamerce
streamerkę
streamerką
streamerko
streamerek
streamerkom
streamerkami
streamerkach
streamerem

Jak widać, w wynikach wyszukiwania uwzględniona zostałaby także forma narzędnikowa wyrazu *streamer*, a chciałem uniknąć tego. Dlatego zabezpieczyłem się przed tym i dodałem drugi nawias, który uwzględniał litery [aieęąok], pomijał natomiast literę *m*.

W przypadku czerwonej nitki, czyli podrodziny *streamuj*, wyjaśnienia wymaga, dlaczego rozdzielone zostały dwie frazy wyszukiwania – <[Ss]treamuj[aa]c*> oraz <[Ss]treamuj[aa]c>. Pierwsza z nich wyszuka następujące wyrazy:

streamujący
streamująca
streamujące
streamującego
streamującej
streamujących
streamującemu
streamującym
streamującej
streamującą
streamującymi

streamując

+ wszystkie powyższe warianty, które zamiast *a* mają *a*

Jak widać, wyszukane zostały wszystkie formy imiesłowu przymiotnikowego czynnego *streamujący* oraz *streamując* – imiesłów przysłówkowy współczesny. Ostatnia z form jest niepożądana podczas tego wyszukiwania, na szczęście ma ona (imiesłów przysłówkowych współczesny) tylko jedną wersję graficzną (wyraz nieodmienny), więc chcąc wyszukać wyłącznie ten imiesłów zastosowałem frazę `<[Ss]treamuj[aa]c>`. Aby więc uzyskać wyłącznie formy imiesłowu przymiotnikowego czynnego musiałem odjąć od wyników wyszukiwania przy pomocy formuły `<[Ss]treamuj[aa]c*>` wynik wyszukiwania z zastosowaniem `<[Ss]treamuj[aa]c>`.

W przypadku czarnej gałęzi wyjaśnić należy, dlaczego w formule wyszukiwania `<[Ss]treaamowa[ćell]*>` w nawiasach kwadratowych umieściłem aż 4 litery. Literę *ć* umieściłem po to, aby wyszukane zostały bezokoliczniki, *c* dodałem, aby program uwzględnił także bezokoliczniki zapisane bez znaków diakrytycznych, *ł* i *l*, aby program wyszukał formy czasowników w czasie przeszłym oraz przyszłym, który wykorzystuje 3 os. czasu przeszłego, a także, aby uwzględnione zostały formy trybu przypuszczającego.

Ostatnia gałąź – żółta reprezentuje podrodzinę *live-stream*. Można zadać pytanie, dlaczego ten wyraz znalazł się na grafice, przecież on nie rozpoczyna się ciągiem znaków *stream*, tylko *live*, a na stronie 106 napisałem, że umieszczając ciąg znaków między nawiasami kątowymi ($\langle \rangle$), program zinterpretuje go jako jeden wyraz. Dotychczas taka informacja wystarczyła, żeby wyjaśnić różne procesy związane z wyszukiwaniem metodą z wykorzystaniem symboli wieloznacznych, natomiast teraz trzeba ją rozszerzyć. Ciąg znaków między nawiasami kątowymi ($\langle \rangle$) program rzeczywiście rozumie jako jeden wyraz, który przez program rozumiany jest jako ciąg znaków od spacji do spacji, ale także jako ciąg od spacji do znaku interpunkcyjnego lub od znaku interpunkcyjnego do spacji. Dywiz użyty w konstrukcji *live-stream* jest znakiem interpunkcyjnym⁶⁹, dlatego wyraz ten znalazł się w tym zestawieniu. W przypadku tej gałęzi zastosowałem pewne, dodatkowe zabezpieczenie – zarówno pierwszą literę części *live* oraz pierwszą literę części *stream*, wzięłem w nawiasy kwadratowe i dopisałem wielkie litery, ponieważ badania pokazały mi, że jest on (i inne tego typu podobne konstrukcje z dywizem w środku) zapisywany różnie.

⁶⁹ Bywa także określany znakiem ortograficznym

Na grafice poza zwykłymi liniami, umieszczone zostały jeszcze dwie przerywane. Pierwsza łączy <[L]ive-[Ss]treamowan*> i <[Ss]treamowan*>, a druga <[L]ive-[Ss]treamow*> i <[Ss]treamow*>. Połączenie to wynika bezpośrednio z tego, co napisałem kilka linijek wyżej, czyli z rozumienia przez program określenia wyraz. Po wpisaniu do wyszukiwarki <[Ss]treamow*>, poza formami fleksyjnymi przymiotnika *streamowy*, wyszukane zostaną wszystkie formy fleksyjne *live-streamowy*, *streamowany* i *live-streamowany*. Dlatego, aby uzyskać realny wynik wystąpień podrodziny wyrazów *streamowy*, musiałem odjąć poza wynikami uzyskanymi przy pomocy formuł <[Ss]treamowa[ćł]*> i <[Ss]treamowan*> także wyniki wyszukiwania uzyskane przy pomocy formuły <[L]ive-[Ss]treamow*>, analogicznie musiałem postąpić w przypadku pary <[Ss]treamowan*> i <[L]ive-[Ss]treamowan*>, czyli od wyników pozyskanych przy pomocy pierwszej formuły odjąć wyniki zgromadzone przy pomocy drugiej.

W ramach podsumowania tego podrozdziału załączam poniżej tabelę, w której schematycznie pokazałem przykładowe wyniki wyszukiwania po wpisaniu danej frazy, a także działania, jakie należy wykonać, aby uzyskać realną liczbę wystąpień form fleksyjnych danego wyrazu.

Tabela 12. Przedstawienie procesu wyszukiwania

Część 1 Wpisywana fraza	Część 2		Część 3		
	Wyniki wyszukiwania	Działania 1	Wyniki po 1 działaniu	Działania 2	Realna liczba
<[Ss]tream*>	150		150	150-10-21-9-49-25	36
<[Ss]treaming*>	10		10	10-3-2	5
<[Ss]treaminguj*>	3		3	brak	3
<[Ss]treamingow*>	2		2	brak	2
<[Ss]treamer*>	21		21	21-4-11	6
<[Ss]treamers*>	4		4	brak	4
<[Ss]treamer[kce][aieęąok*>	11		11	brak	11
<[Ss]treamuj*>	9		9	9-8	1
<[Ss]treamuj[ąa]c*>	8		8	8-5	3
<[Ss]treamuj[ąa]c>	5		5	brak	5
<[Ss]treamow*>	69	69-20	49	49-18-7	24
<[Ss]treamowa[ćł]*>	18		18	brak	18
<[Ss]treamowan*>	26	26-19	7	7-1	6
<[Ss]treamowani*>	1		1	brak	1
<[L]ive-[Ss]tream*>	25		25	25-20	5
<[L]ive-[Ss]treamow*>	20		20	20-19	1
<[L]ive-[Ss]treamowan*>	19		19	brak	19
					Suma: 150

Powyższą tabelę podzieliłem na 3 części. W pierwszej umieściłem formuły wyszukiwania, które zaprezentowałem poprzedzając odpowiednim wcięciem akapitowym. W ten sposób bardziej obrazowo zademonstrowałem zależności między frazami wyszukiwania – wyniki uzyskane przy pomocy formuł, które są bardziej wcięte należy odjąć od tych, które są wyżej i mniej wcięte. Wcięcia te zachowałem także w części drugiej i trzeciej tabeli, prezentując przykładowe dane liczbowe.

Część druga zawiera wyniki wyszukiwania uzyskane przy pomocy poszczególnych formuł umieszczonych w części pierwszej oraz kolumnę *działania 1*. Powstanie tej kolumny wymuszone zostało tym, że program Word interpretuje wyraz m.in. jako ciąg znaków od znaku interpunkcyjnego do spacji. Zobrazowałem w tej części tabeli to, o czym pisałem nieco wyżej, czyli konieczność odjęcia wyników wyszukiwania <[L]ive-[Ss]treamow*> od <[Ss]treamow*> i <[L]ive-[Ss]treamow*> od <[Ss]treamow*>, gdybym tego nie zrobił na początku, to dane byłyby zdublowane i wyniki badania zniekształcone.

W części trzeciej przedstawiłem zależności między poszczególnymi formami wyszukiwania oraz działania matematyczne, jakie trzeba wykonać, aby uzyskać właściwe wyniki liczby wystąpień danego wyrazu z uwzględnieniem wszystkich jego form fleksyjnych.

W ostatnim wierszu tabeli umieściłem sumę wszystkich form wyrazowych całej rodziny wyrazów *stream*. Suma stanowiła swoistą bramkę bezpieczeństwa. Gdyby liczba umieszczona w polu *suma* nie była równa wynikowi wyszukiwania przy pomocy formuły <[Ss]treamow*>, to znaczyłoby, że zapomniałem o jakiejś formie wyrazowej z tej rodziny lub pomyliłem się w obliczeniach.

Na zakończenie tej części warto jeszcze zaznaczyć, że specyfika języka polskiego w odniesieniu do czasowników zakończonych na *-ować*, wymaga tego, aby na samym końcu pracy z rodziną wyrazów *stream*, zsumować wyniki wyszukiwania dwóch różnych ciągów literowych, czyli tego z częstka *-uj-* (<[Ss]treamuj*>) oraz tego z *-owa-* (<[Ss]treamowa[ćł]*>). W ten sposób otrzymałem realną liczbę wystąpień czasownika *streamować*.

Nakładanie się w badaniach korpusowych różnego typu form morfologicznych: fleksyjnych i słowotwórczych jest nieuniknione. Wynika to z wieloznaczności formuł wyszukiwania, ponieważ system rozpoznaje części wyrazów bez ich podziału na morfemy fleksyjne bądź słowotwórcze. Pierwotnie badacz otrzymuje łączną grupę wystąpień, natomiast ich podział jest determinowany przez precyzyjność formuł oddzielających kolejne podgrupy. Na tym przykładzie widać, że same dane ilościowe są dość surowym materiałem, który zawsze będzie wymagał interpretacji językoznawczej.

PROBLEM TRZECI – KRÓTKIE WYRAZY

Problem omówiony w tym podrozdziale dotyczy najczęściej wyrazów, które w formułach wyszukiwania mają 3 lub mniej znaków. Jako przykład do omówienia tego problemu posłuży mi w tej części wyraz *demo*⁷⁰ ‘demonstracyjna wersja programu np. gry’, którego forma wyszukiwania, po uwzględnieniu wszystkich przytoczonych w poprzednich rozdziałach informacji, wyglądałaby następująco: <[Dd]em*>, ponieważ jego formy deklinacyjne to:

Tabela 13. Deklinowanie wyrazu *demo*

	Liczba pojedyncza	Liczba mnoga
MIANOWNIK	demo	dema
DOPEŁNIACZ	dema	dem
CELOWNIK	demu	demom
BIERNIK	demo	dema
NARZĘDNIK	demem	demami
MIEJSCOWNIK	demie	demach
WOŁACZ	demo	dema

Po wpisaniu do okna wyszukiwania powyższego polecenia, poza formami fleksyjnymi wyrazu *demo* wyszukiwarka programu Word znalazła dużo wyrazów, które rozpoczynają się ciągiem liter *dem*, ale z *demem* mają niewiele wspólnego. Były to wyrazy takie jak np.:

demoniczny
demon
Demon’s Souls (tytuł gry)
demolować
demotywator
demokracja
demencja
demoralizacja
demontaż
itd.

Postanowiłem więc w przypadku krótkich wyrazów zastosować inną, dłuższą metodę.

⁷⁰ W *Wielkim słowniku ortograficznym PWN* napisano, że *demo* można traktować jako wyraz nieodmienny lub odmienny [2017: 332]. W środowisku graczy wideo jest on odmieniany, więc w tej pracy także będę go odmieniał.

Po pierwsze podzieliłem wszystkie formy fleksyjne wyrazu *demo* ze względu na liczbę znaków, które je budują. W efekcie dla wszystkich form fleksyjnych stworzyłem 4 grupy:

Tabela 14. Podział form deklinacyjnych wyrazu *demo* ze względu na liczbę liter

Wyraz trzyliterowy	Wyrazy czteroliterowe	Wyrazy pięcioliterowe	Wyrazy sześcioliterowe
dem	demo dema demu demo demo dema dema dema	demem demie demom	demami demach

a następnie ułożyłem frazy wyszukiwania dla wyrazów z danej grupy:

Tabela 15. Formuły wyszukiwania dla wszystkich form deklinacyjnych wyrazu *demo*

Dla wyrazu trzyliterowego	Dla wyrazów czteroliterowych	Dla wyrazów pięcioliterowych	Dla wyrazów sześcioliterowych
<[Dd]em>	<[Dd]em[oau]>	<[Dd]em[eio][me]>	<[Dd]ema[mc][hi]>

Jak wywnioskować można z obserwacji powyższych formuł, rozwiązaniem problemu krótkich wyrazów było zrezygnowanie z symbolu gwiazdki na poczet większej liczby nawiasów kwadratowych zawierających wariantywności użycia liter.

Wyjaśnienia wymaga to, dlaczego w grupie wyrazów sześcioliterowych zamiast trzech nawiasów kwadratowych zastosowałem tylko dwa (pomijając oczywiście ten, który zawiera wielką i małą literę na początku wyrazu). Zdecydowałem się na taki krok, ponieważ w formach sześcioliterowych, na początku, występuje czteroliterowy ciąg znaków, który powtarza się w obu formach wyrazowych. Ciąg ten (**dema**) jest o jeden znak dłuższy niż w poprzednich grupach (**dem**), powtarza się litera a na czwartej pozycji, więc w przypadku tej pozycji nie ma żadnej wariantywności.

Po wyszukaniu form wyrazowych przy użyciu czterech powyższych formuł, wyniki zsumowałem i w efekcie uzyskałem realną liczbę wystąpień wszystkich form fleksyjnych wyrazu *demo*.

PROBLEM CZWARTY – WIELOZNACZNOŚĆ

Problem wieloznaczności poruszany był przez wielu badaczy, którzy podejmowali badania ilościowe lub korpusowe (pisałem o tym między innymi na stronie 56) i nigdy tak naprawdę nie został on w pełni rozwiązany. Duże korpusy (takie jak PWN czy Narodowy Korpus Języka Polskiego) radzą sobie z wieloznacznością poprzez stosowanie tagów, czyli dodatkowych informacji dodawanych do wyrazu, np. jaką jest on częścią mowy lub w jakim znaczeniu wystąpił w danym kontekście. Tagowanie jest zajęciem żmudnym i czasochłonnym, dlatego ja nie zdecydowałem się na otagowanie całego zgromadzonego materiału, wolałem poszukać innej metody na rozwiązanie problemu wieloznaczności.

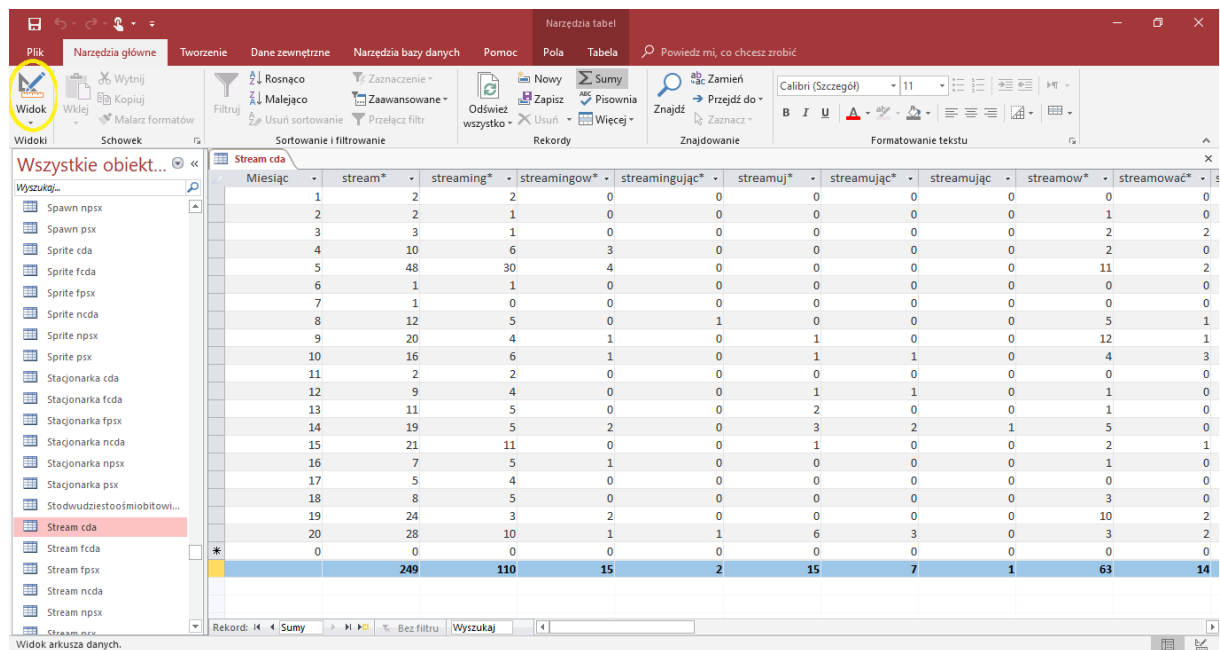
Niewątpliwie pomocna w tym zadaniu byłaby analiza różnych łączliwości danych wyrazów oraz ich otoczenia leksykalnego, lecz aby wykonać to automatycznie (przy pomocy komputera) trzeba by było stworzyć olbrzymi zestaw poleceń, który pokazałby komputerowi, czego ma szukać oraz konieczne byłoby posiadanie komputera (najlepiej superkomputera), który obsłużyłby tak zaawansowany sposób wyszukiwania. Rozwiązaniem byłoby wykorzystanie do tego celu sztucznej inteligencji, ale prace nad nią wciąż nie osiągnęły zadowalających wyników.

Jak więc rozwiązałem ten problem w mojej pracy? Pomocna okazała się sama wyszukiwarka programu Word. Po wpisaniu do niej danego słowa/formuły, poza liczbą wystąpień danego wyrazu, pokazywane jest także jego otoczenie (pisałem o tym na stronie 104). Przeglądałem więc poszczególne wyniki wyszukiwania, wyraz po wyrazie i wynotowywałem, na podstawie kontekstu, jakie znaczenie danego wyrazu zostało użyte. Gdy miałem za mało informacji, żeby stwierdzić, z jakim znaczeniem mam do czynienia, ponieważ nie wystarczał krótki kontekst podany w okienku wyszukiwania, to klikałem na wyraz i program przenosił mnie do miejsca, w którym dany wyraz wystąpił, a wtedy mogłem zapoznać się z jego większym otoczeniem wyrazowym.

W niniejszym rozdziale do tej pory zaprezentowałem sposób wyszukiwania z wykorzystaniem metody symboli wieloznacznych oraz przedstawiłem, jakie problemy wiążą się z tą metodą, a także, jak sobie z nimi poradziłem. Należy w tym miejscu podkreślić, że

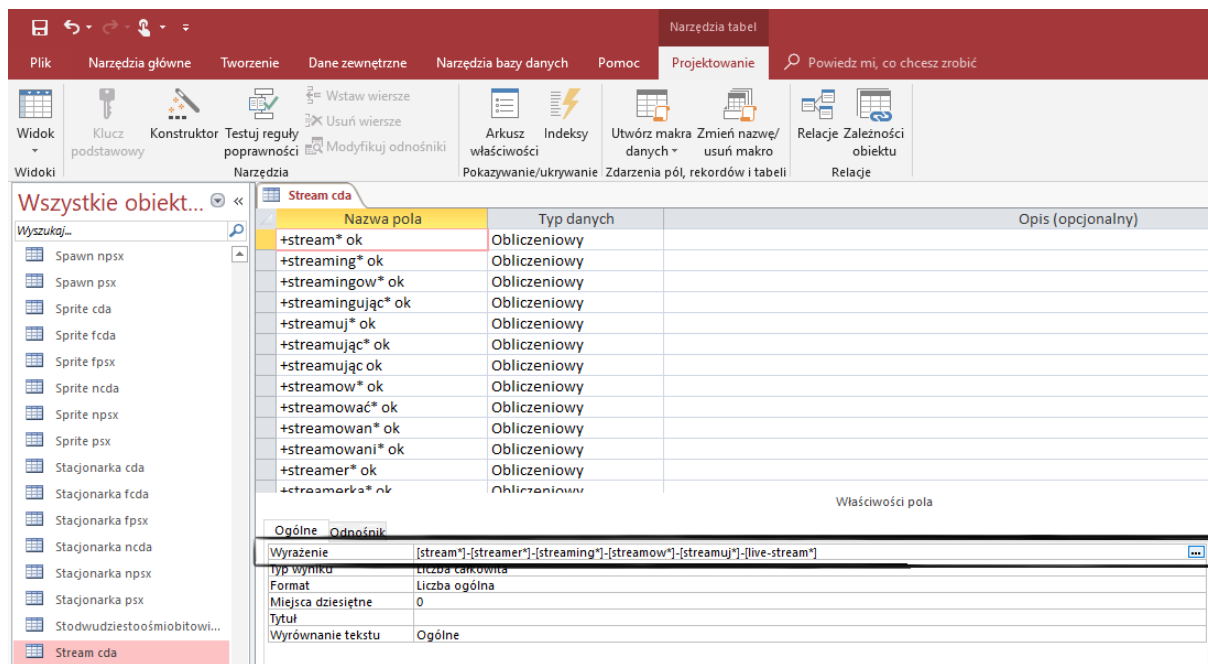
pokazane powyżej problemy mogły też nakładać się na siebie. Musiałem wtedy łączyć poszczególne rozwiązania, dokładając wszelkich starań, aby wyniki były dokładne i satysfakcjonujące.

Aby nieco ułatwić sobie dalszą pracę, zwłaszcza w przypadku wykonywania różnych działań matematycznych w obrębie wyników liczbowych (to, o czym pisałem przy problemie drugim, s. 110), postanowiłem przenieść informacje z fiszek do stworzonej bazy danych w programie Microsoft Access.



Zdjęcie 33. Główne okno programu Microsoft Access

Stworzona baza danych składała się z wielu tabel, natomiast jedna tabela zawierała wyniki wyszukiwania jednej rodziny wyrazów w jednym korpusie tekstów (np. rodzina *stream* w korpusie cda). W samej tabeli zaś umieszczone zostały wyniki wyszukiwania z podziałem na kwartały (patrz zdj. 33). Żeby móc dodawać różne formuły matematyczne do danej tabeli, musiałem zmienić *widok* okna programu. Uczyniłem to klikając przycisk w lewym górnym rogu programu (patrz zdj. 33, żółta elipsa). Tabela zmieniała swoją formę i teraz mogłem modyfikować ją w fazie projektu. Faza projektu od fazy tabeli różni się tym, że w fazie projektu program nie pokazuje danych liczbowych, a jedynie sam szkielet tabeli, czyli nagłówki kolumn oraz informacje o tym, jaki rodzaj danych one zawierają.



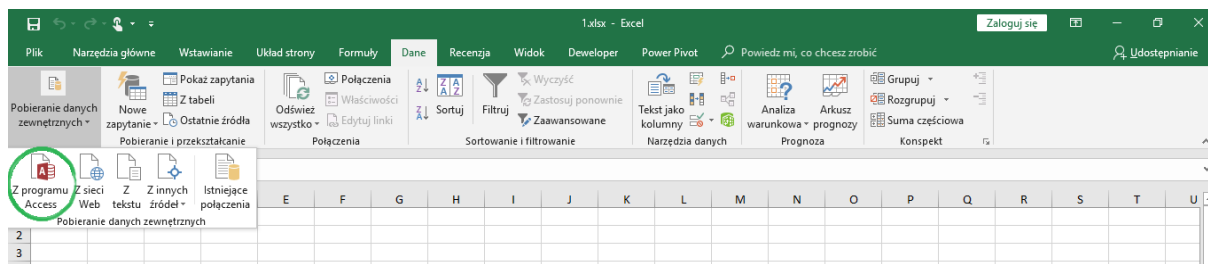
Zdjęcie 34. Widok projektu w programie Microsoft Access

Jak już wcześniej wspomniałem, zdecydowałem się na gromadzenie informacji w postaci bazy danych, ponieważ w ten sposób mogłem łatwo odejmować poszczególne wyniki wyszukiwania. Czyniłem to, umieszczając/wpisując poszczególne działania matematyczne w polu wyrażenie (zdj. 34, czarna ramka).

Na powyższym zdjęciu widać, że program w polu *stream* będzie przechowywał tylko te wyniki, które powstaną przez odjęcie danych zgromadzonych w kolumnach *streamer**, *streaming**, *streamow** *streamuj** i *live-stream** od kolumny *stream**.

W tym programie różne działania matematyczne mogłem dodać do każdej kolumny wybranej tabeli.

Gdy już uzyskałem rzeczywiste dane na temat liczby występów w korpusie danych wyrazów, które wyliczone zostały przez program na podstawie działań przeze mnie wprowadzonych, to mogłem przenieść wyniki (bazę danych Access) do ostatniego już programu – Microsoft Excel. Programy z pakietu Office współpracują ze sobą, dlatego przeniesienie danych nie stanowiło dużego problemu. Wybrałem w Excelu opcję *pobieranie danych zewnętrznych*, a następnie z rozwiniętej zakładki przycisk z *programu Access* (patrz zdj. 35, zielona elipsa).



Zdjęcie 35. Okno programu Excel

W tym programie pracowałem już do samego końca badań, ponieważ tutaj najłatwiej było mi sumować dane zgromadzone w poszczególnych kolumnach, wierszach czy też całych tabelach, wyciągać średnie oraz wykonywać inne potrzebne działania matematyczne, a następnie w łatwy sposób tworzyć z nich zestawienia tabelaryczne czy też wykresy.

ETAP 5. STWORZENIE PÓL LEKSYKALNYCH

Skorzystanie z wyszukiwarki pozwoliło na dopełnienie pól wyrazowych o dodatkowe formy podstawowe. Na przykład po wpisaniu do wyszukiwarki formuły wyszukującej słowoformy formy podstawowej *calak* dowiedziałem się, że istnieją także takie formy podstawowe, jak: *calak*, *calakować*, *calakowanie*, *calakowany*, *calakowy*, *calakujący*, *calakujący* itd.

W efekcie z 260 form podstawowych (tych zaprezentowanych w tabeli 4) dzięki wyszukiwaniu, ich liczba rozszerzona została do 709 (patrz poniższa tabela).

Tabela 16. Wykaz pól leksykalnych i ich form podstawowych

Nazwa pola	Nazwa formy podstawowej	Złożoność pola	Nazwa pola	Nazwa formy podstawowej	Złożoność pola
Action RPG		4	Cheat		7
	Action RPG			Cheat	
	Action-RPG			Cheater	
	RPG akcji			Cheaterskie	
	RPG-akcji			Cheaterstwo	
Ban		11		Cheatować	
	Ban			Cheatowanie	
	Banować			Cheatowany	
	Banowalność		Checkpoint		2
	Banowalny			Checkpoint	
	Banowanie			Cheackpointowy	
	Banowany		Craft		10
	Banując			Craft	
	Banujący			Crafter	
	Zbanować			Crafting	
	Zbanowanie			Craftingowy	
	Zbanowany			Craftować	
Bijatyka		4		Craftowalny	
	Bijatyka			Craftowanie	
	Bijatykowiec			Craftowy	
	Bijatykowy			Craftujący	
	Naparzanka			Scrafcić	
Budżetówka		2	Crap		15
	Budżetowiec			Crap	
	Budżetówka			Crapcom	
Calak		17		Crapiarz	
	Calak			Crapiastry	
	Calakować			Crapik	
	Calakowanie			Crapisko	
	Calakowany			Crapiszcz	
	Calakowy			Crapiszon	
	Calakując			Crapowatość	
	Calakujący			Crapowaty	
	Masterować			Crapowy	
	Masterowanie			Crapszot	
	Masterowany			Crapulec	
	Masterowicz			Krap	
	Masterowy			Krapiszcz	
	Masterując		Cross-controller		3
	Masterujący			Cross-control	
	Wymasterować			Cross-controller	
	Wymasterowanie			Cross-controller	
	Wymasterowany		Cross-gen		3
Casual		24		Cross-gen	
	Casual			Cross-generacyjny	
	Casualizacja			Cross-genowy	
	Casualka		Crossover		2
	Casualowanie			Crossover	
	Casualowatość			Cross-over	
	Casualowaty		Cross-platform		4
	Casualowiec			Cross-platform	
	Casualowienie			Cross-platforming	
	Casualowo			Cross-platformowość	
	Casualowość			Cross-platformowy	
	Casualowy		Current-gen		2
	Casualówka			Current-gen	
	Każual			Current-genowy	
	Każualista		Cutscena		4
	Każualizacja			Cutscena	
	Każualizm			Cutscenes	
	Każualizowanie			Cutscenka	
	Każualność			Cutscenkowy	
	Każualowatość		Czarnula		2
	Każualowiec			Czarnula	
	Każualowo			Czarnulka	
	Każualowość		Deathmatch		2
	Każualowy			Deathmatch	
	Każualówka			Deathmatchowy	
Celowniczek		2	Demo		2
	Celowniczek			Demko	
	Celowniczkowy			Demo	
Cel-shading		3	D-pad		3
	Cel-shading			D-pad	
	Cel-shadingowo			D-padowy	
	Cel-shadingowy			Krzyżak	

Nazwa pola	Nazwa formy podstawowej	Złożoność pola	Nazwa pola	Nazwa formy podstawowej	Złożoność pola
Dreamcast		5		Grindując	
	Dreamcast			Grindujący	
	Dreamcastowy			Grindzik	
	Drim		Grywalny		7
	DC			Grywalnie	
	Makaron			Grywalnościowy	
Dungeon		3		Grywalność	
	Dungeon			Grywalny	
	Dungeonowanie			Niegrywalnie	
	Dungeonowy			Niegrywalność	
Egranizacja		5		Niegrywalny	
	Egranizacja		Handheld		6
	Egranizować			Handheld	
	Egranizowanie			Handheldowy	
	Zegranizować			Kieszonsolka	
	Zegranizowany			Kieszonsolkowy	
Emulator		3		Przenośka	
	Emulator			Przenośniak	
	Emulatorek		HP		3
	Emulatorowy			Hapek	
Exclusive		7		HP	
	Ekskluziw			HPk	
	Ekskluzyw		Hardkor		15
	Ekskluzywnie			Ultrahardkor	
	Ekskluzywny			Hardkor	
	Ekskulzywność			Hardkorok	
	Ex			Hardkorować	
	Exclusive			Hardkorowanie	
Farmić		7		Hardkorowiec	
	Farmiąc			Hardkorowo	
	Farmiący			Hardkorowość	
	Farmić			Hardkorowy	
	Farmienie			Hardkorówa	
	Farming			Hardkorówka	
	Farmingowy			hardkoryzacja	
	Farmiony			Hardkoryzm	
Feature		3		Hardkorzasty	
	Feature			Hardkorzyć	
	Featurek		Hardware		3
	Ficzer			Hardware	
Free-to-play		3		Hardware'owo	
	Free2play			Hardware'owy	
	Free-to-play		Headset		2
	Free-to-playowy			Headset	
GameCube		8		Headsetowy	
	Gacek		Healer		2
	Gackowy			Healer	
	GameCube			Healerka	
	gamecube'owy		Hejt		16
	GC			Hejcić	
	NGC			Hejcik	
	Kostka			Hejt	
	Nocniczek			Hejter	
Gameplay		3		Hejterek	
	Gameplay			Hejterka	
	Gameplayowo			Hejterować	
	Gameplayowy			Hejterski	
Gaming		4		Hejterstwo	
	Gaming			Hejteryzm	
	Gamingowo			Hejting	
	Gamingowość			Hejtować	
	Gamingowy			Hejtowanie	
Geek		4		Hejtowany	
	Geek			Hejtując	
	Geekowaty			Hejtujący	
	Geekowski		High-end		3
	Geekowy			High-end	
Grind		12		High-endowiec	
	Grind			High-endowy	
	Grinder		Indie		6
	Grinderski			Indie	
	Grinding			Indyczek	
	Grindować			Indyczy	
	Grindowalny			Indyk	
	Grindowanie			Indykowate	
	Grindownik			Indykowe	
	Grindowy				

Nazwa pola	Nazwa formy podstawowej	Złożoność pola	Nazwa pola	Nazwa formy podstawowej	Złożoność pola
JRPG		4	Multiplatformowy		5
	JRPG			Multiplatformowy	
	JRPG'owiec			Multiplatformowiec	
	JRPG-owość			Multiplatformowo	
	JRPG-owy			Multiplatformowość	
Kinect		2		Multiplatformówka	
	Kinect		Next-gen		5
	Kinectowy			Next-gen	
Klatkować		5		Next-genowiec	
	Klatkować			Next-genowo	
	Klatkowanie			Next-genowość	
	Klatkowany			Next-genowy	
	Klatkowy		Off-road		6
	Klatkujący			Offroad	
Klawiatura		5		Off-road	
	Klawiatura			Offroadowy	
	Klawiaturka			Off-roadowy	
	Klawiaturka			Offroadówka	
	Klawiaturowy			Ofrold	
	Klawierka		Ośmiobitowy		2
Komputer		9		Ośmiobitowiec	
	Komp			Ośmiobitowy	
	Kompek		Padlet		2
	Kompik			Padlet	
	Komputer			Tableto-pad	
	Komputerek		Past-gen		2
	Komputerowiec			Past-gen	
	Komputerowo			Past-genowy	
	Komputerowy		Patch		20
	Komputeryzacja			Patch	
Konsola		6		Patcher	
	Konsola			Patchować	
	Konsolka			Patchowanie	
	Konsolowiec			Patchowany	
	Konsolowo			Patchując	
	Konsolowość			Patchujący	
	Konsolowy			Pacz	
Liniowy		3		Paczer	
	Liniowo			Paczowanie	
	Liniowość			Paczowany	
	Liniowy			Paczyk	
Live-stream		3		Spaczować	
	Live-stream			Spaczowanie	
	Live-streamowany			Spaczowany	
	Live-streamowy			Spatchować	
Loading		3		Spatchowanie	
	Loading			Spatchowany	
	Loadingować			Łatka	
	Loadingowy			Łatkowanie	
Marian		3	Pauza		6
	Marian			Pauza	
	Marianek			Pauzować	
	Marianowy			Pauzowanie	
Mikrotransakcja		2		Pauzowany	
	Mikrotransakcja			Pauzując	
	Mikrotransakcyjny			Pauzujący	
Minimapa		2	PC		20
	Minimapa			PC	
	Minimapka			PC-t	
MMORPG		4		PC-towiec	
	MMO RPG			PC-towy	
	MMORPG			Pececiarz	
	MMO-RPG			Pececik	
	MMORPG-owy			Pecet	
Multi		6		Pecetowiec	
	Multi			Pecetowo	
	Multiplayer			Pecetowość	
	Multiplayerowiec			Pecetowy	
	Multiplayerowość			Piec	
	Multiplayerowy			Piecet	
	Multiplayerówka			Piecowiec	
Multimedia		4		Piecowy	
	Multimedia			Piecyk	
	Multimedialnie			Piecykowy	
	Multimedialność			Błaszak	
	Multimedialny			Błaszakowiec	
				Błaszakowy	

Nazwa pola	Nazwa formy podstawowej	Złożoność pola	Nazwa pola	Nazwa formy podstawowej	Złożoność pola
Pesczwórka		4	PSP go		2
	Peeszttery			PSP go	
	Peesczwórka			PSPgo	
	Peesczwórkowiec		Quest		8
	Peesczwórkowy			Quest	
Peesdwójka		3		Questing	
	Peesdwójka			Questować	
	Peesdwójkowy			Questowanie	
	Pees dwójka			Questowo	
Peesiks		3		Questowy	
	Peesiks			Questując	
	Szarak			Questujący	
	Szarakowy		Randomizacja		3
Peesplusowy		3		Randomizacja	
	Peesplusowicz			Randomizować	
	Peesplusowiec			Randomizowany	
	Peesplusowy		Real-time		2
Peestrójka		6		Real-time	
	Peestrójkowiec			Real-time'owy	
	Peestrójka		Redzi		3
	Peestrójkowicz			Redowski	
	Peestrójkowy			Redowy	
	Pees-trójkowy			Redzi	
	Chlebak		Remake		4
Pełniak		3		Remake	
	Pełniak			Remakeowanie	
	Pełniakowo			Rimejk	
	Pełniakowy			Rimejkowanie	
Platyna		7	Remaster		9
	Platyna			Remaster	
	Platynka			Remasterek	
	Platynować			Remastering	
	Platynowanie			Remasterowanie	
	Platynowany			Remasterowany	
	Platynowi			Remasterując	
	Platynowy			Zremasterować	
PlayStation		16		Zremasterowanie	
	PlayStation			Zremasterowany	
	Playstationowiec		Roguelike		4
	Playstationowy			Rogalik	
	Plej			Rogalikowy	
	Plejacz			Roguelike	
	Plejak			Roguelike'owy	
	Plejakować		Rozdzielczość		2
	Plejeczka			Rozdziałka	
	Plejka			Rozdzielczość	
	Plejowy		RPG		10
	Plejtacja			RPG	
	Plejstejszon			RPG-owanie	
	Plejstejszonowy			RPG-owiec	
	Plejstejszyn			RPG-owy	
	Peeska			Erpeg	
	Peesowy			Erpegować	
Pre-alpha		2		Erpegowanie	
	Pre-alfa			Erpegowiec	
	Pre-alpha			Erpegowość	
Pre-oreder		4		Erpegowy	
	Pre-order		Samograj		4
	Pre-orderowicz			Samograj	
	Pre-orderowiec			Samogrające	
	Pre-orderowy			Samograjka	
Przeglądarka		3		Samograjowy	
	Przeglądarka		Sandbox		11
	Przeglądarkowy			Sandboks	
	Przeglądarkówka			Sandboksik	
Przygodówka		2		Sandboksowanie	
	Przygodówka			Sandboksowość	
	Przygodówkowy			Sandboksowy	
PSP		4		Sandbox	
	PSP			Sandboxowo	
	PSP'owiec			Sandboxowość	
	Peespek			Sandboxowy	
	Peespekowy			Piaskownica	
PSP2		2		Piaskownicowy	
	PSP 2				
	PSP2				

Nazwa pola	Nazwa formy podstawowej	Złożoność pola	Nazwa pola	Nazwa formy podstawowej	Złożoność pola
Save		11		Streamingujący	
	Save			Streamować	
	Save'ować			Streamowanie	
	Save'owanie			Streamowany	
	Savepoint			Streamowy	
	Save-point			Streamując	
	Sejw			Streamujący	
	Sejwować			Strzelanina	6
	Sejwowanie			Strzelanina	
	Sejwujący			Strzelanka	
	Zasejwować			Strzelankowy	
Zasejwowanie		Shooter			
		Shooterek			
Sequel	Sequel	3		Shooterowy	
	Sequelizm			Szesnastobitowy	2
	Sequelowy			Szesnastobitowiec	
Shmup		2		Szesnastobitowy	
	Shmup			Ścigałka	3
	Shmupowy		Ścigała		
Sieciówka		2		Ścigałka	
	Sieciówka			Ścigałkowy	
	Sieciówkowy			Tablet	5
Singiel		10		Tablecik	
	Singiel			Tablet	
	Singleplajer			Tabletek	
	Singleplayer			Tabletkowy	
	Single-player			Tabletowy	
	Singleplayerowy			Tank	7
	Single-playerowy			Tank	
	Singlowiec			Tanker	
	Singlowo			Tankista	
	Singlowość			Tankować	
Singlowy		Tankowanie			
Skin		2		Tankowaty	
	Skin			Tankujący	
	Skórka		Tekstówka	2	
Skrypt		5			Tekstówka
	Skrypt			Tekstówkowy	
	Skryptować			Tekstura	10
	Skryptowanie			Tekstura	
	Skryptowany			Teksturka	
Skryptowy		Teksturować			
Smartfon		2		Teksturowanie	
	Smartfon			Teksturowany	
	Smartfonowy			Teksturowy	
Software		3		Teksturujący	
	Software			Oteksturować	
	Software'owo			Oteksturowanie	
	Software'owy			Oteksturowany	
Spawn		17		Touchpad	4
	Spawn			Touch pad	
	Spawner			Touchpad	
	Spawnowanie			Myziadelko	
	Spawnowany			Gładzik	
	Spawnujący			Trailer	3
	Spwanować			Trailer	
	Zrespawnić			Trailerek	
	Zrespawnować			Trailerowy	
	Zrespawnowanie			Uzywka	2
	Zrespawnowany			Używka	
	Respawn			Używkowy	
	Respawner			WSAD	3
	Respawnować			WASD	
	Respawnowanie			WSAD	
	Respawnowy			WSAD-owy	
	Respawnując			Xbox	8
Respawnujący		Xboks			
Sprite		2		Xboksiak	
	Sprajt			Xboksowiec	
	Sprite			Xboksowy	
Stream		13		Xbox	
	Stream			Xkloc	
	Streamer			Xklocek	
	Streamerka			Xklockowy	
	Streamerski			Zapychacz	2
	Streaming			Zapychacz	
	Streamingowy			Zapychaczowy	
				Zapychaczowy	

Nazwa pola	Nazwa formy podstawowej	Złożoność pola	Nazwa pola	Nazwa formy podstawowej	Złożoność pola
Bumper		1	Przedłużajka		1
	Bumper			Przedłużajka	
Cooldown		1	Przejściówka		1
	Cooldown			Przejściówka	
Creditsy		1	Przeszkadzajka		1
	Creditsy			Przeszkadzajka	
Cross-buy		1	Przetrwaniówka		1
	Cross-buy			Przetrwaniówka	
Cross-chat		1	PvE		1
	Cross-chat			PvE	
Cross-play		1	PvP		1
	Cross-play			PvP	
Cross-save		1	QTE		1
	Cross-save			QTE	
Devkit		1	QWERTY		1
	Devkit			QWERTY	
DLC		1	Reedycja		1
	DLC			Reedycja	
Dokanałówki		1	Router		1
	Dokanałówki			Router	
Eksbokslan		1	RROD		1
	Eksbokslan			RROD	
Flashówka		1	Samochodówka		1
	Flashówka			Samochodówka	
Framerate		1	Shader		1
	Framerate			Shader	
GotY		1	Side-quest		1
	GotY			Side-quest	
Grzybek		1	Soundtrack		1
	Grzybek			Soundtrack	
Jednostrzałowiec		1	Stacjonarka		1
	Jednostrzałowiec			Stacjonarka	
Kartridż		1	Stodwudziestośmiobitowie		1
	Kartridż			Stodwudziestośmiobitowie	
Kontroler		1	Stylus		1
	Kontroler			Stylus	
Konwersja		1	Sub-boss		1
	Konwersja			Sub-boss	
Lokalizacja		1	Sub-quest		1
	Lokalizacja			Sub-quest	
LPM		1	Subskrypcja		1
	LPM			Subskrypcja	
Matryca		1	Three-sixty		1
	Matryca			Three-sixty	
Metagra		1	Trigger		1
	Metagra			Trigger	
Multikart		1	Turówka		1
	Multikart			Turówka	
Multitap		1	Twarzowanie		1
	Multitap			Twarzowanie	
Multitasking		1	Unboxing		1
	Multitasking			Unboxing	
Multitouch		1	Wiesiek		1
	Multitouch			Wiesiek	
Must-have		1	Wtyczka		1
	Must-have			Wtyczka	
Niedzielnik		1	YLOD		1
	Niedzielnik			YLOD	
Notebook		1	Zblurowany		1
	Notebook			Zblurowany	
NPC		1	Znajdźka		1
	NPC			Znajdźka	
Odblokowajka		1	Żywołność		1
	Odblokowajka			Żywołność	
One-liner		1			
	One-liner				
Pad		1			
	Pad				
Pay-to-win		1			
	Pay-to-win				
Podstawka		1			
	Podstawka				
PPM		1			
	PPM				

Dzięki zastosowaniu wyszukiwarki, a później wykonaniu odpowiednich działań matematycznych, uzyskałem dokładne dane liczbowe na temat wszystkich form podstawowych oraz sumy wystąpień ich wszystkich wyrazów graficznych w poszczególnych częściach korpusach.

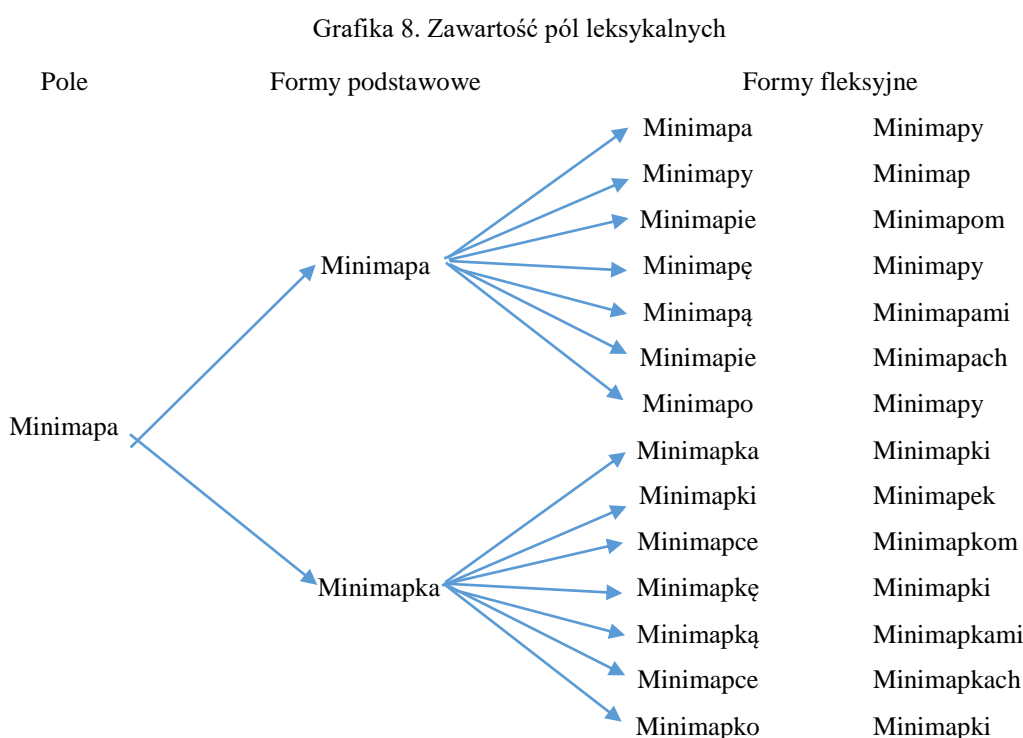
Powyższa tabela prezentuje ostatni etap pracy, czyli połączenie wszystkich form podstawowych w pola wyrazowe na tej zasadzie, że

- połączyłem wersje graficzne tych samych pojęć (np. *save*, *sejw*)
- połączyłem ekwiwalenty leksykalne tego samego pojęcia (np. *komputer*, *komp*)

Powyższa tabela zbudowana została z trzech zasadniczych części: *nazwa pola*, *nazwa formy podstawowej* i *złożoność pola*. Na jednej stronie dodałem po dwa zestawy tych części.

W części *złożoność pola* umieściłem liczbę, która informuje, ile form podstawowych zawiera dane pole (np. pole *action RPG* ma tych form 4: *action RPG*, *action-RPG*, *RPG akcji*, *RPG-akcji*). W części *nazwa formy podstawowej* wypisałem wszystkie formy podstawowe danych pól (łącznie w całym badaniu jest ich 709). W części *nazwa pola* podałem wszystkie nazwy pól (189), które stanowią trzon dalszej części badań. Nazwę całego pola stanowi jedna z form podstawowych, która w porównaniu z innymi formami danego pola wystąpiła najliczniej w całym badanym korpusie.

Zatem jedno pole zawiera w sobie różne formy podstawowe, natomiast jedna forma podstawowa wszystkie jej realizacje tekstowe – wyrazy graficzne (np. realizacje różnych form fleksyjnych).



Wyjaśnienia wymaga jeszcze jeden element tabeli 16, czyli pola przedstawione na stronie 125.

Zdaję sobie sprawę z tego, że używanie pojęcia pole leksykalne w odniesieniu do jednoelementowego zbioru jest być może niezręczne, ale jest to działanie świadome, gdyż zakładam, że zbiory uległyby rozbudowaniu podczas analizy kolejnych próbek materiału, które zapewne pokazałyby, że zawierają one więcej elementów*.

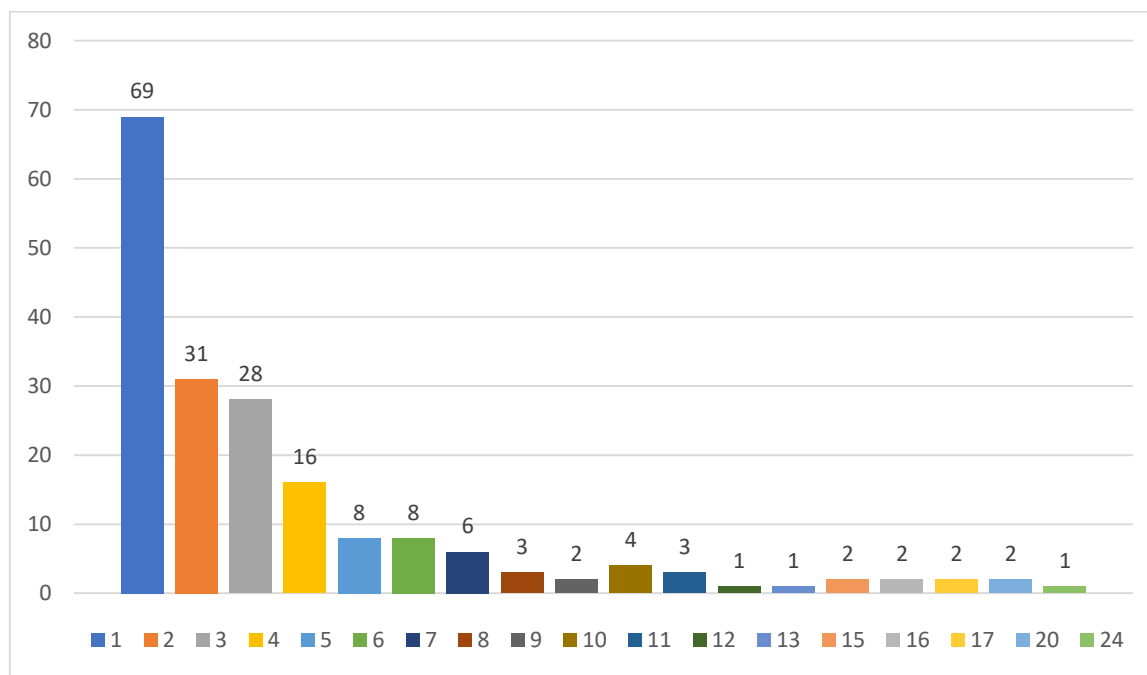
Reasumując, powyższa tabela (16) jest wynikiem ekscerpcji, porządkowania i poszerzenia zebranego materiału, który jest kluczowy dla badań. Przy czym trzeba zaznaczyć, że jest to wersja maksymalna tabeli (uwzględnia nazwy pól oraz nazwy form podstawowych). W dalszej części pracy zazwyczaj będę odnosił się do jej skróconego odpowiednika – tabeli zawierającej wyłącznie nazwy pól. Z faktem tym wiązała się konieczność wyjaśnienia, co kryje się pod taką nazwą pola, która jest reprezentantem całego zbioru wyrazowego.

* W innych naukach jest to normalna procedura, np. w matematyce istnieje nie tylko pojęcie zbioru jednoelementowego (jak w mojej pracy), ale nawet zbioru pustego. Terminy te używane są zatem w naukach ścisłych, a moja praca łączy językoznawstwo i metodologię badań ilościowych, w której często wykorzystuje się różnego rodzaju terminologię matematyczną.

CZEŚĆ
ANALITYCZNA

7. POLA LEKSYKALNE A ICH ZŁOŻONOŚĆ

Na początku części analitycznej warto przyrzeć się złożoności pól (189), biorąc pod uwagę wszystkie ich formy podstawowe (709). Jak już nadmieniałem, powstały one dzięki połączeniu w jednym polu wariantów zapisu (będących skutkiem różnic uzualnych i niestabilizowania normy), form pochodnych słowotwórczo, czyli rodzin wyrazów potwierdzonych w korpusie oraz połączenia ekwiwalentów znaczeniowych (kiedy np. 2 nazwy odnoszą się do tego samego desygnatu). W poniższym zestawieniu podano, ile jest procent pól cechujących się taką samą złożonością, czyli mających po tyle samo form podstawowych.

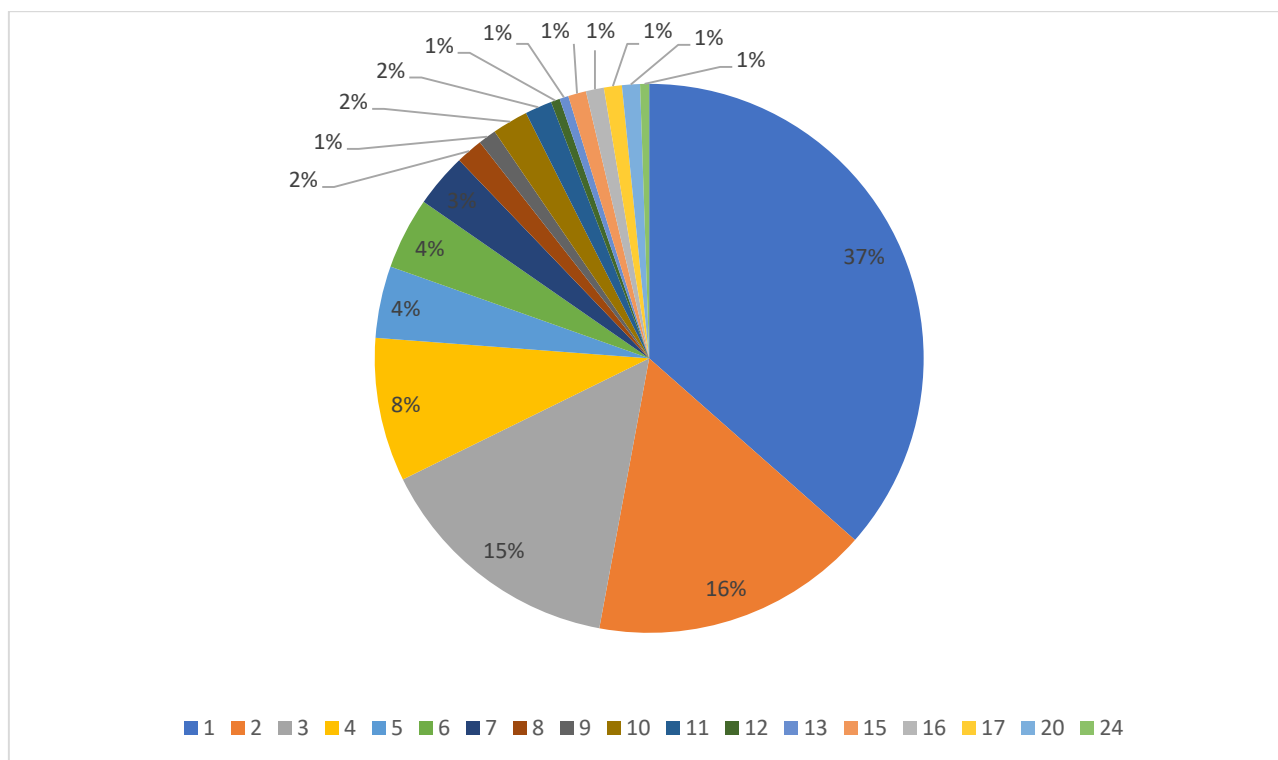


Wykres 1. Złożoność pól wyrazowych (ujęcie liczbowe)

Na wykresie zobrazowałem to, jak bardzo rozbudowane są pola, czyli z ilu form podstawowych się składają. Na osi x umieszczone zostały grupy pól, które zawierają jednakową liczbę form podstawowych (jednakowa złożoność), natomiast na osi y pokazałem, ile jest danych grup. Ponadto liczba pól z jednakową złożonością oznaczona została danym kolorem, a znaczenie kolorów podano w legendzie na dole wykresu.

Po uporządkowaniu pól pod względem ich złożoności, okazało się, że najwięcej jest pól jednoelementowych (patrz powyższy wykres), bo aż 69, natomiast najmniej pól trzynasto- i dwudziestoczeroelementowych, bo tylko po 1.

Powyżej skupiłem się na liczbowym zaprezentowaniu złożoności pól, warto też przyrzeć się podziałowi procentowemu:

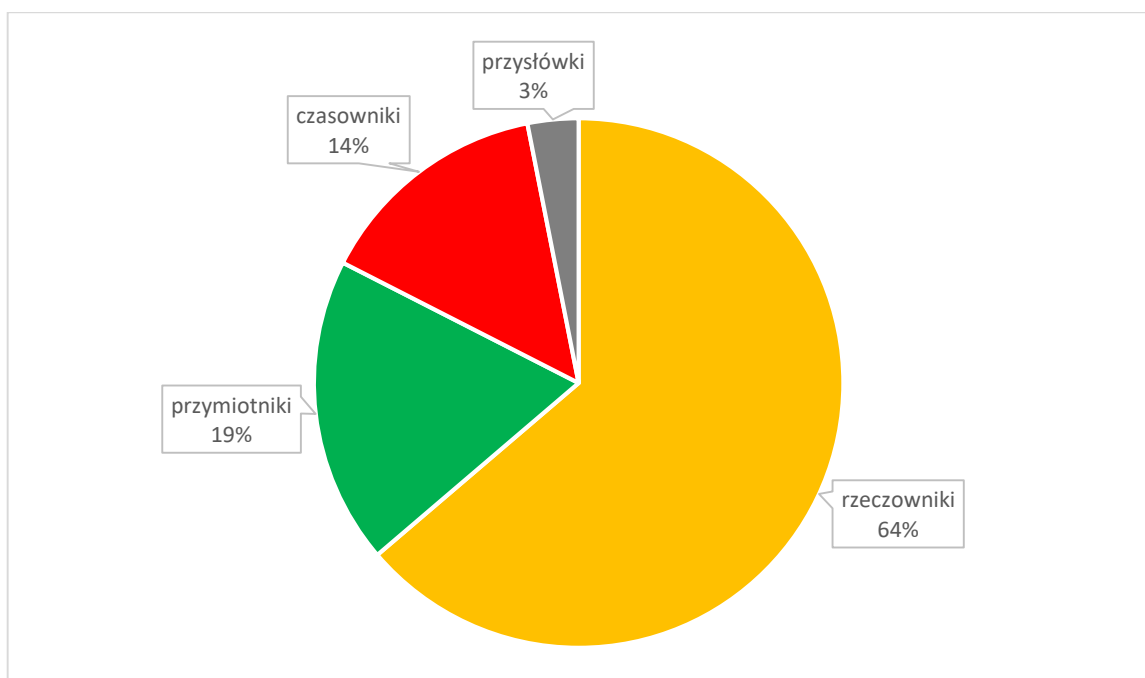


Wykres 2. Złożoność pól wyrazowych (ujęcie procentowe)⁷¹

Pola jednoelementowe stanowią aż 37% wszystkich pól. Są to pola, które zawierają głównie tylko rzeczowniki. Rozpatrując natomiast łączną liczbę wystąpień pól jedno- i dwuelementowych, to stanowią one już większość całego zbioru, bo 53%. Pola dwuelementowe zawierają głównie rzeczownik i przymiotnik od niego powstały lub dwie formy zapisu danego rzeczownika.

Analizując tabelę 16 pod kątem części mowy, ale wyłącznie w odniesieniu do form podstawowych, które budowały poszczególne pola leksykalne, zaobserwować można zdecydowaną przewagę rzeczowników. Dokładne dane na temat liczebności części mowy przedstawiłem na poniższym wykresie. W tym zestawieniu wzięto pod uwagę 709 form.

⁷¹ Po zsumowaniu wszystkich liczb z części właściwej wykresu uzyskuje się wynik 101%. Powstał on, ponieważ program Word automatycznie zaokrągliła wszystkie wprowadzane dane. Jednak wykresy ten służy jedynie poglądowemu przedstawieniu procentowego udziału danych pól, bo dokładne dane znajdują się na wykresie 1.



Wykres 3. Formy podstawowe a części mowy

Podczas analizy form podstawowych odnotowałem najwięcej rzeczowników, których jest aż 452. Na kolejnym miejscu są przymiotniki (133), następnie czasowniki (102), a na ostatnim – czwartym miejscu – przysłówki (22).

W grupie czasowników uwzględniłem zarówno bezokoliczniki, jak i imiesłowy. Gdyby je rozdzielić, to imiesłowów byłaby zdecydowana większość, bo aż 60, natomiast bezokoliczników tylko 42.

Ten krótki rozdział miał na celu przedstawienie liczebności form podstawowych danych pól wyrazowych, a ponadto zbadanie złożoności pól wyrazowych oraz wskazanie, jakich części mowy było najwięcej w moich badaniach. Do tematu złożoności pól oraz liczebności form podstawowych będę jeszcze wracał w dalszej części pracy.

8. KORPUSY TEKSTOWE A LICZBA WSZYSTKICH WYRAZÓW GRAFICZNYCH

W tym rozdziale skupię się na liczebności wyrazów graficznych w poszczególnych częściach korpusu tekstów. W tabeli poniżej zaprezentowałem dokładną liczbę wszystkich wyrazów graficznych (zarówno z leksyki ogólnej, jak i środowiskowej), które włączyłem do badań. W poniższej prezentacji nie ograniczyłem się jedynie do podania sumy wszystkich wyrazów graficznych w korpusach, ale pokazałem wersję rozszerzoną, czyli z podziałem na kwartały, oczywiście tylko tam, gdzie było to możliwe, czyli w cda, npsx, ncda i npsx. Danych kwartalnych w odniesieniu do forów internetowych nie mogłem podać, a wynikało to z powodów, o których pisałem na stronie 98.

	CDA	PSX	NCDA	NPSX
1 kwartał 2011	213.796	233.173	580.488	115.779
2 kwartał 2011	209.769	214.770	552.181	165.676
3 kwartał 2011	214.669	223.248	566.036	221.434
4 kwartał 2011	215.591	220.420	573.272	227.261
1 kwartał 2012	226.891	244.719	598.267	271.402
2 kwartał 2012	226.476	240.336	568.149	220.002
3 kwartał 2012	227.679	259.779	568.741	166.876
4 kwartał 2012	232.267	260.470	590.189	225.754
1 kwartał 2013	230.631	240.380	582.889	161.670
2 kwartał 2013	223.739	248.474	555.814	276.962
3 kwartał 2013	224.302	240.215	558.436	106.434
4 kwartał 2013	220.713	233.117	587.675	90.387
1 kwartał 2014	235.679	237.730	555.237	107.873
2 kwartał 2014	228.988	250.850	577.713	147.106
3 kwartał 2014	235.437	225.763	553.295	116.747
4 kwartał 2014	240.211	212.789	599.253	114.493
1 kwartał 2015	253.936	222.546	622.227	132.101
2 kwartał 2015	260.900	214.617	580.202	113.123
3 kwartał 2015	255.959	207.323	560.475	201.365
4 kwartał 2015	241.696	200.105	573.749	292.158
suma	4.619.329	4.630.824	11.504.288	3.474.603

FCDA	FPSX
19.892.423	19.892.385

Tabela 17. Suma wszystkich wyrazów graficznych w poszczególnych częściach korpusu

Pierwszą ciekawą informacją, jaką można odczytać z powyższego zestawienia jest to, że na przestrzeni wszystkich kwartałów, sumy wyrazów graficznych użytych w czasopiśmie

„CD-Action” (a także suma końcowa – 4.619.329) są bardzo zbliżone do sum wyrazów w czasopiśmie „PSX Extreme” (także do sumy końcowej – 4.630.824). Fakt ten może dziwić, ponieważ „CD-Action” w wersji drukowanej zawiera o około 30 stron więcej niż czasopismo „PSX Extreme”. Wywnioskować więc można, że „CD-Action” na tych stronach umieszcza po prostu więcej grafik lub zdjęć niż „PSX Extreme”, natomiast samego tekstu jest podobna ilość.

Kolejną wartą odnotowania informacją jest to, że analizując dane z wykresu zauważyć można dużą różnicę w łącznej liczbie wyrazów graficznych między newsami pozyskanymi ze strony www.ppe.pl (3.474.603), a tymi z www.cdaction.pl (11.504.288). Wyniki te mogą dziwić, ponieważ do badań wybrałem po 300 newsów z jednego kwartału dla każdej ze stron. Aby zrównać tę różnicę, należałoby dołożyć kolejne kilkaset newsów opublikowanych na stronie ppe.pl, niestety nie mogłem tego zrobić, ponieważ nie dysponowałem aż tyloma dodatkowymi tekstami.

Rozwiązaniem zaistniałej sytuacji było statystyczne, czysto matematyczne, zrównanie liczby wyrazów poprzez zastosowanie odpowiednich mnożników⁷². I w tym przypadku także mógłbym się ograniczyć wyłącznie do zrównania sumy wyrazów ze wszystkich kwartałów i zastosować jeden mnożnik. Chcąc jednak zachować dokładniejsze wyniki podczas dalszej pracy, w której będę wykorzystywał różne mnożniki, postanowiłem stworzyć po jednym mnożniku dla danego kwartału. Inaczej mówiąc, zastosowanie jednego, ogólnego mnożnika byłoby zbyt dużym uproszczeniem.

Mnożniki powstały przez podzielenie liczby wyrazów z danego kwartału w zbiorze ncda przez liczbę wyrazów z tego samego kwartału, ale w zbiorze npsx. Przykładowo, dla 1 kwartału 2011 roku mnożnik powstał w wyniku podzielenia liczby 580.488 przez 115.779. W efekcie tego działania otrzymałem liczbę 5,013759, więc po zaokrągleniu mnożnik przybrał formę cyfry 5. Zaokrąglenia dokonywałem do części połowicznej, czyli gdy wynik działania dla danego kwartału wynosił np. 3,468043, to w tabeli umieszczałem wynik 3,5.

Wszystkie mnożniki wyglądają następująco:

	NPSX (przed)		Mnożnik		NPSX (po)	NCDA
1 kwartał 2011	115.779		5		578.895	580.488
2 kwartał 2011	165.676		3,5		579.866	552.181
3 kwartał 2011	221.434		2,5		553.585	566.036
4 kwartał 2011	227.261		2,5		568.153	573.272

⁷² Mnożniki są stosowane w celu wyrównania pewnych danych np. w ekonomii (patrz artykuł R. Kala i L. Błażejczyk-Majki [2010]), matematyce czy ornitologii (patrz artykuł L. Tomiałojcia [1970]). W pracach językoznawczych nie korzysta się z nich, co jest moim zdaniem błędem.

1 kwartał 2012	271.402		2		542.804	598.267
2 kwartał 2012	220.002		2,5		550.005	568.149
3 kwartał 2012	166.876		3,5		584.066	568.741
4 kwartał 2012	225.754		2,5		564.385	590.189
1 kwartał 2013	161.670		3,5		565.845	582.889
2 kwartał 2013	276.962		2		553.924	555.814
3 kwartał 2013	106.434		5		532.170	558.436
4 kwartał 2013	90.387		6,5		587.516	587.675
1 kwartał 2014	107.873		5		539.365	555.237
2 kwartał 2014	147.106		4		588.424	577.713
3 kwartał 2014	116.747		4,5		525.362	553.295
4 kwartał 2014	114.493		5		572.465	599.253
1 kwartał 2015	132.101		4,5		594.455	622.227
2 kwartał 2015	113.123		5		565.615	580.202
3 kwartał 2015	201.365		3		604.095	560.475
4 kwartał 2015	292.158		2		584.316	573.749
suma	3.474.603				11.335.311	11.504.288

Tabela 18. Suma wszystkich wyrazów graficznych w korpusie NPSX przed i po zastosowaniu mnożników

W powyższej tabeli poza samymi mnożnikami pokazałem także liczbę wyrazów danego kwartału, odnoszącą się do realnie zgromadzonych przeze mnie tekstów newsów, czyli przed zastosowaniem mnożników (patrz kolumna *NPSX (przed)*), a także liczbę wyrazów, która powstała w wyniku przemnożenia wszystkich danych kwartalnych korpusu npsx (patrz kolumna *NPSX (po)*). Dodatkowo, dla lepszego zwizualizowania wyników, dodałem jeszcze jedną kolumnę – *NCDA*, która zawiera dane kwartalne w odniesieniu do korpusu ncda.

Zbiory tekstów z forów internetowych, które pozyskałem do stworzenia korpusów, zostały tak zaplanowane, żeby ostatecznie fpsx i fcda zawierały podobną liczbę wyrazów, czyli po $\approx 1,99$ mln wyrazów. Tutaj, podobnie jak w przypadku czasopism, nie było konieczności stosowania żadnych mnożników, które zrównałyby liczbę wyrazów w obu korpusach tekstowych.

Powyżej opisałem sytuację zrównywania, jeśli oczywiście była taka konieczność, liczby wyrazów z dwóch korpusów, które odnosiły się do tego samego miejsca pozyskiwania materiałów – czasopisma, newsy lub fora. Poza tego typu zrównywaniem w pracy zastosowałem jeszcze jedno, mianowicie takie, które wszystkie sześć korpusów (cda, psx, ncda, npsx, fcda i fpsx) wyrównało do jednej zbliżonej sumy wyrazów. Powyżej zrównywałem liczbę wyrazów w korpusie npsx do góry (do liczby wyrazów w korpusie ncda), więc zdecydowałem,

że dane umieszczone w korpusach cda, ncda, psx i npsx zrównane zostaną do liczby wyrazów w korpusach forów internetowych (fpsx i fcda).

Aby to uczynić zastosowałem następujące mnożniki:

	CDA (przed)	PSX (przed)	NCDA (przed)	NPSX (przed)	M. CDA	M. PSX	M. NCDA	M. NPSX	CDA (po)	PSX (po)	NCDA (po)	NPSX (po)
1k. 11	213.796	233.173	580.488	115.779	4,5	4,5	1,5	8,5	962.082	1.049.279	870.732	984.122
2k. 11	209.769	214.770	552.181	165.676	4,5	4,5	2	6	943.961	966.465	1.104.362	994.056
3k. 11	214.669	223.248	566.036	221.434	4,5	4,5	2	4,5	966.011	1.004.616	1.132.072	996.453
4k. 11	215.591	220.420	573.272	227.261	4,5	4,5	1,5	4,5	970.160	991.890	859.908	1.022.675
1k. 12	226.891	244.719	598.267	271.402	4,5	4	1,5	3,5	1.021.010	978.876	897.401	949.907
2k. 12	226.476	240.336	568.149	220.002	4,5	4	2	4,5	1.019.142	961.344	1.136.298	990.009
3k. 12	227.679	259.779	568.741	166.876	4,5	4	1,5	6	1.024.556	1.039.116	853.112	1.001.256
4k. 12	232.267	260.470	590.189	225.754	4,5	4	1,5	4,5	1.045.202	1.041.880	885.284	1.015.893
1k. 13	230.631	240.380	582.889	161.670	4,5	4	1,5	6	1.037.840	961.520	874.334	970.020
2k. 13	223.739	248.474	555.814	276.962	4,5	4	2	3,5	1.006.826	993.896	1.111.628	969.367
3k. 13	224.302	240.215	558.436	106.434	4,5	4	2	9,5	1.009.359	960.860	1.116.872	1.011.123
4k. 13	220.713	233.117	587.675	90.387	4,5	4,5	1,5	11	993.209	1.049.027	881.513	994.257
1k. 14	235.679	237.730	555.237	107.873	4	4	2	9	942.716	950.920	1.110.474	970.857
2k. 14	228.988	250.850	577.713	147.106	4,5	4	1,5	7	1.030.446	1.003.400	866.570	1.029.742
3k. 14	235.437	225.763	553.295	116.747	4	4,5	2	8,5	941.748	1.015.934	1.106.590	992.350
4k. 14	240.211	212.789	599.253	114.493	4	4,5	1,5	8,5	960.844	957.551	898.880	973.191
1k. 15	253.936	222.546	622.227	132.101	4	4,5	1,5	7,5	1.015.744	1.001.457	933.341	990.758
2k. 15	260.900	214.617	580.202	113.123	4	4,5	1,5	9	1.043.600	965.777	870.303	1.018.107
3k. 15	255.959	207.323	560.475	201.365	4	5	2	5	1.023.836	1.036.615	1.120.950	1.006.825
4k. 15	241.696	200.105	573.749	292.158	4	5	1,5	3,5	966.784	1.000.525	860.624	1.022.553
suma	4.619.329	4.630.824	11.504.288	3.474.603					19.925.076	19.930.948	19.491.248	19.903.521

FCDA	FPSX
19.892.423	19.892.385

Tabela 19. Suma wszystkich wyrazów graficznych we wszystkich częściach korpusu przed i po zastosowaniu mnożników

Stosując przedstawione w powyższej tabeli mnożniki, które zaprezentowane zostały w kolumnach *M. CDA*, *M. PSX*, *M. NCDA* i *M. NPSX* chciałem, aby dane z poszczególnych kwartałów zbliżone zostały do liczby 994.620. Powstała ona przez wyciągnięcie średniej z ogólnej liczby wyrazów na forach internetowych $[(19.892.423 + 19.892.385) : 2]$, a następnie podzieleniu uzyskanego wyniku przez liczbę kwartałów, czyli w tym przypadku przez 20 $(19.892.404 : 20)$.

Podsumowując, w moich badaniach na temat liczebności konkretnych wyrazów lub ich zbiorów, wyniki będę rozpatrywał na trzech płaszczyznach, z których dwie bezpośrednio łączą się ze stosowaniem, zademonstrowanych w powyższej części pracy, mnożników.

Pierwszą płaszczyznę, na tle której zostaną zaprezentowane wyniki badań ilościowych, stanowią korpusy tekstów pozyskane przeze mnie w sposób, który opisałem w rozdziałach metodologicznych, czyli bez stosowania żadnych mnożników. Innymi słowy wyniki podawane będą w odniesieniu do realnie utworzonego korpusu, który liczy łącznie 64.013.852 wyrazy (cda – 4.619.329, psx – 4.630.824, ncda – 11.504.288, npsx – 3.474.603, fcda – 19.892.423 i fpsx – 19.892.385). Dla ułatwienia, płaszczyznę tę nazwałem zieloną, a co się z tym wiąże, wszystkie wykresy, tabele, zestawienia będą posiadały zieloną kolorystykę tła lub ramki.

Drugą płaszczyznę stanowią dwa korpusy tekstów (npsx i ncda), gdzie w jednym zastosowałem mnożniki danych (ten umieszczone w tabeli na stronie 139). Wyrównałem w ten sposób liczbę wyrazów w korpusach npsx i ncda. Korpus wyrazów tej płaszczyzny liczy łącznie 22.839.599 jednostek (npsx – 11.335.311, ncda – 11.504.288). Płaszczyzna ta została stworzona, aby pokazać zależności lub porównać wyniki między dwoma zbiorami newsów. Dążyłem więc do tego, aby miały one zbliżoną liczbę wyrazów w korpusach. Płaszczyznę tę nazwałem żółtą i podobnie jak w przypadku płaszczyzny zielonej, podczas prezentacji wyników będę używał jednej barwy, tym razem będzie to kolor żółty.

Ostatnia, trzecia płaszczyzna, to wyniki badań, które zaprezentowane zostaną na tle jednego, dużego korpusu, który powstał przez wykorzystanie czterech zestawów mnożników (pokazałem je na stronie poprzedniej). W wyniku mnożenia, zrównywane zostały liczby wyrazów w korpusach cda, ncda, psx i npsx ze średnią liczbą wyrazów na forach internetowych. Płaszczyzna ta liczy 119.035.598 jednostek (cda – 19.925.076, psx – 19.930.948, ncda – 19.491.248, npsx – 19.903.521, fcda – 19.892.423 i fpsx – 19.892.385). Wszystkie korpusy są sobie równe pod względem liczby wyrazów. Barwa przypisana tej platformie to czerwień.

W rzeczywistości, podczas wyszukiwania danych jednostek wyrazowych, pracowałem tylko w odniesieniu do korpusu, który obecnie stanowi platformę zieloną, czyli bez żadnych zaokrągleń. Jednakże trudno byłoby obiektywnie zestawić dane z mniejszych korpusów, które zawierają nierówną bazę wyrazową, dlatego pojawiła się konieczność stworzenia innych platform. To właśnie zestawianie danych z różnych korpusów wymusiło ich powstanie. Należy jednak pamiętać, że powstały one w wyniku pewnych działań matematycznych, więc wyniki odnoszące się do nich będą efektem przybliżenia statystycznego. W tej pracy najczęściej jednak wykorzystuję platformę zieloną, a najrzadziej czerwoną.

W pracy bardzo często podaję wyniki dla płaszczyzny zielonej, a dopiero później zestawiam ją z inną. W tym przypadku, ta *inna* pełni swoistą funkcję kontrolną, ponieważ jeśli jakiś badany aspekt językowy został wykazany na platformie zielonej i na czerwonej znajdzie także swoje odzwierciedlenie, to z całą pewnością można stwierdzić, że ma on istotne znaczenie.

9. WYRAZY GRAFICZNE ŚRODOWISKA GRACZY A PLATFORMY BADAWCZE

W stworzonych korpusach, w wyniku wyszukiwania tekstowych wystąpień 709 form podstawowych (skomasowanych w 189 polach wyrazowych – mniej lub bardziej licznych) odnotowałem łącznie 434.924 wyrazy graficzne polskiego socjolektu graczy wideo, które włączone zostały do badań ilościowych⁷³.

Rzeczywistą sumę wyrazów graczy w poszczególnych częściach korpusu prezentuje poniższa tabelka:

Suma wyrazów	
CDA	50.946
NCDA	90.506
FCDA	82.940
PSX	47.338
NPSX	68.046
FPSX	95.148

Tabela 20. Suma wyrazów graficznych graczy w poszczególnych częściach korpusu (platforma zielona)

Należy jednak pamiętać, że wyrazy te zostały pozyskane z 6 podkorpusów o różnej liczebności. Te liczby wyrazów ogólnych z sześciu korpusów stanowią jak gdyby matematyczny mianownik operacji porównania. Żeby porównanie to było bardziej miarodajne, należy sprowadzić korpusy do zbliżonej liczebności, a następnie powiększyć o odpowiednie mnożniki sumę wyrazów specjalistycznych.

Dane z powyżej zamieszczonej tabeli zostały więc zaokrąglone i wyrównane do wspólnego mianownika⁷⁴ W wyniku tej operacji wyrazy graczy w powiększonych korpusach uległy zwielokrotnieniu, a ich sumy wyglądają następująco:

Suma wyrazów	
CDA	220.377
NCDA	154.317

⁷³ Aby nie używać tak długiej nazwy, w dalszej części pracy często będę tę grupę wyrazów nazywał *wyrazami graficznymi graczy* lub *wyrazami graczy*.

⁷⁴ Zdecydowano się na ten krok ze względu na konieczność stworzenia porównywalnej bazy wyrazów ogólnych. Założono bowiem, że np. 10 wystąpień wyrazów graczy w korpusie 100-wyrazowym ma inne nasycenie niż wystąpienie 10 wyrazów w korpusie 50-wyrazowym. Aby porównywać nasycenie tekstów leksyką specjalistyczną, należy najpierw wyrównać liczbę ogólnych wyrazów w korpusach. Zwielokrotnienie podstawy porównania automatycznie wyzwala konieczność zwielokrotnienia o taki sam mnożnik wyrazów specjalistycznych w powiększonym korpusie.

FCDA	82.940
PSX	204.100
NPSX	386.899
FPSX	95.148

Tabela 21. Suma wyrazów graficznych graczy w poszczególnych częściach korpusu (platforma czerwona)

Liczby w powyższej czerwonej tabeli powstały w wyniku zastosowania mnożników, o których pisałem w poprzednim rozdziale. Punktem odniesienia stał się najliczniejszy korpus, a mianowicie ten pozyskany z forów internetowych. Efektem matematycznego zwiększenia wszystkich wyrazów graficznych w danej części korpusu było także zwiększenie sumy wyrazów graficznych graczy, na przykład korpus CDA platformy zielonej (czyli przed zastosowaniem jakichkolwiek mnożników) liczył łącznie 4.619.329 wyrazów graficznych (patrz tabela 17 w poprzednim rozdziale). Dzięki wykorzystaniu wyszukiwarki programu Word (pisałem o tym na stronie poprzedniej) znalazłem w nim 50.946 wyrazów graficznych graczy, natomiast w części korpusu NCDA, który zbudowany został z 11.504.288 wyrazów graficznych (patrz tabela 17) odnotowałem aż 90.506 wyrazów graficznych graczy wideo. Miałem więc do czynienia z następującą sytuacją:

$$CDA = \frac{50.946}{4.619.329} \quad NCDA = \frac{90.506}{11.504.288}$$

Powyżej dane zostały ułożone w następujący sposób:

$$\text{Skrótowa nazwa części korpusu} = \frac{\text{suma wyrazów socjolektalnych w korpusie CDA/NCDA}}{\text{całkowita suma wszystkich wyrazów w korpusie CDA/NCDA}}$$

Widać zatem, że liczba odnotowanych wyrazów graficznych graczy (50.946 i 90.506) jest większa w NCDA, ale trzeba pamiętać, że wyrazy te znalezione zostały także w bardziej obszernym korpusie tekstów, który liczy 11.504.288 wyrazów graficznych.

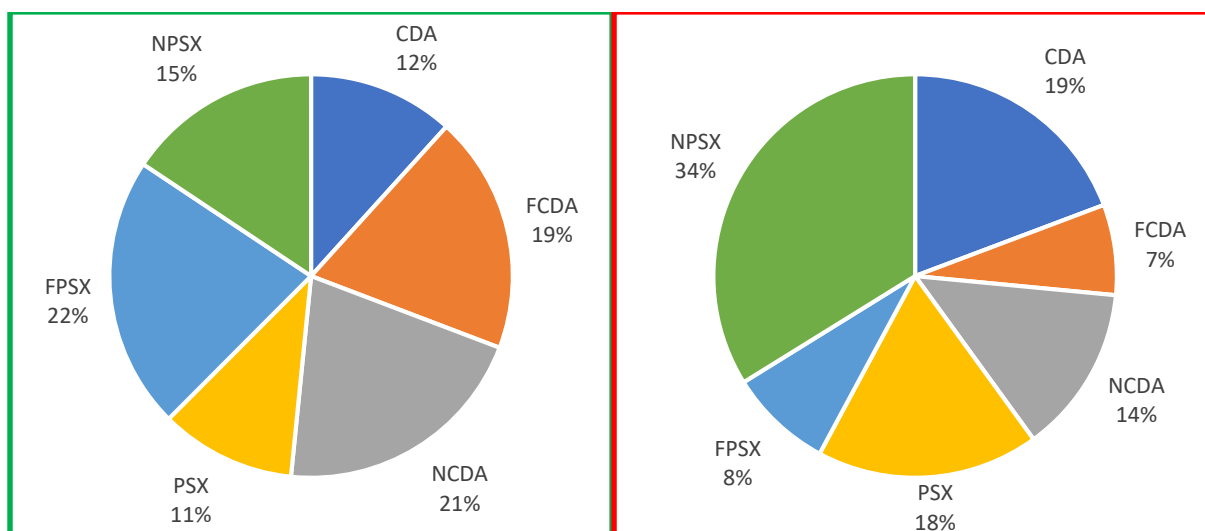
Dlatego aby wyrównać dane we wszystkich korpusach pojawiła się potrzeba stworzenia platformy czerwonej i zastosowania mnożników. Chcąc teraz wyrównać dane z obu przykładowych korpusów (CDA i NCDA), należy zastosować omówiony w poprzednim rozdziale zestaw mnożników. Po ich zastosowaniu liczebność wyrazów w korpusów oraz liczebność wyrazów graficznych graczy zmieniła się diametralnie:

$$CDA = \frac{220.377}{19.925.076} \quad NCDA = \frac{154.317}{19.491.248}$$

Mnożniki zastosowane zostały w odniesieniu do licznika, który reprezentuje sumę wyrazów socjolektalnych w korpusie CDA (220.377) i NCDA (154.317) oraz mianownika, który w obu częściach korpusu jest zbliżony. Dane na temat całkowitej sumy wszystkich wyrazów w korpusie CDA/NCDA pochodzą z tabeli umieszczonej w poprzednim rozdziale.

W dysertacji będę zatem prezentował wyniki analizy odnoszące się do danych przed zastosowaniem mnożników oraz po ich zastosowaniu.

Dla lepszego zobrazowania sum wyrazów graczy w danych częściach korpusu, umieszczam dwa wykresy kołowe, które są prezentacją procentową informacji z powyższych tabel:



Wykres 4. Suma wyrazów graficznych graczy w poszczególnych częściach korpusu w ujęciu procentowym (platforma zielona)

Wykres 5. Suma wyrazów graficznych graczy w poszczególnych częściach korpusu w ujęciu procentowym (platforma czerwona)

Analiza płaszczyzny zielonej pokazuje, że najwięcej wyrazów graczy odnotowanych zostało we wpisach na forach internetowych (fpsx i fcda), bo łącznie aż 178.088, co stanowi 40,9% wszystkich badanych wyrazów. W newsach (npsx i ncda) odnotowano łącznie 158.552 jednostki (36,5%), natomiast najmniej wyrazów znaleziono w czasopiśmie – 98.284 jednostki (22,6%).

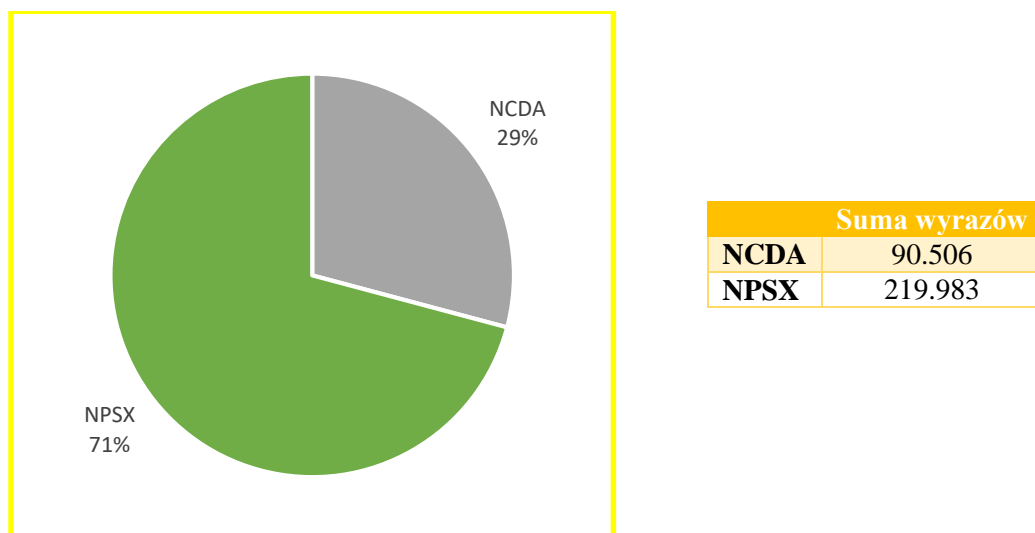
Zgoła odmiennie prezentują się wyniki przedstawione w odniesieniu do płaszczyzny czerwonej, czyli po zrównaniu liczby wyrazów ze wszystkich korpusów do średniej liczby wyrazów budujących korpus tekstów z forów internetowych. Wyrazy graczy z największym nasileniem wystąpiły w newsach, odnotowano ich tam 541.216 (47,3%). W czasopiśmie wystąpiło 424.477 wyrazów, czyli 37,1%, natomiast we wpisach na forach było ich 178.088 (15,6%).

Niezaprzeczalnie można stwierdzić, że teksty opublikowane w newsach, są najbardziej, na tle wszystkich badanych sposobów publikacji treści, nasycone wyrazami graczy. Jednostki pozyskane z newsów w jednym z zestawień uplasowały się na pozycji drugiej pod względem wszystkich zgromadzonych wyrazów graczy, a w drugim zestawieniu na miejscu pierwszym.

Natomiast porównując osobno wyniki dla dwóch czasopism (tutaj wystarczy odniesienie tylko do zielonej płaszczyzny, ponieważ korpusy tekstów z czasopism zawierają podobną liczbę wyrazów) zauważyć można, że więcej jednostek użytych zostało w tekstach czasopisma „CD-Action”.

Wynioskować można, że teksty umieszczone w tym czasopiśmie są w większym stopniu nasycone wyrazami graczy, przez co są także trudniejsze do zrozumienia dla czytelników, którzy nie są graczami wideo.

Żeby porównać nasycenie leksykalne w tekstach newsów trzeba posłużyć się nie zieloną, ale żółtą platformą, w której wyłącznie korpusy tekstowe newsów zostały ze sobą zrównane:



Wykres 6. Suma wyrazów graficznych graczy w poszczególnych częściach korpusu (NCDA i NPSX) w ujęciu procentowym i tabelarycznym (platforma żółta)

Z wykresu i danych w tabeli, wyczytać można, że liczba wyrazów graczy wideo była większa w newsach opublikowanych na stronie www.ppe.pl niż w tych ze strony www.cdaction.pl. Prawie 3/4 wszystkich wyrazów odnotowanych zostało w npsx, natomiast nieco ponad 1/4 w ncda.

Zestawiając dane na temat liczby wystąpień wyrazów graczy na forach internetowych, stwierdzić można, że fpsx nieco pod tym względem przewyższa fcda.

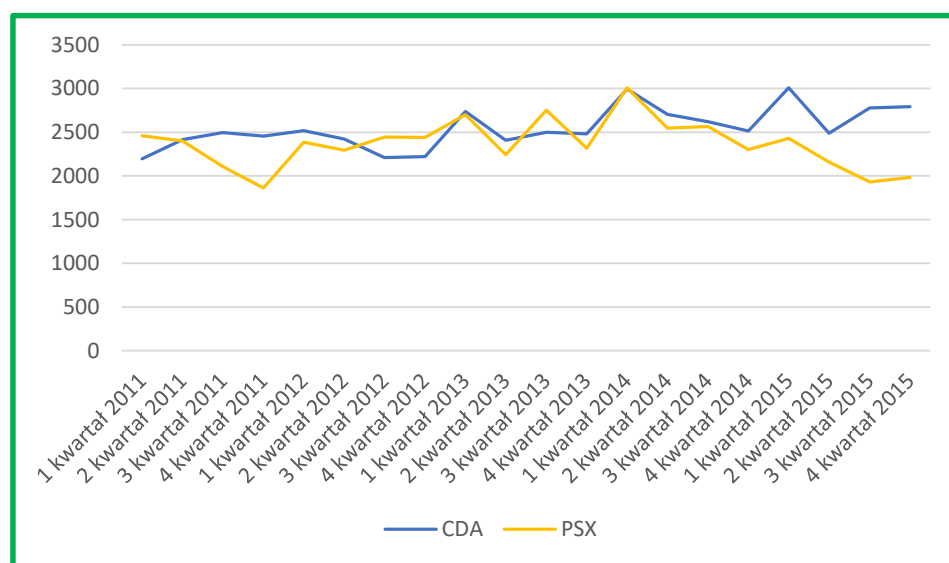
Celem podrozdziału była odpowiedź na pytanie, jaki jest procent nasycenia tekstów czasopism, newsów, forów w wyrazy graczy.

Podsumowując ten podrozdział, trzeba wyciągnąć wniosek, że środowisko *konsolowców* w znacznie większym stopniu w swoich wypowiedziach pisemnych, posługuje się wyrazami środowiskowymi. W dwóch z trzech możliwych sposobów publikacji tekstów (fpsx i npsx) to ich teksty wygrały pod względem nasycenia w wyrazy graficzne graczy. Jedynie w tekstach czasopism góruje cda.

10. WYRAZY GRAFICZNE ŚRODOWISKA GRACZY A PODZIAŁ KWARTALNY

W poprzednim podrozdziale przedstawiłem informacje na temat tego, jak pod względem ilościowym prezentuje się cały zbiór wyrazów graczy w obrębie każdego z korpusów. W tej części przyjrę się tym samym wyrazom, co w poprzednim podrozdziale, ale bardziej dokładnie, ponieważ nie będę ich traktował jako jednego dużego zbioru, ale podzielę je na kwartały, których jest łącznie po 20 w każdym z korpusów (z wyjątkiem forów). Na początku przyjrę się wyrazom pozyskanym z czasopism.

Na poniższym wykresie przedstawiłem, jak na przestrzeni pięciu lat zmieniała się liczba wystąpień wyrazów graczy, natomiast dokładne dane liczbowe przedstawiłem w tabeli po prawej stronie wykresu.



Kwartały	CDA	PSX
1 k. 11	2.194	2.457
2 k. 11	2.415	2.401
3 k. 11	2.496	2.109
4 k. 11	2.454	1.863
1 k. 12	2.517	2.387
2 k. 12	2.421	2.295
3 k. 12	2.208	2.444
4 k. 12	2.222	2.440
1 k. 13	2.738	2.700
2 k. 13	2.408	2.244
3 k. 13	2.499	2.752
4 k. 13	2.480	2.317
1 k. 14	2.992	3.009
2 k. 14	2.703	2.548
3 k. 14	2.619	2.566
4 k. 14	2.515	2.302
1 k. 15	3.010	2.429
2 k. 15	2.487	2.160
3 k. 15	2.776	1.932
4 k. 15	2.792	1.983

Wykres 7. Suma wyrazów graficznych graczy w poszczególnych kwartałach w korpusach CDA i PSX

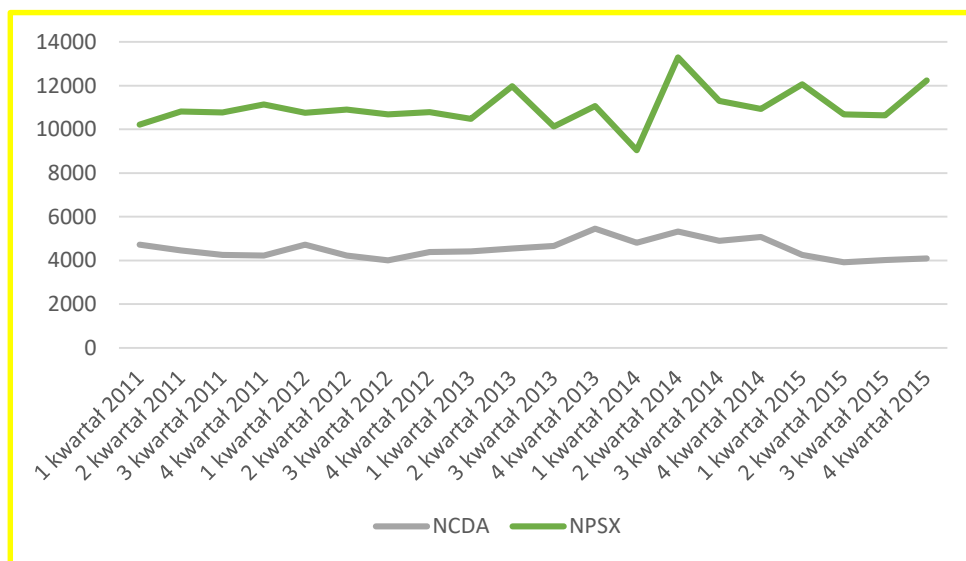
Wcześniej stwierdziłem, że w czasopiśmie „CD-Action” znaleźć można więcej wyrazów graczy wideo niż w „PSX Extreme”. Wtedy, jak już wcześniej wspominałem, uwzględniałem cały, duży zbiór wyrazów graczy, natomiast teraz rozpatruję owe zagadnienie z większą – kwartalną, dokładnością. Dopiero podczas tak zaprezentowanych wyników, zauważyć można, że nie we wszystkich kwartałach czasopismo „CD-Action” górowało nad „PSX Extreme” pod względem użycia jednostek. Kwartałami, w których psx przewyższało pod tym względem cda były: 1 kwartał 2011, 3 i 4 kw. 2012, 3 kw. 2013 oraz 1 kwartał 2014, który jest swoistym punktem granicznym. Widać, że od tego momentu psx już nigdy w badanym okresie nie górował nad cda. Być może jego pozycja utrzymała się na długo, ale to wymagałoby dłuższych badań.

Po przeanalizowaniu danych przedstawionych w tabeli zaobserwować można, że liczba wystąpień wyrazów graczy w danym kwartale w przypadku czasopisma „PSX Extreme” wyniosła 2.367. Czasopismo „CD-Action” miało natomiast średnią na poziomie 2.547 wyrazów. Jeśli natomiast weźmie się pod uwagę tylko okres od punktu granicznego, to średnia ta jeszcze bardziej przechyla się na korzyść czasopisma cda, gdzie w jednym kwartale tego okresu użyto średnio 2.737 jednostek, natomiast w psx tylko 2.366. Analiza ta jeszcze dokładniej pokazuje, że „CD-Action” w tym okresie nieustannie górowały pod względem wyrazów graficznych nad „PSX Extreme”

Ponadto widać, że wykres psx ma linię trendu spadkową, natomiast cda rosnącą, więc jeszcze dobitniej potwierdza to tę teorię. Więcej na temat samej linii trendu/tendencji wykresów piszę na stronie np. 213.

Pismo o większym nasyceniu leksyką graczy jest zatem bardziej fachowe, bardziej adekwatnie oddaje rzeczywistość graczy oraz jest w większym stopniu zakodowane (Grabias nazwał to *tajnościami*, patrz strona 26).

W przypadku danych pozyskanych z korpusów tekstów newsów, to wykres i tabela z podziałem kwartalnym wyglądają następująco:



Kwartały	NCDA	NPSX
1 k. 11	4.729	10.220
2 k. 11	4.453	10.822
3 k. 11	4.256	10.778
4 k. 11	4.229	11.140
1 k. 12	4.729	10.764
2 k. 12	4.233	10.908
3 k. 12	4.009	10.686
4 k. 12	4.388	10.793
1 k. 13	4.416	10.476
2 k. 13	4.549	11.978
3 k. 13	4.659	10.135
4 k. 13	5.456	11.070
1 k. 14	4.816	9.050
2 k. 14	5.318	13.292
3 k. 14	4.901	11.295
4 k. 14	5.071	10.940
1 k. 15	4.251	12.065
2 k. 15	3.920	10.690
3 k. 15	4.026	10.647
4 k. 15	4.097	12.234

Wykres 8. Suma wyrazów graficznych graczy w poszczególnych kwartałach w korpusach NCDA i NPSX

W przeciwieństwie do wykresu odnoszącego się do tekstów czasopism, tutaj linie reprezentujące dany korpus (ncda i npsx) nie nachodzą na siebie i nie przecinają się, więc w przypadku newsów jednoznacznie można stwierdzić, że pod względem użytych w tekstach jednostek leksykalnych graczy zawsze więcej ich jest w newsach na stronie www.ppe.pl niż na www.cdaction.pl.

Średnio w jednym kwartale npsx użyto aż 10.999 badanych wyrazów, natomiast w ncda było ich tylko 4.525. Różnica między dwoma zbiorami newsów wynosi średnio aż 6.474 jednostek.

Analizując tendencję, widać, że jest ona w obu przypadkach rosnąca. Dlatego można przypuszczać, npsx pod względem wyrazów graficznych graczy będzie nieustannie górował nad ncda.

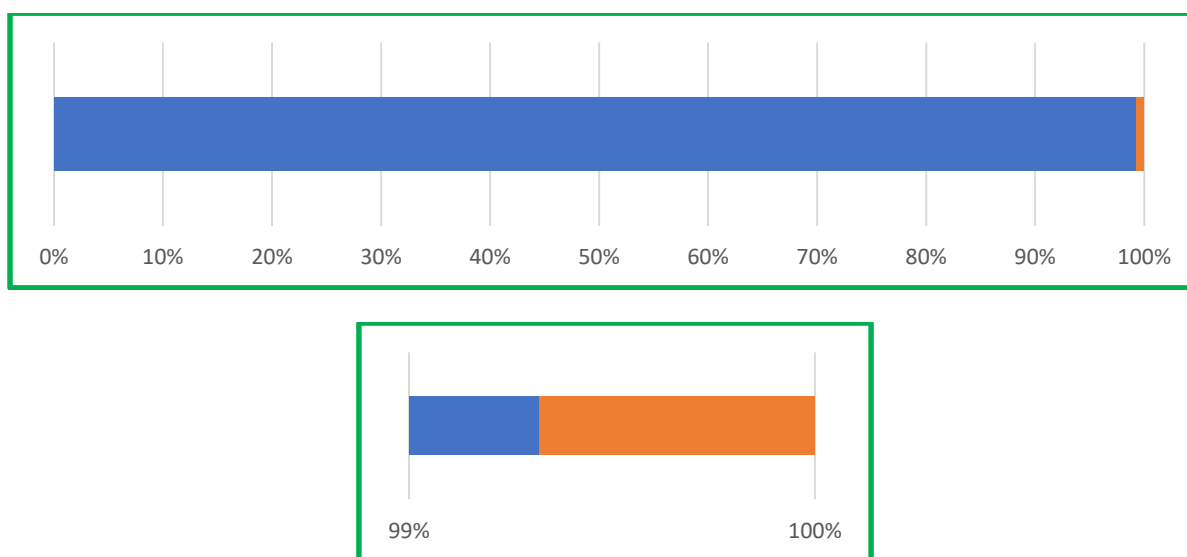
Pominałem w tej części zestawienie forów internetowych, ponieważ tam nie został dokonany podział na kwartały.

11. WYRAZY GRAFICZNE ŚRODOWISKA GRACZY WIDEO A WSZYSTKIE WYRAZY GRAFICZNE KORPUSÓW

W tym podrozdziale zestawię sumę wszystkich zbadanych wyrazów graficznych graczy wideo z pozostałymi wyrazami graficznymi w korpusach tekstów.

Warto zaznaczyć, że do badań ilościowych zgromadziłem i zbadałem tylko część wszystkich wyrazów graficznych środowiska graczy wideo, reszta pozostała w korpusie i nie została zbadana pod względem ilościowym, można nawet pokusić się o stwierdzenie, że większa część pozostała niezbadana. Poza tego typu niezbadanymi jednostkami w korpusach pozostały przede wszystkim wyrazy z polszczyzny ogólnej.

Dla lepszego zobrazowania kluczowego zagadnienia tego podrozdziału początkowo posłużę się dwoma wykresami platformy zielonej:



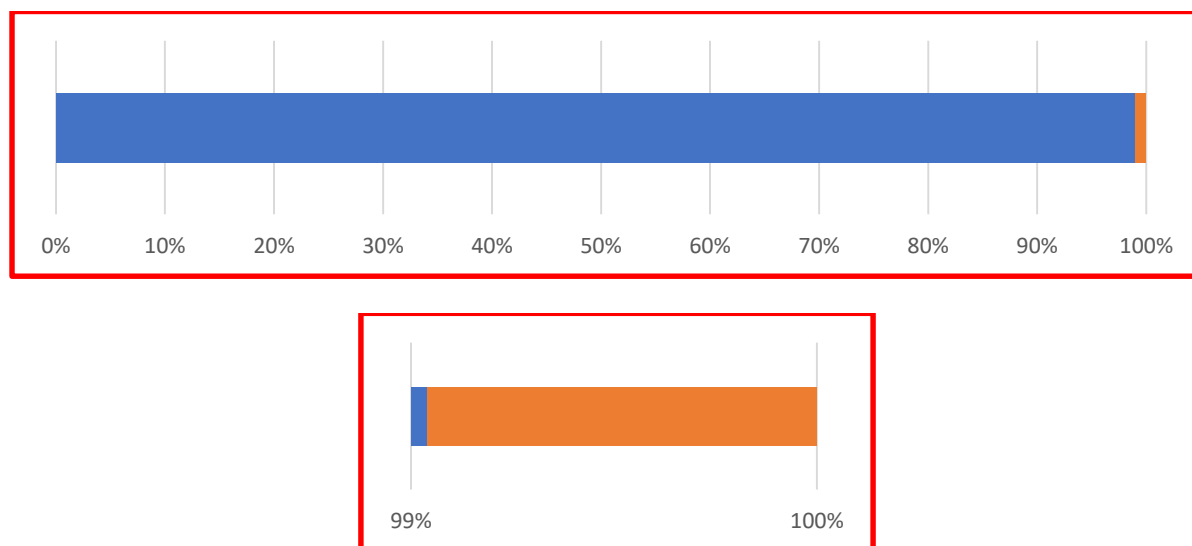
Wykres 9. Suma wyrazów graficznych środowiska graczy wideo w konfrontacji ze wszystkimi wyrazami graficznymi korpusów (platforma zielona)

Wykres pierwszy to zestawienie niezbadanych wyrazów (część niebieska) ze zbadanymi wyrazami graczy (część pomarańczowa), natomiast wykres drugi to powiększenie ostatniego procenta wykresu pierwszego.

Widać, że ponad 99% całego korpusu stanowią przede wszystkim wyrazy polszczyzny ogólnej i część niezbadanych wyrazów środowiska graczy, natomiast niecały procent stanowią wyrazy graczy.

Wszystkich wyrazów w korpusach było 64.013.852⁷⁵ (100%). W liczbie tej zawarte są niezbadane wyrazy graczy i inne wyrazy z polszczyzny ogólnej, których jest łącznie 63.578.928, co daje 99,3205783% całości oraz jednostki zbadane, których jest 434.924, co stanowi dokładnie 0,6794217% całego zbioru wyrazów. Czy to dużo? Trudno jednoznacznie odpowiedzieć, bo nie mam punktu odniesienia do innych tego typu badań socjolektalnych (z wykorzystaniem korpusów tekstów). Szczątkowym punktem odniesienia mogą być jedynie wyniki różnych badań ilościowych, np. prace J. Sambor (dokładniej opisane w części teoretycznej). Autorka badając teksty publicystyczne wykazała, że popularny, pozyskany z polszczyzny ogólnej, wyraz *woda* (ze wszystkimi swoimi słowoformami) użyty został 122 razy w korpusie tekstów, który liczył 100.000 wyrazów. Czyli wyraz ten stanowi 0,122% całego zbioru tekstów. Natomiast wyraz *i* w *Panu Tadeuszu* A. Mickiewicza (dzieło liczy łącznie, jak podaje autorka, 64.000 wyrazów), wystąpił aż 2.058 razy, co stanowi 3,215625% całego zbioru. Zestawiając wyniki moich badań z tak popularnymi wyrazami jak *woda* czy *i*, stwierdzić mogę, że liczba wyrazów graczy jest wysoka, jak na zbiór, który zawiera wyrazy zarówno z frekwencją wysoką, jak i niską. Przy czym należy pamiętać, że w miarę dalszych badań tego socjolektu liczba ta będzie wzrastać, ponieważ odkrywane i analizowane będą kolejne jednostki leksykalne.

Po zrównaniu sumy wyrazów graczy do płaszczyzny czerwonej, zaobserwować można, że liczba pokrycia całego tekstu zwiększyła się:

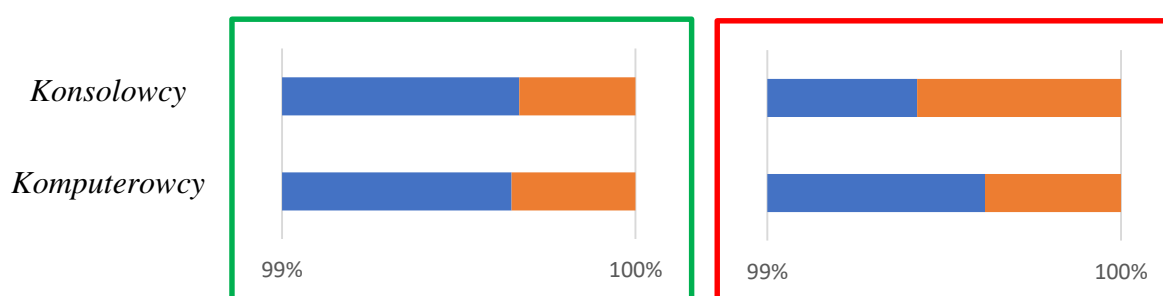


Wykres 10. Suma wyrazów graficznych środowiska graczy wideo w konfrontacji z wszystkimi wyrazami graficznymi korpusów (platforma czerwona)

⁷⁵ Jest to połączenie sumy wszystkich wyrazów graficznych w korpusach psx, npsx, fpsx, cda, ncda i fcda.

Na powyższych wykresach dane odnoszą się do sytuacji, w której cały korpus zbudowany jest z 119.035.601 wyrazów (100%). Zbiór zbadanego słownictwa graczy to 1.143.781⁷⁶ wyrazów graficznych, czyli 0,96087304% całego zbioru wyrazowego. Widać, że tutaj procent liczby wystąpień wyrazów graczy zwiększył się w porównaniu z zestawieniem zielonym.

Na zakończenie tej części pracy warto przyrzeć się jeszcze wyrazom graczy w odniesieniu do dwóch grup korpusów, z których jedna związana jest z *komputerowcami* (korpusy cda, ncda i fcda), a druga z *konsolowcami* (psx, npsx i fpsx). Wyniki tego podziału rozpatrzę zarówno na płaszczyźnie zielonej, jak i czerwonej:



Wykres 11. Suma wyrazów graficznych środowiska graczy wideo w konfrontacji ze wszystkimi wyrazami graficznymi korpusów z podziałem na korpusy *konsolowców* i *komputerowców* (platformy zielona i czerwona).

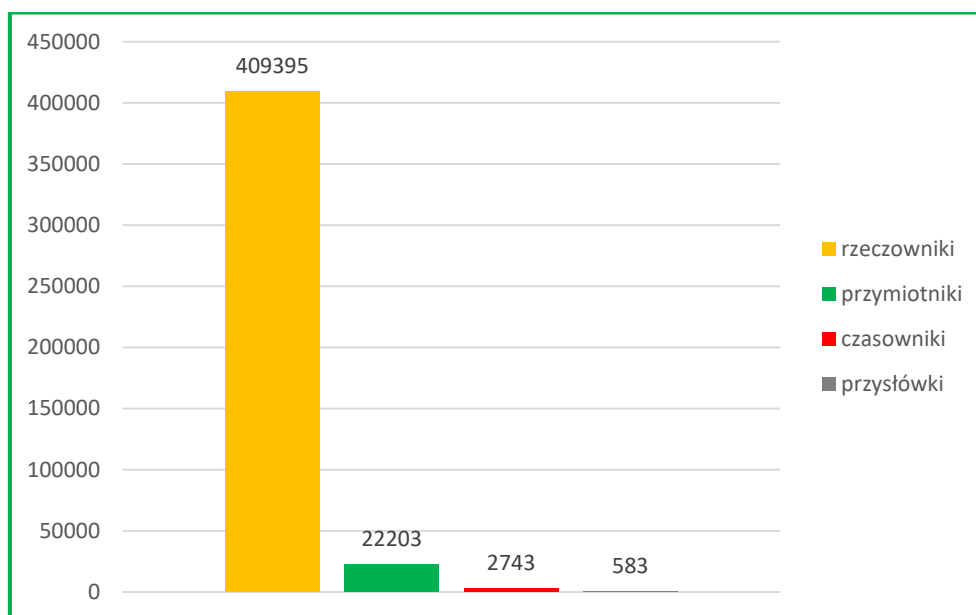
Kolor pomarańczowy odnosi się do wyrazów środowiska graczy.

Widać, że na płaszczyźnie zielonej liczba wyrazów graficznych *konsolowców* i *komputerowców* jest prawie równa, z delikatnie większym zbiorem tych wyrazów po stronie *komputerowców* (różnica między zbiorami to 13.860 wyrazów graficznych), natomiast po zastosowaniu mnożników, wyrównujących dane do płaszczyzny czerwonej widać, że różnica między zbiorami wynosi aż 228.513 jednostek. Więcej wyrazów na tej płaszczyźnie odnotowano w korpusach związanych z *konsolowcami*.

⁷⁶ Liczba ta powstała w wyniku zastosowania mnożników (patrz tabela czerwona na stronach 144 i 145).

12. WYRAZY GRAFICZNE A CZĘŚCI MOWY

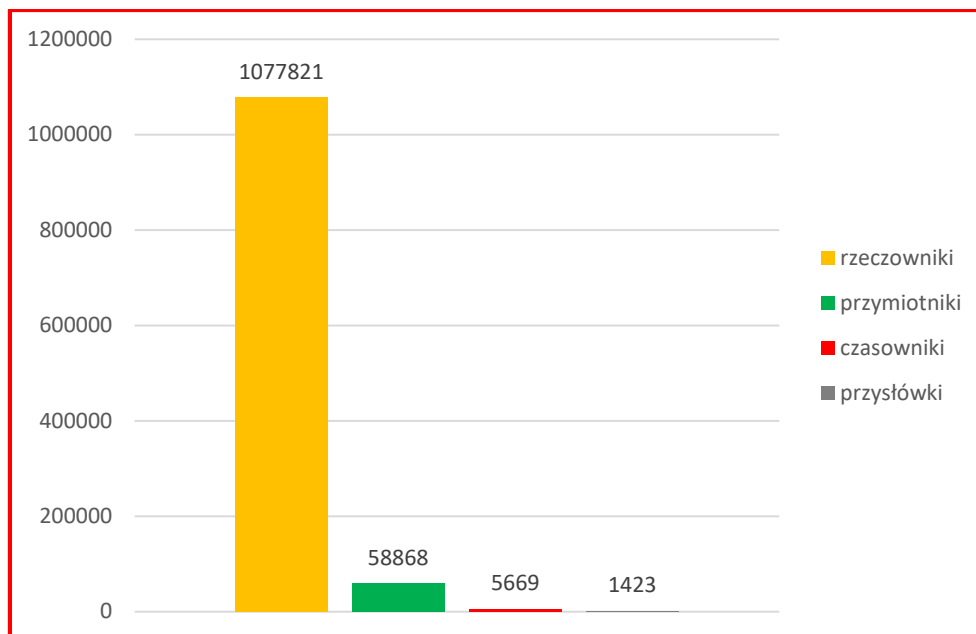
Przyglądając się częściom mowy, jakie reprezentują wyrazy graficzne graczy, można zauważyć zdecydowaną dominację rzeczowników. Szczegółowe wyniki dla płaszczyzny zielonej wyglądają następująco:



Wykres 12. Części mowy wyrazów graficznych na platformie zielonej

Wśród całego zbioru wyrazów graficznych graczy, których było łącznie 434.924, aż 84,13% stanowią rzeczowniki, reszta to pozostałe części mowy, natomiast w grupie czasowników, imiesłowów było tylko 0,2% (849).

W przypadku platformy czerwonej sytuacja nie uległa dużej zmianie:



Wykres 13. Części mowy wyrazów graficznych na platformie czerwonej

Również tutaj rzeczowniki stanowią znaczną część całego zbioru wyrazów graczy. Stanowią one aż 94,23% wszystkich wyrazów graficznych graczy. Kolejne miejsce zajmują przymiotniki, których jest 5,15%, dalej uplasowały się czasowniki (0,50%), a na końcu przysłówki, których jest tylko 0,12%.

13. SYSTEM 1.0

13.1. WPROWADZENIE DO SYSTEMU 1.0

Po stworzeniu zestawienia pól wyrazowych, które przedstawiłem na stronach 126–131, oraz przeprowadzeniu badań zauważyłem, że wyrazy graficzne poszczególnych pól leksykalnych nie występują równomiernie we wszystkich korpusach. To dość oczywiste spostrzeżenie, ale warte odnotowania, ponieważ, to dzięki niemu powstało ważne zestawienie, które prowadzi do pewnych interesujących wniosków. Poniższa tabela jest pierwszym elementem, który wprowadza do systemu 1.0.

TABELA 22

Obecność wyrazów graficznych z danych pól wyrazowych w poszczególnych częściach korpusu

Nazwa pola	PSX	CDA	NPSX	NCDA	FPSX	FCDA	Nazwa pola	PSX	CDA	NPSX	NCDA	FPSX	FCDA
Action RPG	✓	✓	✓	✓	✓	✓	Headset	✓	✓	✓	✓	✓	✓
Ban	✓	✓	✓	✓	✓	✓	Healer	✓	✓	✗	✓	✓	✓
Bijatyka	✓	✓	✓	✓	✓	✓	Hejt	✓	✓	✓	✓	✓	✓
Budżetówka	✓	✓	✓	✓	✓	✓	High-end	✓	✓	✓	✓	✓	✓
Bumper	✓	✓	✗	✓	✓	✗	Indie	✓	✓	✓	✓	✓	✓
Calak	✓	✓	✓	✓	✓	✓	Jednostrzałowiec	✓	✗	✗	✗	✓	✓
Casual	✓	✓	✓	✓	✓	✓	JRPG	✓	✓	✓	✓	✓	✓
Celowiniczek	✓	✓	✓	✓	✓	✓	Kartridż	✓	✓	✓	✓	✓	✓
Cel-shading	✓	✓	✓	✗	✓	✓	Kinect	✓	✓	✓	✓	✓	✓
Checkpoint	✓	✓	✓	✓	✓	✓	Klatkować	✓	✓	✗	✓	✓	✓
Cheat	✓	✓	✓	✓	✓	✓	Klawiatura	✓	✓	✓	✓	✓	✓
Cooldown	✓	✓	✓	✓	✓	✓	Komputer	✓	✓	✓	✓	✓	✓
Craft	✓	✓	✓	✓	✓	✓	Konsola	✓	✓	✓	✓	✓	✓
Crap	✓	✓	✓	✓	✓	✓	Kontroler	✓	✓	✓	✓	✓	✓
Creditsy	✓	✓	✓	✓	✓	✓	Konwersja	✓	✓	✓	✓	✓	✓
Cross-buy	✓	✓	✓	✓	✓	✗	Liniowy	✓	✓	✓	✓	✓	✓
Cross-chat	✗	✓	✗	✗	✓	✗	Live-stream	✗	✗	✓	✗	✓	✗
Cross-controller	✓	✗	✓	✗	✗	✗	Loading	✓	✓	✓	✓	✓	✓
Cross-gen	✓	✓	✓	✓	✓	✗	Lokalizacja	✓	✓	✓	✓	✓	✓
Crossover	✓	✓	✓	✓	✓	✓	LPM	✗	✓	✗	✓	✓	✓
Cross-platform	✓	✗	✓	✓	✓	✓	Marian	✓	✗	✓	✓	✓	✓
Cross-play	✓	✓	✓	✓	✓	✗	Matryca	✓	✓	✓	✓	✓	✓
Cross-save	✓	✗	✓	✓	✓	✗	Metagra	✗	✓	✗	✓	✗	✗
Current-gen	✓	✓	✓	✓	✓	✓	Mikrotransakcja	✓	✓	✓	✓	✓	✓
Cutscena	✓	✓	✓	✓	✓	✓	Minimapa	✓	✓	✗	✓	✓	✓
Czarnula	✓	✗	✓	✓	✓	✓	MMORPG	✓	✓	✓	✓	✓	✓
Deathmatch	✓	✓	✓	✓	✓	✓	Multi	✓	✓	✓	✓	✓	✓
Demo	✓	✓	✓	✓	✓	✓	Multikart	✓	✗	✗	✗	✗	✗
Devkit	✓	✓	✗	✓	✓	✓	Multimedia	✓	✓	✓	✓	✓	✓
DLC	✓	✓	✓	✓	✓	✓	Multiplatformowy	✓	✓	✓	✓	✓	✓
Dokanałówki	✓	✓	✗	✗	✓	✓	Multitap	✓	✓	✗	✗	✓	✗
D-pad	✓	✓	✓	✓	✓	✓	Multitasking	✓	✓	✓	✓	✓	✓
Dreamcast	✓	✓	✓	✓	✓	✓	Multitouch	✓	✓	✗	✓	✓	✓
Dungeon	✓	✓	✓	✓	✓	✓	Must-have	✓	✓	✓	✓	✓	✓
Egranizacja	✓	✓	✗	✓	✓	✓	Next-gen	✓	✓	✓	✓	✓	✓
Eksbokslan	✗	✓	✗	✗	✗	✗	Niedzielnik	✓	✓	✓	✓	✓	✓
Emulator	✓	✓	✓	✓	✓	✓	Notebook	✓	✓	✓	✓	✓	✓
Exclusive	✓	✓	✓	✓	✓	✓	NPC	✓	✓	✓	✓	✓	✓
Farmić	✓	✓	✓	✓	✓	✓	Odblokowajka	✗	✓	✗	✗	✗	✗
Feature	✓	✓	✓	✓	✓	✓	Off-road	✓	✓	✓	✓	✓	✓
Flashówka	✗	✓	✗	✓	✓	✓	One-liner	✓	✓	✓	✓	✓	✓
Framerate	✓	✓	✓	✓	✓	✓	Ośmiobitowy	✓	✓	✓	✓	✓	✓
Free-to-play	✓	✓	✓	✓	✓	✓	Pad	✓	✓	✓	✓	✓	✓
GameCube	✓	✓	✓	✓	✓	✓	Padlet	✓	✓	✓	✓	✓	✓
Gameplay	✓	✓	✓	✓	✓	✓	Past-gen	✓	✗	✓	✓	✓	✓
Gaming	✓	✓	✗	✓	✓	✓	Patch	✓	✓	✓	✓	✓	✓
Geek	✓	✓	✓	✓	✓	✓	Pauza	✓	✓	✓	✓	✓	✓
GotY	✓	✓	✓	✓	✓	✓	Pay-to-win	✓	✓	✗	✓	✓	✗
Grind	✓	✓	✓	✓	✓	✓	PC	✓	✓	✓	✓	✓	✓
Grywalny	✓	✓	✓	✓	✓	✓	Peeszcwórka	✓	✓	✗	✓	✓	✗
Grzybek	✓	✓	✗	✓	✓	✓	Peesdwójka	✓	✓	✗	✓	✓	✓
Handheld	✓	✓	✓	✓	✓	✓	Peesiks	✓	✓	✗	✗	✓	✓
HP	✓	✓	✓	✓	✓	✓	Peesplusowy	✓	✓	✗	✓	✗	✗
Hardkor	✓	✓	✓	✓	✓	✓	Peestrójka	✓	✓	✓	✓	✓	✓
Hardware	✓	✓	✓	✓	✓	✓	Pelniak	✓	✓	✗	✓	✓	✓

Platyna	✓	✓	✓	✓	✓	✓	Tekstura	✓	✓	✓	✓	✓	✓
PlayStation	✓	✓	✓	✓	✓	✓	Three-sixty	✗	✓	✗	✗	✗	✗
Podstawka	✓	✓	✓	✓	✓	✓	Touchpad	✓	✓	✓	✓	✓	✓
PPM	✗	✓	✗	✓	✓	✓	Trailer	✓	✓	✓	✓	✓	✓
Pre-alpha	✓	✓	✓	✓	✓	✓	Trigger	✓	✓	✓	✓	✓	✓
Pre-order	✓	✓	✓	✓	✓	✓	Turówka	✓	✓	✓	✓	✓	✓
Przedłużajka	✗	✓	✗	✗	✗	✗	Twarzowanie	✓	✗	✗	✗	✗	✗
Przeładarka	✓	✓	✓	✓	✓	✓	Unboxing	✓	✓	✓	✓	✓	✓
Przejsiówka	✓	✓	✓	✓	✓	✓	Używka	✓	✓	✓	✓	✓	✓
Przeszkadzajka	✓	✓	✓	✓	✓	✓	Wiesiek	✓	✓	✓	✓	✓	✓
Przetrawniówka	✗	✓	✗	✗	✗	✗	WSAD	✗	✓	✗	✓	✓	✓
Przygodówka	✓	✓	✓	✓	✓	✓	Wtyczka	✓	✓	✗	✓	✓	✓
PSP	✓	✓	✓	✓	✓	✓	Xbox	✓	✓	✓	✓	✓	✓
PSP2	✓	✓	✓	✓	✓	✗	YLOD	✓	✗	✗	✓	✓	✓
PSP go	✓	✓	✓	✓	✓	✗	Zapychacz	✓	✓	✓	✓	✓	✓
PvE	✓	✓	✓	✓	✓	✓	Zblurowany	✓	✗	✗	✗	✓	✓
PvP	✓	✓	✓	✓	✓	✓	Znajdźka	✓	✓	✓	✓	✓	✓
QTE	✓	✓	✓	✗	✓	✓	Żywołność	✓	✓	✓	✓	✓	✓
Quest	✓	✓	✓	✓	✓	✓							
QWERTY	✓	✓	✗	✗	✓	✓							
Randomizacja	✗	✓	✗	✓	✓	✓							
Real-time	✓	✓	✗	✓	✓	✓							
Redzi	✓	✓	✗	✓	✓	✓							
Reedycja	✓	✓	✓	✓	✓	✓							
Remake	✓	✓	✓	✓	✓	✓							
Remaster	✓	✓	✓	✓	✓	✓							
Roguelike	✓	✓	✓	✓	✓	✓							
Router	✓	✓	✗	✓	✓	✓							
Rozdzielczość	✓	✓	✓	✓	✓	✓							
RPG	✓	✓	✓	✓	✓	✓							
RROD	✓	✗	✓	✓	✓	✓							
Samochodówka	✓	✓	✓	✓	✓	✓							
Samograj	✓	✓	✓	✓	✓	✓							
Sandbox	✓	✓	✓	✓	✓	✓							
Save	✓	✓	✓	✓	✓	✓							
Sequel	✓	✓	✓	✓	✓	✓							
Shader	✓	✓	✓	✓	✓	✓							
Shmup	✓	✓	✓	✓	✓	✓							
Side-quest	✓	✓	✗	✓	✓	✓							
Sieciówka	✓	✓	✗	✓	✓	✓							
Singiel	✓	✓	✓	✓	✓	✓							
Skin	✓	✓	✓	✓	✓	✓							
Skrypt	✓	✓	✓	✓	✓	✓							
Smartfon	✓	✓	✓	✓	✓	✓							
Software	✓	✓	✓	✓	✓	✓							
Soundtrack	✓	✓	✓	✓	✓	✓							
Spawn	✓	✓	✓	✓	✓	✓							
Sprite	✓	✓	✓	✓	✓	✓							
Stacjonarka	✓	✓	✓	✓	✓	✓							
Stodwudziestoośmiobitowiec	✗	✓	✗	✗	✗	✗							
Stream	✓	✓	✓	✓	✓	✓							
Strzelanina	✓	✓	✓	✓	✓	✓							
Stylus	✓	✓	✓	✓	✓	✓							
Sub-boss	✓	✗	✓	✗	✓	✗							
Sub-quest	✓	✗	✗	✗	✓	✓							
Subskrypcja	✓	✓	✓	✓	✓	✓							
Szesnastobitowy	✗	✓	✗	✓	✓	✗							
Ściągarka	✓	✓	✓	✓	✓	✓							
Tablet	✓	✓	✓	✓	✓	✓							
Tank	✓	✓	✓	✓	✓	✓							
Tekstówka	✓	✓	✗	✓	✓	✓							

W powyższej tabeli pokazałem pola (reprezentują je wyrazy podstawowe, hasłowe), które wystąpiły w danym korpusie tekstów (nazwy korpusów znajdują się w główce tabeli, są to: cda, ncda, fcda, psx, npsx, fpsx), a których tam zabrakło. Te, które obecne są w korpusie, oznaczyłem symbolem ✓, natomiast te, których brakuje symbolem X. O ile drugi symbol można w polszczyźnie nazwać krzyżykiem, o tyle problem jest z nazwaniem pierwszego symbolu. Na potrzeby tej pracy będę używał angielskiej nazwy w odniesieniu do tego symbolu, czyli *tick*⁷⁷. Aby dane pole otrzymało znaczek *tick*, to przynajmniej jeden z wyrazów graficznych tego pola musiał wystąpić minimum jeden raz w danym korpusie. Maksymalnie jedno pole leksykalne mogło otrzymać sześć znaczków, a minimalnie jeden, ponieważ pól, które nie wystąpiły w żadnym korpusie nie uwzględniałem w badaniu.

W powyższej tabeli zrezygnowałem z prezentacji dokładnych danych liczbowych, ponieważ temu poświęcona zostanie dalsza część pracy, a teraz interesuje mnie tylko to, czy dane pole wyrazowe wystąpiło w danym korpusie oraz w ilu jeszcze zbiorach tekstów jest ono obecne.

Na początku trzeba przyjrzeć się możliwie szeroko powyższemu zestawieniu i sprawdzić, w którym z korpusów wystąpiło największe bogactwo pól wyrazowych związanych z graczami. W jednym korpusie tekstów mogło wystąpić maksymalnie 189 pól, a minimalnie 1. Patrząc na zestawienie widać, że najbogatszym, pod tym względem pól jest fpsx, w których jest aż 178 znaczków *tick*, a tylko 11 krzyżyków. Aby uzyskać pełne spektrum informacji, należy jeszcze odnotować, jakich pól zabrakło w tym korpusie:

peesplusowy, [metagra](#), [cross-controller](#), [multikart](#), [twarzowanie](#), [odblokowajka](#), [przetrvaniówka](#), [Eksbokslan](#), [przedłużajka](#), [stodwudziestoośmiobitowiec](#), [Three-sixty](#).

Drugie miejsce zajęły ex aequo korpusy zbudowane z tekstów czasopism, czyli cda i psx, w których wynotowano aż po 174 znaczków *tick*, a tylko 15 krzyżyków.

W korpusie cda nie znaleziono takich pól, jak:

[cross-controller](#), cross-platform, cross-save, czarnula, jednostrzałowiec, [live-stream](#), Marian, [multikart](#), past-gen, RROD, sub-boss, sub-quest, [twarzowanie](#), YLOD, zblurowany,

natomiast w psx zabrakło pól:

WSAD, [metagra](#), PPM, flashówka, [odblokowajka](#), szesnastobitowy, randomizacja, [przetrvaniówka](#), [Eksbokslan](#), cross-chat, LPM, [przedłużajka](#), [stodwudziestoośmiobitowiec](#), [Three-sixty](#), [live-stream](#).

Zestawiając ze sobą kilka korpusów dostrzec można, że pewnej części pól wyrazowych zabrakło tylko w jednym z dwóch korpusów (czarne nazwy w powyższych wyliczeniach),

⁷⁷ Moim zdaniem brzmi to lepiej niż np. *ptaszek*

a części w obu korpusach (nazwy niebieskie). Spostrzeżenie to dało podstawy do przypuszczenia, że różne pola związane są z różnymi rodzajami tekstów, a co się z tym wiąże, także z różnymi grupami graczy wideo, które te teksty stworzyły. Jak już wspominałem są podstawą do hipotezy, że pewien zbiór słownictwa związany jest wyłącznie z *konsolowcami*, a inny z *komputerowcami*. Wynika z tego, że w przypadku analizy pól leksykalnych można dokonać takiego podziału. Jednak na tę chwilę nie dysponuję jeszcze dostatecznymi danymi, aby to wykonać, bo analiza trzech korpusów pod względem występujących lub brakujących w nich pól nie wystarczy. Wróć do tego w dalszej części pracy, a teraz pokażę, jak wygląda sprawa liczebności pól w innych korpusach oraz jakich pól w nich zabrakło.

Trzecie miejsce pod względem bogactwa pól zajął korpus *ncda*, w którym odnotowano 168 pól wyrazowych graczy, a zabrakło wśród nich takich, jak:

live-stream, Peesiks, QTE, cross-controller, cel-shading, sub-boss, multitap, sub-quest, jednostrzałowiec, multikart, QWERTY, dokanałówki, twarzowanie, zblurowany, odblokowajka, przetrwaniówka, Eksbokslan, cross-chat, przedłużajka, stodwudziestoośmiobitowiec, Three-sixty.

W korpusie *fcda* odnotowano 164 pól, a wśród nich nie było takich, jak:

cross-buy, PSP2, pay-to-win, Peeschwórka, peesplusowy, cross-play, szesnastobitowy, PSP go, cross-gen, cross-save, bumper, metagra, live-stream, cross-controller, sub-boss, multitap, multikart, twarzowanie, odblokowajka, przetrwaniówka, Eksbokslan, cross-chat, przedłużajka, stodwudziestoośmiobitowiec, Three-sixty.

Na ostatnim miejscu, czyli jest to korpus posiadający najmniej pól leksykalnych, uplasował się *npsx*. Znalaziono w nim tylko 145 pól, a zabrakło:

gaming, bumper, YLOD, sieciówka, devkit, side-quest, real-time, pay-to-win, egranizacja, minimapa, router, multitap, grzybek, sub-quest, multitouch, tekstówka, wtyczka, peeschwórka, klatkować, jednostrzałowiec, pełniak, peesplusowy, multikart, healer, QWERTY, redzi, dokanałówki, Peesdwójka, twarzowanie, zblurowany, WSAD, metagra, PPM, flashówka, odblokowajka, szesnastobitowy, randomizacja, przetrwaniówka, Eksbokslan, cross-chat, LPM, przedłużajka, stodwudziestoośmiobitowiec, Three-sixty.

Powyżej przeanalizowałem każdy z korpusów osobno, aby sprawdzić, w którym jest najwięcej pól, a w którym najmniej. Im większa liczba pól znalezionych w korpusach, tym teksty je budujące są bardziej różnorodne, ale także i trudniejsze w zrozumieniu dla przeciętnego (nie gracza) czytelnika.

Patrząc na korpusy z nieco innego punktu widzenia, czyli dzieląc je na 3 grupy, wyciągnąć można inne ciekawe wnioski. Celem tego podziału jest sprawdzenie, w jakich tekstach opublikowanych w czasopismach, newsach lub na forach, występuje najwięcej pól, a które w nich są nieobecne. Innymi słowy podział na grupy dokonany został ze względu na sposób publikacji tekstów. W pierwszej grupie (newsy) umieszczone zostały korpusy *ncda* i *npsx*, w drugiej (fora) *fpsx* i *fcda*, natomiast w trzeciej (czasopisma) *cda* i *psx*.

Najwięcej pól leksykalnych odnotowano w tekstach czasopism. Było ich aż 188. Maksymalnie mogły tam wystąpić 189 pól, więc zabrakło tylko jednego – live-stream.

W tekstach na forach internetowych znalazło się 178 pól, a wśród nich zabrakło:

cross-controller, Eksbokslan, metagra, multikart, odblokowajka, peesplusowy, przedłużajka, przetrwaniówka, stodwudziestoośmiobitowiec, Three-sixty, twarzowanie.

Na ostatnim miejscu pod względem bogactwa pól wyrazowych uplasowały się teksty newsów. Było w nich tylko 174 pól wyrazowych, zabrakło natomiast pól

cross-chat, dokanałówki, Eksbokslan, jednostrzałowiec, multikart, multitap, odblokowajka, przedłużajka, przetrwaniówka, QWERTY, stodwudziestoośmiobitowiec, sub-quest, Three-sixty, twarzowanie, zblurowany

Jednak trzeba zaznaczyć, że zakres wahania pod względem wystąpień pól w danych rodzajach tekstów jest mały, ponieważ wynosi mniej niż 10%. To świadczy o porównywalnym bogactwie leksykalnym.

W powyższych rozważaniach udało się dostarczyć odpowiedzi np. na to, w jakich rodzajach tekstów występuje najwięcej pól lub w jakim korpusie odnotowano najmniejsze bogactwo pod względem użycia wyrazów graficznych z pól wyrazowych. Przed przejściem do kolejnego zestawienia, trzeba jeszcze odpowiedzieć na pytanie, które teksty, jakiej grupy środowiska graczy wideo, są bogatsze w pola, a przez to bardziej różnorodne. Czy są to *konsolowcy*, z którymi związane są korpusy psx, npsx i fpsx, czy też *komputerowcy* z tekstami zgromadzonymi w korpusach cda, fcda i ncda.

W tekstach graczy komputerowych wystąpiły 184 pola (ze wszystkich możliwych 189), natomiast w publikacjach *konsolowców* było ich nieco mniej, bo 182. Bogatsze więc w pola leksykalne są teksty graczy komputerowych.

U *komputerowców* nie znalazłem przedstawicieli wyrazowych z zakresu tych pól:

cross-controller, live-stream, multikart, sub-boss, twarzowanie,

natomiast u *konsolowców* zabrakło pól:

Eksbokslan, metagra, odblokowajka, przedłużajka, przetrwaniówka, stodwudziestoośmiobitowiec, Three-sixty

Można więc stwierdzić, że pola wyrazowe, które nie wystąpiły w korpusach *konsolowców* są charakterystyczne wyłącznie dla grupy *komputerowców* i odwrotnie.

Teraz, gdy już opisałem i wyciągnąłem możliwie dużo informacji z powyższej tabeli (22), mogę przejść do tego, o czym pisałem nieco wyżej, czyli do podziału pól wyrazowych na te, które są związane ze środowiskiem graczy komputerowych oraz te, które występują wyłącznie w tekstach *konsolowców*.

Analiza doprowadziła do powstania klasyfikacji, którą nazwałem systemem 1.0. w nawiązaniu do leksyki środowiska graczy.

13.2. ZASADY SYSTEMU 1.0

Z powyższych rozważań i analiz zarówno zestawień, jak i samej tabeli, widać, że zarysowane zostały pewne granice grup pól wyrazowych, które wspólne są wyłącznie danym korpusom, miejscom publikacji tekstów lub środowiskom. Grupy te nie są jednak w równy sposób do siebie zbliżone, nie są też identycznie zbudowane. w swoich badania postanowiłem wprowadzić pewien system podziału pól, w zależności od tego, z jaki środowiskiem są one powiązane, czyli czy bliżej im do środowiska *komputerowców* czy też do *konsolowców*.

System ten opiera się na sześciu korpusach, które stanowią podstawę materiałową moich badań, czyli na: psx, npsx, fpsx, cda, ncda i fcda. Poza korpusami, niezbędnym jego elementem jest tabela rejestrująca to, czy dane pola wystąpiły w jakimś korpusie, czy też nie wystąpiły (patrz tabela 22). Przy pomocy systemu można ustalić stopień powiązania danego pola z określonym środowiskiem graczy wideo.

System ten pozwala na odpowiedź na jedną z postawionych w tej pracy hipotez, a mianowicie, że istnieją jednostki związane wyłącznie z daną podgrupą graczy wideo (*konsolowców*, *komputerowców*) oraz jak silnie są z nią związane. Na czym polega działanie tego systemu podziału? Najłatwiej wyjaśnić to na przykładach.

Jako pierwszy przykład posłuży pole *konsola*, które posiada 6 znaczków *tick*, czyli maksymalną ich liczbę:

	PSX	CDA	NPSX	NCDA	FPSX	FCDA
Konsola	✓	✓	✓	✓	✓	✓

Czy po spojrzeniu na powyższą tabelkę można już stwierdzić, że pole *konsola* związane jest z danym środowiskiem? Na podstawie analizy powyższych danych, można jedynie powiedzieć, że jest ono w równym stopniu związane ze środowiskiem *konsolowców* i *komputerowców*. Innymi słowy, wyrazy z tego pola można spotkać w tekstach zarówno środowiska graczy komputerowych, jak i konsolowych, nie jest więc ono ograniczone do żadnej z grup.

Nieco inaczej wygląda sprawa z polem *sub-boss*

	PSX	CDA	NPSX	NCDA	FPSX	FCDA
Sub-boss	✓	✗	✓	✗	✓	✗

Widać, że wyrazy tego pola wystąpiły we wszystkich trzech korpusach reprezentujących tylko teksty graczy konsolowych (psx, npsx i fpsx). Można więc stwierdzić, że wyrazy tego pola są silnie związane ze środowiskiem *konsolowców*.

Jeszcze inaczej wygląda sytuacja pola *healer*:

	PSX	CDA	NPSX	NCDA	FPSX	FCDA
Healer	✓	✓	✗	✓	✓	✓

Pole to wystąpiło we wszystkich trzech korpusach reprezentujących teksty graczy komputerowych (cda, ncda i fca), ale także w 2 korpusach reprezentujących teksty środowiska graczy konsolowych (psx i fpsx). Wyrazów tego pola nie odnotowano jedynie w korpusie npsx. Można więc stwierdzić, że pole to jest delikatnie / trochę bardziej związane ze środowiskiem *komputerowców*, niż *konsolowców*.

Powyżej przedstawiłem tylko trzy możliwości rozłożenia znaczków *tick* w danych tabelach, w rzeczywistości jest ich znacznie więcej. Niemniej te trzy przedstawione sytuacje dobrze służą zobrazowaniu zasad działania systemu. Większa liczba sytuacji przedstawiona została w tabeli poniżej:

TABELA 23. STOPIEŃ POWIĄZANIA PÓL WYRAZOWYCH Z DANYM ŚRODOWISKIEM (ZASADY)

Duży	Średni	Mały	Pola wspólne	Mały	Średni	Duży
<ul style="list-style-type: none"> • Pole wystąpiło tylko w cda, ncda i fcda (posiada aż 3 ticki w polach cda, ncda i fcda) oraz jeden raz w 1 z 3 możliwych korpusów • Pole wystąpiło w 2 z 3 możliwych korpusów związanych ze środowiskiem graczy komputerowych (posiada 2 ticki w cda lub ncda lub fcda) • Pole wystąpiło tylko w 1 z 3 możliwych korpusów związanych ze środowiskiem graczy komputerowych (posiada tylko 1 tick w psx lub npsx lub fpsx) • Pole wystąpiło w 2 z 3 możliwych korpusów związanych ze środowiskiem graczy komputerowych (posiada 2 ticki w cda lub ncda lub fcda) • Pole wystąpiło tylko w 1 z 3 możliwych korpusów związanych ze środowiskiem graczy komputerowych (posiada 2 ticki w polach cda lub ncda lub fcda) oraz 1 raz w jednym z 3 możliwych korpusów związanych ze środowiskiem graczy komputerowych (posiada tylko 1 tick w cda lub ncda lub fcda) 	<ul style="list-style-type: none"> • Pole wystąpiło we wszystkich korpusach środowiska graczy komputerowych (posiada aż 3 ticki w polach cda, ncda i fcda) oraz jeden raz w 1 z 3 możliwych korpusów związanych ze środowiskiem graczy komputerowych (posiada 2 ticki w cda lub ncda lub fcda) oraz w 2 z trzech w możliwych korpusów związanych ze środowiskiem graczy komputerowych (posiada 2 ticki w cda lub ncda lub fcda) oraz 2 razy w możliwych korpusach związanych ze środowiskiem graczy komputerowych (posiada 2 ticki w psx lub npsx lub fpsx) • Pole wystąpiło w 1 z 3 w możliwych korpusów związanych ze środowiskiem graczy komputerowych (posiada 2 ticki w psx lub npsx lub fpsx) 	<ul style="list-style-type: none"> • Pole wystąpiło we wszystkich korpusach środowiska graczy komputerowych (posiada aż 3 ticki w polach cda, ncda i fcda) oraz 2 razy w możliwych korpusach związanych ze środowiskiem graczy komputerowych (posiada 2 ticki w psx lub npsx lub fpsx) 	<ul style="list-style-type: none"> • Pole wystąpiło we wszystkich możliwych korpusach (posiada 6 ticków) • Pole wystąpiło w 2 z trzech w możliwych korpusów związanych ze środowiskiem graczy komputerowych (posiada 2 ticki w cda lub ncda lub fcda) oraz w 2 z trzech w możliwych korpusów związanych ze środowiskiem graczy komputerowych (posiada 2 ticki w psx lub npsx lub fpsx) • Pole wystąpiło w 1 z 3 w możliwych korpusów związanych ze środowiskiem graczy komputerowych (posiada 2 ticki w cda lub ncda lub fcda) oraz w 1 z trzech w możliwych korpusów związanych ze środowiskiem graczy komputerowych (posiada tylko 1 tick w cda lub ncda lub fcda) oraz w 1 z trzech w możliwych korpusów związanych ze środowiskiem graczy komputerowych (posiada tylko 1 tick w psx lub npsx lub fpsx) 	<ul style="list-style-type: none"> • Pole wystąpiło we wszystkich korpusach środowiska graczy komputerowych (posiada aż 3 ticki w polach psx, npsx i fpsx) oraz jeden raz w 1 z 3 możliwych korpusów związanych ze środowiskiem graczy komputerowych (posiada tylko 1 tick w cda lub ncda lub fpsx) • Pole wystąpiło w 2 z 3 możliwych korpusów związanych ze środowiskiem graczy komputerowych (posiada 2 ticki w psx lub npsx lub fpsx) oraz 2 razy w możliwych korpusach związanych ze środowiskiem graczy komputerowych (posiada 2 ticki w cda lub ncda lub fcda) 	<ul style="list-style-type: none"> • Pole wystąpiło we wszystkich korpusach środowiska graczy komputerowych (posiada aż 3 ticki w polach psx, npsx i fpsx) oraz jeden raz w 1 z 3 możliwych korpusów związanych ze środowiskiem graczy komputerowych (posiada tylko 1 tick w cda lub ncda lub fpsx) • Pole wystąpiło w 2 z 3 możliwych korpusów związanych ze środowiskiem graczy komputerowych (posiada 2 ticki w psx lub npsx lub fpsx) • Pole wystąpiło tylko w 1 z 3 możliwych korpusów związanych ze środowiskiem graczy komputerowych (posiada tylko 1 tick w psx lub npsx lub fpsx) 	<ul style="list-style-type: none"> • Pole wystąpiło tylko w psx, npsx i fpsx (posiada aż 3 ticki w polach cda, ncda i fcda) • Pole wystąpiło w 2 z 3 możliwych korpusów związanych ze środowiskiem graczy konsolowych (posiada 2 ticki w psx lub npsx lub fpsx) • Pole wystąpiło tylko w 1 z 3 możliwych korpusów związanych ze środowiskiem graczy konsolowych (posiada tylko 1 tick w psx lub npsx lub fpsx)

System, jak widać, patrząc na główkę powyższej tabeli, zakłada istnienie siedmiu stopni/grup powiązania pól wyrazowych z danym środowiskiem. Pole może być w stopniu *dużym*, *średnim* lub *małym* powiązane z *konsolowcami* lub *komputerowcami*. Każda z grup umieszczona jest w jednej kolumnie tabeli, a te z kolei umieszczone są między napisami *konsolowcy* i *komputerowcy*. W głównej części tabeli wypisane zostały wymogi, jakie musi spełnić dane pole, aby znalazło się w konkretnej kolumnie. i tak np. pole *konsola*, omówiona powyżej, znajdzie się w kolumnie środkowej – *pola wspólne*, ponieważ to „pole wystąpiło we wszystkich możliwych korpusach (posiada 6 ticków)”. Pole *sub-boss* znajdzie się w kolumnie po prawej stronie – *powiązanie duże*, ponieważ „pole wystąpiło tylko w psx, npsx i fpsx (posiada aż 3 ticki w polach cda, ncda i fcda)”, natomiast pole *healer* umieszczone zostanie w kolumnie *powiązanie małe*, po lewej stronie, bo „pole wystąpiło we wszystkich korpusach środowiska graczy komputerowych (posiada aż 3 ticki w polach cda, ncda i fcda) oraz 2 razy w możliwych korpusach związanych ze środowiskiem graczy konsolowych (posiada 2 ticki w psx lub npsx lub fpsx)”.

Pola leksykalne umieszczone najbliżej lewej strony są najbardziej związane ze środowiskiem graczy komputerowych (wystąpiły one tylko w korpusach cda lub ncda lub npsx), natomiast te pola, które umieszczone zostały po prawej stronie w kolumnie *powiązanie duże* są związane tylko ze środowiskiem graczy konsolowych. Pola umieszczone pośrodku, w kolumnie *pola wspólne* to te, które wystąpiły w równym stopniu w korpusach cda, fcda, ncda oraz psx, fpsx i npsx, czyli nie są przypisane ani do jednej, ani do drugiej grupy graczy wideo. Są to po prostu pola, które w równym stopniu występują w korpusach tekstów *konsolowców* oraz *komputerowców*.

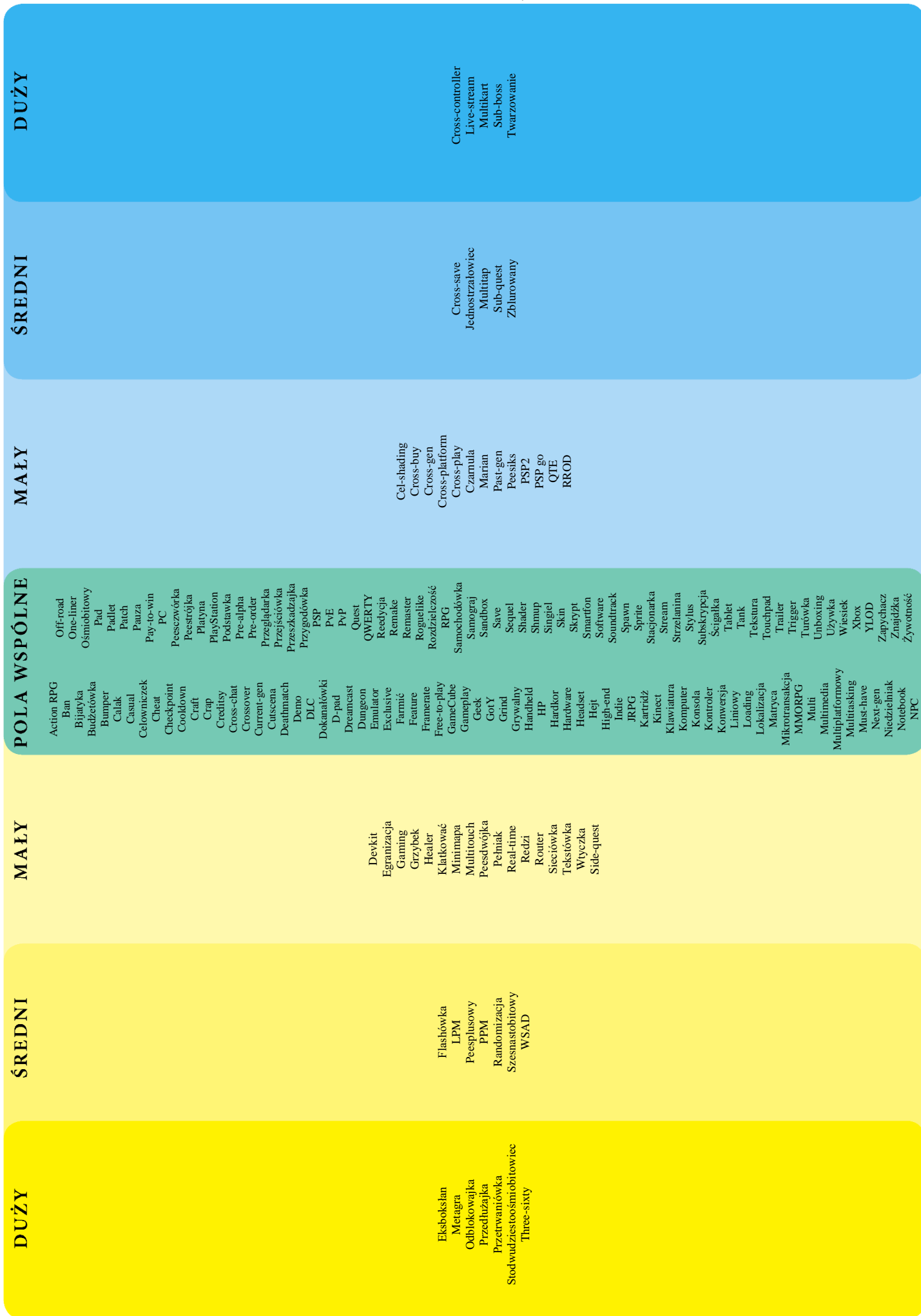
System ten nazwany został 1.0, ponieważ w dalszej części pracy zostanie zaprezentowany system 2.0, więc dla rozróżnienia wprowadziłem cyfry. Tym bardziej znajduje to sens, ponieważ system 1.0 będzie jednym ze elementów systemu 2.0. Natomiast na tę chwilę wystarczy jego wersja pierwsza.

Dla większej czytelności wyników powstałych w oparciu o powyższy system poszczególne kolumny tabeli zostały pokolorowane. Te pola, które są bardziej związane z *konsolowcami* będą umieszczone w kolumnach koloru niebieskiego i w zależności od tego, czy będą z tym środowiskiem związane w stopniu *małym*, *średnim* lub *dużym* nasycenie koloru będzie się zwiększać. Pola związane ze środowiskiem komputerowców będą w kolumnach koloru żółtego i także zróżnicowane zostanie nasycenie tego koloru, natomiast pola *wspólne* obu środowiskom, będą w kolumnie koloru zielonego – połączenia barwy żółtej i niebieskiej.

13.3. WYNIKI SYSTEMU 1.0

W efekcie zastosowanie zasad systemu 1.0, które przedstawiłem w poprzednim podrozdziale, wszystkie badane pola leksykalne (wyrazowe) zostały podzielone na następujące obszary:

GRAFIKA 9. STOPIEŃ POWIĄZANIA PÓL WYRAZOWYCH Z DANYM ŚRODOWISKIEM (WYNIKI)



Jak widać z zestawienia tylko 7 pól związanych jest wyłącznie ze środowiskiem graczy komputerowych (żółta kolumna *powiązanie duże*). Zbiór ten stanowi, w przybliżeniu, 3,70% wszystkich pól umieszczonych w tabeli, natomiast pól związanych wyłącznie ze środowiskiem graczy konsolowych jest nieco mniej, bo tylko 5. Co stanowi, w przybliżeniu, 2,65% całego zbioru.

Rozpatrując pola, które znalazły się w kolumnach żółtych, czyli są w jakiś sposób związane z *komputerowcami*, najczęściej odnotowano ich w kolumnie jasnożółtej, czyli są one tylko w *stopniu średnim* powiązane z daną grupą graczy. Znalazło się tam 17 pól, co stanowi 8,99% całego zbioru.

Podobnie wygląda sytuacja w odniesieniu do pól związanych z *konsolowcami*, bo tutaj także najobszerniejsza w pola wyrazowe jest kolumna jasnoniebieska (*powiązanie małe*). W kolumnie tej umieszczonych zostało 13 pól, czyli 6,88% całego zbioru.

Patrząc całościowo na tabelę i wydzielając tylko trzy grupy obszarów (żółty, zielony i niebieski, pomijając ich odcienie) to najczęściej pól znalazło się w kolumnie zielonej. Jest ich tam aż 135. W kolumnach żółtych jest łącznie 31 pól, natomiast w niebieskich tylko 23. W ujęciu procentowym pola w kolumnie zielonej stanowią 71,43%, w kolumnach żółtych 16,40%, a niebieskich 12,17%.

Prawie 3/4 wszystkich pól występuje w obu środowiskach (są to te, które umieszczone zostały w kolumnie zielonej), natomiast ponad 1/4 całego zbioru to pola charakterystyczne dla danych środowisk, czyli są mniej lub bardziej związane z *konsolowcami* lub *komputerowcami*.

Na zakończenie tego fragmentu pracy stwierdzić trzeba, że wypracowany system dobrze sprawdza się w typologii pól leksykalnych pod względem stopnia ich powiązania z daną grupą graczy wideo.

14. ZŁOŻONOŚĆ A SYSTEM 1.0

W pierwszym rozdziale części analitycznej omówiłem, czym jest złożoność pól oraz zaprezentowałem, jak ona wygląda w odniesieniu do mojego materiału badawczego. Na tym etapie pracy należałoby nieco rozszerzyć tabelę zaprezentowaną w tamtym rozdziale o i dodać informacje na temat wystąpień poszczególnych pól oraz form podstawowych w danych częściach korpusu.

Tabela ta wygląda następująco:

TABELA 24. Wykaz wystąpień poszczególnych pól oraz form podstawowych w danych częściach korpusu

Nazwa pola	Nazwa formy podstawowej	cda	psx	fcda	fpsx	ncda	npsx	Złożoność pola	Nazwa pola	Nazwa formy podstawowej	cda	psx	fcda	fpsx	ncda	npsx	Złożoność pola
Action RPG								4	Cheat								7
	Action RPG	✓	✓	✓	✓	✓	✓			Cheat	✓	✓	✓	✓	✓	✓	
	Action-RPG	✓	✓	✓	✓	✓	✓			Cheater	✓	✓	✓	✓	✓	✓	
	RPG akcji	✓	✓	✓	✓	✓	✓			Cheaterskie	✓	✓	✓	✓	✓	✓	
	RPG-akcji	✗	✓	✗	✗	✗	✓			Cheaterstwo	✗	✗	✗	✗	✓	✗	
Ban								11	Cheatować	✓	✗	✓	✓	✓	✓	✗	
	Ban	✓	✓	✓	✓	✓	✓			Cheatowanie	✓	✓	✓	✓	✓	✗	
	Banować	✓	✓	✓	✗	✓	✓			Cheatowany	✗	✗	✗	✓	✗	✗	
	Banowalność	✗	✗	✓	✗	✗	✗		Checkpoint								2
	Banowalny	✗	✗	✓	✗	✗	✗			Checkpoint	✓	✓	✓	✓	✓	✓	
	Banowanie	✗	✓	✓	✓	✓	✓			Cheackpointowy	✗	✗	✗	✓	✗	✗	
	Banowany	✗	✓	✓	✓	✓	✓		Craft								10
	Banując	✗	✗	✓	✗	✗	✗			Craft	✓	✗	✓	✓	✓	✗	
	Banujący	✓	✗	✓	✗	✗	✗			Crafter	✗	✗	✓	✗	✓	✗	
	Zbanować	✓	✓	✓	✓	✓	✓			Crafting	✓	✓	✓	✓	✓	✓	
	Zbanowanie	✓	✓	✓	✓	✓	✓			Craftingowy	✓	✗	✓	✓	✗	✗	
	Zbanowany	✓	✓	✓	✓	✓	✓			Craftować	✓	✗	✗	✗	✗	✗	
Bijatyka								4	Craftowalny	✓	✗	✓	✗	✗	✗	✗	
	Bijatyka	✓	✓	✓	✓	✓	✓			Craftowanie	✓	✗	✓	✓	✓	✗	
	Bijatykowiec	✗	✗	✗	✓	✗	✗			Craftowy	✓	✗	✓	✓	✗	✗	
	Bijatykowy	✓	✓	✗	✓	✓	✓			Craftujący	✓	✗	✗	✗	✗	✗	
	Naparzanka	✓	✓	✓	✓	✓	✓			Scraftić	✗	✗	✓	✓	✗	✗	
Budżetówka								2	Crap								15
	Budżetowiec	✗	✗	✓	✓	✗	✗			Crap	✓	✓	✓	✓	✓	✓	
	Budżetówka	✓	✓	✓	✓	✓	✓			Crapcom	✗	✗	✗	✓	✗	✗	
Calak								17	Crapiarz	✗	✗	✗	✓	✗	✗	✗	
	Calak	✓	✓	✓	✓	✓	✓			Crapiasty	✗	✗	✗	✓	✗	✗	
	Calakować	✓	✓	✓	✓	✗	✗			Crapik	✗	✗	✗	✓	✗	✗	
	Calakowanie	✓	✓	✓	✓	✗	✗			Crapisko	✗	✗	✗	✓	✗	✗	
	Calakowany	✗	✓	✗	✗	✗	✗			Crapiszcz	✗	✗	✗	✓	✗	✗	
	Calakowy	✗	✗	✗	✓	✗	✗			Crapiszson	✗	✗	✗	✓	✗	✗	
	Calakując	✗	✓	✓	✗	✗	✗			Crapowatość	✗	✗	✗	✓	✗	✗	
	Calakujący	✗	✓	✗	✗	✗	✗			Crapowaty	✗	✓	✓	✓	✗	✗	
	Masterować	✓	✓	✓	✓	✓	✗			Crapowy	✗	✗	✗	✓	✗	✗	
	Masterowanie	✓	✓	✓	✓	✓	✓			Crapszot	✗	✗	✗	✓	✗	✗	
	Masterowany	✓	✓	✓	✓	✗	✗			Crapulec	✗	✗	✗	✓	✗	✗	
	Masterowicz	✗	✗	✗	✓	✗	✗			Krap	✗	✗	✓	✓	✓	✗	
	Masterowy	✗	✓	✗	✓	✗	✗			Krapiszcz	✓	✓	✗	✓	✓	✗	
	Masterując	✗	✓	✓	✓	✓	✗		Cross-controller								3
	Masterujący	✗	✓	✗	✓	✗	✗			Cross-control	✗	✗	✗	✗	✗	✓	
	Wymasterować	✓	✓	✗	✓	✓	✓			Cross-controler	✗	✗	✗	✗	✗	✓	
	Wymasterowanie	✗	✓	✗	✓	✓	✓			Cross-controller	✗	✓	✗	✗	✗	✓	
	Wymasterowany	✗	✓	✓	✓	✗	✗		Cross-gen								3
Casual								24	Cross-gen								3
	Casual	✓	✓	✓	✓	✓	✓			Cross-gen	✗	✗	✗	✓	✓	✗	
	Casualizacja	✓	✓	✓	✓	✓	✗			Cross-generacyjny	✗	✓	✗	✗	✗	✗	
	Casualka	✓	✗	✓	✗	✗	✗			Cross-genowy	✓	✓	✗	✓	✓	✓	
	Casualowanie	✗	✗	✗	✓	✗	✗		Cross-over								2
	Casualowość	✗	✗	✗	✓	✗	✗			Crossover	✓	✓	✓	✓	✓	✓	
	Casualowaty	✗	✗	✗	✓	✗	✗			Cross-over	✓	✓	✓	✓	✓	✓	
	Casualowiec	✗	✓	✗	✓	✓	✓		Cross-platform								4
	Casualowienie	✗	✗	✗	✓	✗	✗			Cross-platform	✗	✓	✗	✓	✓	✓	
	Casualowo	✗	✓	✗	✓	✗	✗			Cross-platforming	✗	✗	✗	✓	✗	✗	
	Casualowość	✗	✗	✗	✓	✓	✗			Cross-platformowość	✗	✓	✗	✗	✗	✗	
	Casualowy	✓	✓	✓	✓	✓	✓			Cross-platformowy	✗	✓	✓	✓	✓	✓	
	Casualówka	✗	✗	✗	✓	✗	✗		Current-gen								2
	Każual	✓	✓	✓	✓	✓	✗			Current-gen	✓	✓	✓	✓	✓	✓	
	Każualista	✗	✓	✗	✗	✗	✗			Current-genowy	✓	✓	✗	✓	✓	✓	
	Każualizacja	✓	✗	✓	✓	✓	✗		Cutscena								4
	Każualizm	✗	✗	✓	✗	✗	✗			Cutscena	✓	✓	✓	✓	✓	✗	
	Każualizowanie	✓	✗	✗	✗	✗	✗			Cutscenes	✓	✗	✗	✗	✗	✗	
	Każualność	✗	✗	✗	✗	✓	✗			Cutscenka	✓	✓	✓	✓	✓	✓	
	Każualowatość	✓	✗	✗	✗	✗	✗			Cutscenkowy	✓	✗	✗	✗	✓	✗	
	Każualowiec	✗	✗	✓	✗	✗	✗		Czarnula								2
	Każualowo	✓	✗	✓	✓	✗	✗			Czarnula	✗	✓	✗	✓	✓	✓	
	Każualowość	✗	✗	✓	✗	✓	✗			Czarnulka	✗	✓	✓	✓	✓	✓	
	Każualowy	✓	✗	✓	✓	✓	✗		Deathmatch								2
	Każualówka	✓	✗	✗	✗	✗	✗			Deathmatch	✓	✓	✓	✓	✓	✓	
		✓	✗	✗	✗	✗	✗			Deathmatchowy	✓	✓	✓	✗	✗	✓	
Celowniczek								2	Demo								2
	Celowniczek	✓	✓	✓	✓	✓	✓			Demko	✓	✓	✓	✓	✓	✓	
	Celowniczkowy	✓	✓	✓	✗	✓	✗			Demo	✓	✓	✓	✓	✓	✓	
Cel-shading								3	D-pad								3
	Cel-shading	✓	✓	✓	✓	✗	✓			D-pad	✓	✓	✓	✓	✓	✓	
	Cel-shadingowo	✓	✓	✗	✗	✗	✗			D-padowy	✗	✗	✗	✓	✗	✗	
	Cel-shadingowy	✓	✓	✗	✓	✗	✓			Krzyżak	✓	✓	✓	✓	✓	✓	

Nazwa pola	Nazwa formy podstawowej	cda	psx	fcda	fpsx	ncda	npsx	Złożoność pola	Nazwa pola	Nazwa formy podstawowej	cda	psx	fcda	fpsx	ncda	npsx	Złożoność pola
Dreamcast								5		Grindując	X	✓	X	✓	X	X	
	Dreamcast	✓	✓	✓	✓	✓	✓			Grindujący	✓	X	X	X	X	X	
	Dreamcastowy	✓	X	✓	✓	X	X			Grindzik	X	X	X	✓	X	X	
	Drim	X	X	X	✓	✓	X		Grywalny								7
	DC	✓	✓	X	✓	✓	✓			Grywalnie	X	X	✓	✓	✓	X	
	Makaron	✓	✓	✓	✓	X	X			Grywalnościowy	✓	✓	X	✓	X	✓	
Dungeon								3		Grywalność	✓	✓	✓	✓	✓	✓	
	Dungeon	✓	✓	✓	✓	✓	✓			Grywalny	✓	✓	✓	✓	✓	✓	
	Dungeonowanie	✓	X	✓	✓	X	X			Niegrywalnie	X	X	✓	✓	✓	✓	
	Dungeonowy	X	✓	X	X	X	X			Niegrywalność	✓	X	✓	✓	✓	✓	
Egranizacja								5		Niegrywalny	✓	✓	✓	✓	✓	✓	
	Egranizacja	✓	✓	✓	✓	✓	X		Handheld								6
	Egranizować	✓	X	X	X	X	X			Handheld	✓	✓	✓	✓	✓	✓	
	Egranizowanie	✓	X	X	X	X	X			Handheldowy	✓	✓	✓	✓	✓	✓	
	Zegraniżować	✓	X	X	X	X	X			Kieszonsolka	✓	✓	✓	✓	✓	✓	
	Zegraniżowany	✓	X	X	X	✓	X			Kieszonsolkowy	X	✓	X	✓	X	✓	
Emulator								3		Przenośka	X	✓	X	X	✓	X	
	Emlator	✓	✓	✓	✓	✓	✓			Przenośniak	✓	✓	X	✓	X	✓	
	Emulatorek	X	✓	X	✓	✓	X		HP								3
	Emulatorowy	X	X	X	✓	X	X			Hapek	✓	X	✓	X	✓	X	
Exclusive								7		HP	✓	✓	✓	✓	✓	✓	
	Ekskluziw	✓	✓	✓	✓	✓	X			HPk	X	X	✓	X	✓	X	
	Ekskluzyw	X	✓	✓	✓	X	✓		Hardkor								15
	Ekskluzywie	✓	✓	✓	✓	✓	✓			Ultrahardkor	✓	X	X	X	X	X	
	Ekskluzywny	✓	✓	✓	✓	✓	✓			Hardkor	✓	✓	✓	✓	✓	✓	
	Ekskulzywność	✓	✓	✓	✓	✓	✓			Hardkorok	X	X	X	X	✓	✓	
	Ex	X	✓	✓	✓	✓	✓			Hardkorować	X	✓	X	X	X	X	
	Exclusive	✓	✓	✓	✓	✓	✓			Hardkorowanie	X	✓	X	X	X	X	
Farmić								7		Hardkorowiec	✓	✓	✓	✓	✓	✓	
	Farmiąc	✓	X	X	✓	X	X			Hardkorowo	✓	✓	✓	✓	✓	X	
	Farmiący	✓	X	X	✓	X	X			Hardkorowość	✓	✓	✓	✓	✓	✓	
	Farmić	✓	✓	✓	✓	✓	X			Hardkorowy	✓	✓	✓	✓	✓	✓	
	Farmienie	✓	✓	✓	✓	✓	✓			Hardkorówa	X	✓	X	X	X	X	
	Farming	✓	✓	X	✓	X	X			Hardkorówka	X	✓	X	X	X	X	
	Farmingowy	X	X	X	✓	X	X			Hardkoryzacja	✓	✓	X	X	X	X	
	Farmiony	✓	X	X	X	X	X			Hardkoryzm	✓	✓	X	X	X	X	
Feature								3		Hardkorzasty	X	X	✓	X	X	X	
	Feature	X	✓	✓	✓	✓	✓			Hardkorzyć	✓	✓	X	X	X	X	
	Featurek	X	✓	X	X	✓	X		Hardware								3
	Ficzter	✓	✓	✓	✓	✓	✓			Hardware	✓	✓	✓	✓	✓	✓	
Freaa-to-play								3		Hardware'owo	✓	X	✓	✓	X	X	
	Free2play	✓	✓	✓	✓	✓	✓			Hardware'owy	✓	✓	✓	✓	✓	✓	
	Free-to-play	✓	✓	✓	✓	✓	✓		Headset								2
	Free-to-playowy	✓	X	X	✓	X	X			Headset	✓	✓	✓	✓	✓	✓	
Gamecube								8		Headsetowy	✓	X	X	X	X	X	
	Gacek	X	✓	✓	✓	X	✓		Healer								2
	Gackowy	X	X	X	✓	X	X			Healer	✓	✓	✓	✓	✓	X	
	GameCube	✓	✓	✓	✓	✓	✓			Healerka	✓	X	X	✓	X	X	
	Gamecube'owy	X	✓	X	✓	X	X		Hejt								16
	GC	✓	✓	✓	✓	✓	✓			Hejcić	✓	X	✓	X	✓	X	
	NGC	✓	✓	✓	✓	X	✓			Hejcik	X	X	X	✓	X	X	
	Kostka	✓	✓	✓	✓	X	✓			Hejt	✓	✓	✓	✓	✓	✓	
	Nocniczek	✓	X	X	X	X	X			Hejter	✓	✓	✓	✓	✓	✓	
Gameplay								3		Hejterek	X	X	X	✓	✓	X	
	Gameplay	✓	✓	✓	✓	✓	✓			Hejterka	✓	X	X	X	X	X	
	Gameplayowo	✓	✓	✓	✓	✓	✓			Hejterować	X	X	X	✓	X	X	
	Gameplayowy	✓	✓	✓	✓	✓	✓			Hejterski	✓	✓	✓	✓	✓	X	
Gaming								4		Hejterstwo	✓	✓	✓	✓	✓	X	
	Gaming	✓	✓	✓	✓	✓	X			Hejteryzm	X	X	X	✓	X	X	
	Gamingowo	✓	✓	X	X	X	X			Hejting	X	X	✓	✓	✓	✓	
	Gamingowość	✓	X	X	X	X	X			Hejtować	✓	✓	✓	✓	✓	✓	
	Gamingowy	✓	✓	✓	✓	✓	X			Hejtowanie	✓	✓	✓	✓	✓	X	
Geek								4		Hejtowany	✓	✓	✓	✓	✓	X	
	Geek	✓	✓	✓	✓	✓	✓			Hejtując	✓	✓	X	✓	X	X	
	Geekowaty	X	X	✓	X	X	X			Hejtujący	✓	✓	X	✓	✓	X	
	Geekowski	✓	✓	✓	✓	X	✓		High-end								3
	Geekowy	✓	✓	X	✓	X	X			High-end	✓	✓	✓	✓	✓	X	
Grind								12		High-endowiec	✓	X	X	X	X	X	
	Grind	✓	✓	✓	✓	✓	X			High-endowy	✓	✓	✓	✓	✓	✓	
	Grinder	X	✓	✓	X	✓	X		Indie								6
	Grinderski	✓	X	✓	✓	X	X			Indie	✓	✓	✓	✓	✓	✓	
	Grinding	✓	✓	✓	✓	X	✓			Indyczek	✓	X	✓	✓	✓	X	
	Grindować	✓	✓	✓	✓	✓	X			Indyczny	✓	X	✓	✓	✓	✓	
	Grindowalny	✓	X	X	✓	X	X			Indyk	✓	✓	✓	✓	✓	✓	
	Grindowanie	✓	✓	✓	✓	✓	✓			Indykowate	✓	X	X	X	X	X	
	Grindownik	X	X	✓	X	X	X			Indykowe	✓	X	✓	✓	X	✓	
	Grindowy	✓	✓	✓	✓	✓	✓										

Nazwa pola	Nazwa formy podstawowej	cda	psx	fda	fpsx	ncda	npsx	Złożoność pola	Nazwa pola	Nazwa formy podstawowej	cda	psx	fda	fpsx	ncda	npsx	Złożoność pola
JRPG								4	Multiplatformowy								5
	JRPG	✓	✓	✓	✓	✓	✓			Multiplatformowy	✓	✓	✓	✓	✓	✓	
	JRPG'owiec	✗	✗	✓	✗	✗	✗			Multiplatformowiec	✗	✗	✓	✓	✓	✗	
	JRPG-owość	✓	✗	✗	✗	✗	✗			Multiplatformowo	✗	✓	✗	✓	✓	✗	
	JRPG-owy	✓	✓	✓	✗	✗	✓			Multiplatformowość	✓	✓	✓	✓	✓	✓	
Kinect								2	Multiplatformówka		✗	✗	✗	✓	✗	✗	
	Kinect	✓	✓	✓	✓	✓	✓		Next-gen		✓	✓	✓	✓	✓	✓	5
	Kinectowy	✓	✓	✓	✓	✓	✓		Next-gen		✓	✓	✓	✓	✓	✓	
Klatkować								5	Next-genowiec		✓	✗	✗	✗	✗	✗	
	Klatkować	✓	✗	✓	✓	✓	✗		Next-genowo		✓	✓	✗	✓	✓	✗	
	Klatkowanie	✓	✗	✓	✓	✓	✗		Next-genowość		✓	✗	✓	✗	✓	✗	
	Klatkowany	✗	✓	✗	✗	✗	✗		Next-genowy		✓	✓	✓	✓	✓	✓	
	Klatkowy	✗	✗	✓	✓	✗	✗		Off-road		✓	✓	✓	✓	✓	✓	6
	Klatkujący	✓	✓	✓	✓	✓	✗		Offroad		✓	✗	✗	✓	✓	✗	
Klawiatura								5	Off-road		✓	✓	✓	✓	✓	✓	
	Klawiatura	✓	✗	✓	✓	✓	✗		Offroadowy		✓	✗	✗	✓	✗	✓	
	Klawiatura	✓	✓	✓	✓	✓	✓		Off-roadowy		✓	✓	✗	✗	✓	✓	
	Klawiaturka	✓	✓	✓	✓	✓	✗		Offroadówka		✓	✗	✗	✗	✗	✗	
	Klawiaturowy	✓	✓	✓	✓	✓	✗		Ofrold		✓	✗	✗	✗	✗	✗	
	Klawierka	✗	✗	✓	✓	✗	✗		Ośmiobitowy		✓	✓	✗	✓	✓	✗	2
Komputer								9	Ośmiobitowiec		✓	✓	✓	✓	✓	✓	
	Komp	✓	✓	✓	✓	✓	✗		Ośmiobitowy		✓	✓	✓	✓	✓	✓	
	Kompek	✗	✗	✓	✓	✓	✗		Padlet		✓	✓	✓	✓	✓	✓	2
	Kompik	✓	✗	✓	✓	✓	✗		Padlet		✓	✓	✓	✓	✓	✓	
	Komputer	✓	✓	✓	✓	✓	✓		Tableto-pad		✓	✓	✗	✓	✗	✗	
	Komputerek	✓	✓	✓	✓	✓	✗		Past-gen		✗	✓	✓	✓	✓	✓	2
	Komputerowiec	✓	✓	✓	✓	✓	✓		Past-gen		✗	✗	✗	✓	✗	✓	
	Komputerowo	✓	✓	✓	✓	✓	✓		Past-genowy		✗	✗	✗	✓	✗	✓	
	Komputerowy	✓	✓	✓	✓	✓	✓		Patch		✓	✓	✓	✓	✓	✓	20
	Komputeryzacja	✓	✓	✗	✗	✓	✗		Patch		✓	✓	✓	✓	✓	✓	
Konsola								6	Patcher		✗	✗	✗	✗	✓	✗	
	Konsola	✓	✓	✓	✓	✓	✓		Patchować		✓	✓	✓	✓	✓	✗	
	Konsolka	✓	✓	✓	✓	✓	✓		Patchowanie		✗	✗	✓	✓	✓	✗	
	Konsolowiec	✓	✓	✓	✓	✓	✓		Patchowany		✓	✓	✓	✓	✗	✗	
	Konsolowo	✗	✓	✓	✓	✓	✗		Patchując		✗	✓	✗	✗	✗	✗	
	Konsolowość	✓	✗	✗	✓	✓	✗		Patchujący		✓	✓	✗	✗	✗	✗	
	Konsolowy	✓	✓	✓	✓	✓	✓		Pacz		✗	✓	✓	✓	✓	✗	
Liniowy								3	Paczer		✗	✗	✗	✗	✓	✗	
	Liniowo	✓	✓	✓	✓	✓	✗		Paczowanie		✗	✗	✗	✓	✗	✗	
	Liniowość	✓	✓	✓	✓	✓	✓		Paczowany		✗	✗	✗	✓	✗	✗	
	Liniowy	✓	✓	✓	✓	✓	✓		Paczyk		✗	✗	✗	✓	✗	✗	
Live-stream								3	Spaczować		✗	✗	✓	✓	✗	✗	
	Live-stream	✗	✗	✗	✓	✗	✓		Spaczowanie		✗	✗	✓	✓	✗	✗	
	Live-streamowany	✗	✗	✗	✗	✗	✓		Spaczowany		✗	✗	✓	✓	✗	✗	
	Live-streamowy	✗	✗	✗	✗	✗	✓		Spatchować		✓	✓	✓	✓	✓	✗	
Loading								3	Spatchowanie		✓	✗	✓	✓	✓	✗	
	Loading	✓	✓	✓	✓	✓	✓		Spatchowany		✓	✓	✓	✓	✓	✓	
	Loadingować	✗	✗	✗	✓	✗	✗		Łatka		✓	✓	✓	✓	✓	✓	
	Loadingowy	✓	✓	✗	✗	✗	✗		Łatkowanie		✓	✗	✗	✓	✗	✗	
Marian								3	Pauza		✓	✓	✓	✓	✓	✓	6
	Marian	✗	✓	✓	✓	✓	✓		Pauza		✓	✓	✓	✓	✓	✓	
	Marianek	✗	✗	✗	✓	✗	✗		Pauzować		✓	✓	✓	✓	✓	✓	
	Marianowy	✗	✗	✗	✓	✗	✗		Pauzowanie		✓	✓	✓	✓	✗	✗	
Mikrotransakcja								2	Pauzowany		✓	✓	✗	✓	✗	✗	
	Mikrotransakcja	✓	✓	✓	✓	✓	✓		Pauzując		✗	✗	✗	✗	✓	✗	
	Mikrotransakcyjny	✗	✗	✗	✓	✗	✗		Pauzujący		✓	✗	✗	✗	✗	✗	
Minimapa								2	PC		✓	✓	✓	✓	✓	✓	20
	Minimapa	✓	✓	✓	✓	✓	✗		PC		✓	✓	✓	✓	✓	✓	
	Minimapka	✓	✗	✗	✓	✗	✗		PC-t		✓	✓	✓	✓	✓	✓	
MMORPG								4	PC-towiec		✓	✗	✓	✓	✓	✗	
	MMO RPG	✗	✗	✓	✓	✓	✗		PC-towy		✓	✓	✓	✓	✓	✓	
	MMORPG	✓	✓	✓	✓	✓	✓		Pececiarz		✓	✓	✓	✓	✓	✗	
	MMO-RPG	✓	✗	✓	✗	✓	✗		Pececik		✗	✗	✓	✓	✗	✗	
	MMORPG-owy	✓	✓	✗	✓	✗	✗		Pecet		✓	✓	✓	✓	✓	✓	
Multi								6	Pecetowiec		✓	✓	✓	✓	✓	✓	
	Multi	✓	✓	✓	✓	✓	✓		Pecetowo		✓	✗	✗	✗	✗	✗	
	Multiplayer	✓	✓	✓	✓	✓	✓		Pecetowość		✓	✓	✗	✗	✗	✗	
	Multiplayerowiec	✗	✗	✓	✓	✗	✗		Pecetowy		✓	✓	✓	✓	✓	✓	
	Multiplayerowość	✓	✗	✗	✗	✗	✗		Piec		✓	✓	✓	✓	✓	✓	
	Multiplayerowy	✓	✓	✓	✓	✓	✓		Piecet		✗	✓	✗	✗	✗	✗	
	Multiplayerówka	✗	✓	✗	✓	✗	✗		Piecowiec		✗	✗	✗	✓	✓	✗	
Multimedia								4	Piecowy		✓	✓	✓	✓	✓	✗	
	Multimedia	✓	✓	✓	✓	✓	✓		Piecyk		✓	✓	✓	✓	✓	✗	
	Multimedialnie	✓	✗	✗	✗	✗	✗		Piecykowy		✗	✗	✓	✓	✓	✗	
	Multimedialność	✓	✓	✗	✓	✓	✗		Błaszak		✓	✓	✓	✓	✓	✓	
	Multimedialny	✓	✓	✓	✓	✓	✓		Błaszakowiec		✗	✓	✗	✓	✓	✓	
									Błaszakowy		✓	✓	✓	✓	✓	✓	

Nazwa pola	Nazwa formy podstawowej	cda	psx	fcda	fpsx	ncda	npsx	Złożoność pola	Nazwa pola	Nazwa formy podstawowej	cda	psx	fcda	fpsx	ncda	npsx	Złożoność pola
Pesczwórka								4	PSP go								2
	Peesztery	X	X	X	✓	X	X			PSP go	✓	✓	X	✓	✓	✓	
	Peesczwórka	✓	✓	X	X	✓	X			PSPgo	✓	✓	X	✓	X	✓	
	Peesczwórkowiec	✓	X	X	X	X	X		Quest								8
	Peesczwórkowy	X	✓	X	X	✓	X		Quest	✓	✓	✓	✓	✓	✓	✓	
Peesdwójka								3	Questing	✓	X	X	X	X	X	X	
	Peesdwójka	✓	X	X	X	X	X		Questować	✓	X	✓	✓	✓	X	X	
	Peesdwójkowy	✓	✓	✓	✓	✓	✓		Questowanie	✓	X	✓	✓	✓	✓	X	
	Pees dwójka	X	X	X	✓	X	X		Questowo	X	X	X	✓	✓	X	X	
Peesiks								3	Questowy	✓	✓	✓	✓	✓	X	✓	
	Peesiks	X	✓	X	X	X	X		Questując	✓	X	X	X	X	X	X	
	Szarak	✓	✓	✓	✓	X	✓		Questujący	✓	X	X	X	X	X	X	
	Szarakowy	X	✓	X	✓	X	X		Randomizacja								3
Peesplusowy								3	Randomizacja	X	X	✓	✓	✓	X	X	
	Peesplusowicz	✓	✓	X	X	X	X		Randomizować	X	X	X	✓	✓	X	X	
	Peesplusowiec	X	X	X	X	✓	X		Randomizowany	✓	X	X	✓	✓	✓	X	
	Peesplusowy	X	✓	X	X	✓	X		Real-time								2
Peestrójka								6	Real-time	✓	✓	✓	✓	✓	✓	X	
	Peestójkowiec	X	X	✓	X	X	X		Real-time'owy	X	✓	X	X	X	X	X	
	Peestrójka	✓	✓	X	✓	✓	X		Redzi								3
	Peestrójkowicz	X	X	X	✓	X	X		Redowski	✓	X	X	X	X	X	X	
	Peestrójkowy	✓	✓	X	✓	X	✓		Redowy	✓	X	X	X	X	X	X	
	Pees-trójkowy	X	X	X	X	✓	X		Redzi	✓	✓	✓	✓	✓	✓	X	
	Chlebak	X	✓	✓	✓	✓	✓		Remake								4
Pelniak								3	Remake	✓	✓	✓	✓	✓	✓	✓	
	Pelniak	✓	✓	✓	✓	✓	X		Remakeowanie	✓	X	X	X	✓	✓	X	
	Pelniakowo	X	X	✓	X	X	X		Rimejk	✓	✓	✓	✓	✓	✓	✓	
	Pelniakowy	X	X	✓	X	X	X		Rimejkowanie	✓	X	X	X	X	X	X	
Platyna								7	Remaster								9
	Platyna	✓	✓	✓	✓	✓	✓		Remaster	✓	✓	✓	✓	✓	✓	✓	
	Platynka	X	✓	X	✓	✓	✓		Remasterek	✓	X	X	✓	✓	X	X	
	Platynować	X	X	X	✓	X	✓		Remastering	✓	✓	✓	✓	X	X	X	
	Platynowanie	X	✓	✓	✓	X	X		Remasterowanie	✓	✓	✓	✓	✓	✓	✓	
	Platynowany	X	X	X	✓	X	X		Remasterowany	X	✓	✓	✓	✓	✓	X	
	Platynowi	X	✓	✓	✓	X	✓		Remasterując	✓	✓	X	X	X	X	X	
	Platynowy	✓	✓	✓	✓	✓	✓		Zremasterować	✓	✓	✓	✓	✓	✓	X	
PlayStation								16	Zremasterowanie	✓	✓	X	✓	✓	X	✓	
	PlayStation	✓	✓	✓	✓	✓	✓		Zremasterowany	✓	✓	✓	✓	✓	✓	✓	
	Playstationowiec	✓	X	X	X	X	X		Roguelike								4
	Playstationowy	✓	✓	✓	X	X	X		Rogalik	✓	X	✓	X	✓	✓	X	
	Plej	✓	✓	✓	✓	✓	X		Rogalikowy	✓	X	X	X	✓	✓	X	
	Plejacz	X	X	X	✓	X	X		Roguelike	✓	✓	✓	✓	✓	✓	✓	
	Plejak	X	X	✓	✓	✓	X		Roguelike'owy	✓	X	X	X	X	✓	X	
	Plejakować	X	X	X	X	✓	X		Rozdzielczość								2
	Plejeczka	X	X	X	✓	X	X		Rozdziałka	✓	✓	✓	✓	✓	✓	✓	
	Plejka	X	✓	✓	✓	X	✓		Rozdzielczość	✓	✓	✓	✓	✓	✓	✓	
	Plejowy	X	X	X	✓	X	X		RPG								10
	Plejstacja	X	X	✓	✓	X	X		RPG	✓	✓	✓	✓	✓	✓	✓	
	Plejstejszon	X	X	✓	✓	✓	X		RPG-owanie	X	X	✓	X	X	X	X	
	Plejstejszonowy	X	X	X	✓	X	X		RPG-owiec	X	✓	✓	✓	✓	✓	X	
	Plejstejszyn	✓	X	✓	✓	✓	X		RPG-owy	✓	✓	✓	✓	✓	✓	✓	
	Peeska	X	X	✓	X	X	X		Erpeg	✓	✓	✓	✓	✓	✓	✓	
	Peesowy	X	X	✓	X	X	X		Erpegować	X	X	✓	X	✓	✓	X	
Pre-alpha								2	Erpegowanie	X	X	✓	X	✓	✓	X	
	Pre-alfa	✓	✓	✓	X	✓	✓		Erpegowiec	✓	✓	✓	✓	✓	X	X	
	Pre-alpha	✓	✓	✓	✓	✓	✓		Erpegowość	X	X	X	✓	X	X	X	
Pre-oreder								4	Erpegowy	✓	✓	✓	✓	✓	✓	✓	
	Pre-order	✓	✓	✓	✓	✓	✓		Samograj								4
	Pre-orderowicz	✓	X	X	X	X	X		Samograj	✓	✓	✓	✓	✓	✓	✓	
	Pre-orderowiec	✓	✓	X	✓	✓	✓		Samogrające	X	X	✓	X	X	X	X	
	Pre-orderowy	✓	✓	✓	X	✓	✓		Samograjka	✓	X	X	X	X	X	X	
Przeglądarka								3	Samograjowy	X	✓	X	X	X	X	X	
	Przeglądarka	✓	✓	✓	✓	✓	✓		Sandbox								11
	Przeglądarkowy	✓	✓	✓	✓	✓	✓		Sandboks	✓	✓	✓	✓	✓	✓	✓	
	Przeglądarkówka	✓	X	X	X	X	X		Sandboksik	✓	X	X	X	X	X	X	
Przygodówka								2	Sandboksowanie	✓	X	X	X	X	X	X	
	Przygodówka	✓	✓	✓	✓	✓	✓		Sandboksowość	✓	X	✓	X	X	X	X	
	Przygodówkowy	✓	✓	✓	✓	✓	✓		Sandboksowy	✓	✓	✓	✓	✓	✓	✓	
PSP								4	Sandbox	✓	✓	✓	✓	✓	✓	✓	
	PSP	✓	✓	✓	✓	✓	✓		Sandboxowo	X	X	✓	X	X	X	X	
	PSP'owiec	X	X	✓	X	X	X		Sandboxowość	X	X	X	✓	X	X	X	
	Peespek	✓	X	X	✓	X	X		Sandboxowy	X	X	✓	✓	✓	✓	X	
	Peespekowy	X	X	X	✓	X	X		Piaskownica	✓	✓	✓	✓	✓	✓	✓	
PSP2								2	Piaskownicowy	✓	X	X	X	X	X	X	
	PSP 2	✓	X	X	✓	✓	✓										
	PSP2	✓	✓	X	X	✓	✓										

Nazwa pola	Nazwa formy podstawowej	cda	psx	fcda	fpsx	ncda	npsx	Złożoność pola	Nazwa pola	Nazwa formy podstawowej	cda	psx	fcda	fpsx	ncda	npsx	Złożoność pola
Save								11		Streamingujący	✓	✗	✗	✗	✗	✗	
	Save	✓	✓	✓	✓	✓	✓			Streamować	✓	✓	✓	✓	✓	✓	
	Save'ować	✗	✓	✓	✗	✗	✗			Streamowanie	✓	✓	✓	✓	✓	✓	
	Save'owanie	✗	✓	✗	✗	✗	✗			Streamowany	✓	✓	✓	✓	✓	✓	
	Savepoint	✓	✓	✓	✗	✗	✓			Streamowy	✗	✗	✓	✓	✓	✗	
	Save-point	✓	✓	✗	✗	✗	✓			Streamując	✓	✓	✗	✓	✗	✗	
	Sejw	✓	✓	✓	✓	✓	✓			Streamujący	✓	✓	✗	✗	✗	✗	
	Sejwować	✓	✓	✓	✓	✓	✗		Strzelanina								6
	Sejwowanie	✓	✓	✗	✓	✗	✗			Strzelanina	✓	✓	✓	✓	✓	✓	
	Sejwujący	✗	✗	✗	✓	✗	✗			Strzelanka	✓	✓	✓	✓	✓	✓	
	Zasejwować	✗	✗	✓	✓	✓	✗			Strzelankowy	✓	✓	✓	✓	✓	✓	
	Zasejwowanie	✗	✗	✗	✓	✗	✗			Shooter	✓	✓	✓	✓	✓	✓	
Sequel								3		Shooterek	✗	✓	✓	✓	✗	✗	
	Sequel	✓	✓	✓	✓	✓	✓			Shooterowy	✓	✓	✗	✓	✓	✓	
	Sequelizm	✓	✗	✗	✗	✗	✗		Szesnastobitowy								2
	Sequelowy	✓	✓	✗	✗	✗	✓			Szesnastobitowiec	✓	✗	✗	✗	✓	✗	
Shmup								2		Szesnastobitowy	✗	✗	✗	✓	✓	✗	
	Shmup	✓	✓	✓	✓	✓	✓		Ścigalka								3
	Shmupowy	✓	✗	✗	✗	✗	✗			Ścigalka	✓	✓	✓	✓	✓	✓	
Sieciówka								2		Ścigalkowy	✓	✓	✗	✗	✓	✗	
	Sieciówka	✓	✓	✓	✓	✓	✗		Tablet								5
	Sieciówkowy	✗	✗	✗	✗	✓	✗			Tablecik	✗	✗	✗	✓	✗	✗	
Singiel								10		Tablet	✓	✓	✓	✓	✓	✓	
	Singiel	✓	✓	✓	✓	✓	✓			Tabletek	✗	✗	✗	✗	✓	✗	
	Singleplajer	✗	✗	✗	✓	✗	✗			Tabletkowy	✗	✗	✗	✓	✓	✗	
	Singleplayer	✓	✓	✓	✓	✓	✓			Tabletowy	✗	✗	✓	✗	✓	✗	
	Single-player	✓	✓	✓	✓	✓	✓		Tank								7
	Singleplayerowy	✓	✓	✓	✗	✓	✓			Tank	✓	✓	✓	✓	✓	✓	
	Single-playerowy	✗	✗	✓	✓	✗	✗			Tanker	✗	✗	✗	✓	✗	✗	
	Singlowiec	✓	✓	✓	✗	✓	✗			Tankista	✗	✗	✗	✗	✓	✗	
	Singlowo	✓	✗	✗	✓	✗	✗			Tankować	✓	✗	✗	✓	✓	✗	
	Singlowość	✓	✗	✗	✗	✗	✗			Tankowanie	✓	✗	✓	✓	✓	✗	
	Singlowy	✓	✓	✓	✓	✓	✓			Tankowaty	✗	✗	✓	✗	✗	✗	
Skin								2		Tankujący	✓	✗	✗	✗	✗	✗	
	Skin	✓	✓	✓	✓	✓	✓		Tekstówka								2
	Skórka	✓	✓	✓	✓	✓	✓			Tekstówka	✓	✓	✓	✓	✓	✗	
Skrypt								5		Tekstówkowy	✓	✗	✗	✗	✗	✗	
	Skrypt	✓	✓	✓	✓	✓	✓		Tekstura								10
	Skryptować	✗	✓	✓	✗	✓	✗			Tekstura	✓	✓	✓	✓	✓	✓	
	Skryptowanie	✓	✓	✓	✓	✓	✗			Teksturka	✗	✗	✓	✓	✓	✗	
	Skryptowany	✓	✓	✓	✓	✓	✓			Teksturować	✓	✗	✗	✗	✗	✗	
	Skryptowy	✓	✓	✓	✓	✗	✗			Teksturowanie	✓	✓	✓	✗	✗	✗	
Smartfon								2		Teksturowany	✓	✓	✗	✗	✗	✗	
	Smartfon	✓	✓	✓	✓	✓	✓			Teksturowy	✓	✓	✗	✓	✗	✗	
	Smartfonowy	✓	✓	✓	✓	✓	✓			Teksturujący	✓	✗	✗	✗	✗	✗	
Software								3		Oteksturować	✓	✓	✗	✓	✗	✗	
	Software	✓	✓	✓	✓	✓	✓			Oteksturowanie	✓	✓	✗	✓	✗	✗	
	Software'owo	✓	✗	✗	✓	✗	✓			Oteksturowany	✓	✓	✓	✓	✓	✗	
	Software'owy	✓	✓	✗	✓	✓	✓		Touchpad								4
Spawn								17		Touch pad	✗	✗	✗	✗	✗	✓	
	Spawn	✓	✓	✓	✓	✓	✗			Touchpad	✓	✓	✓	✓	✓	✓	
	Spawner	✗	✗	✓	✗	✓	✗			Myziadelko	✓	✗	✗	✗	✗	✗	
	Spawnowanie	✓	✓	✓	✓	✓	✗			Gładzik	✓	✗	✓	✗	✗	✗	
	Spawnowany	✓	✗	✗	✗	✗	✗		Trailer								3
	Spawujący	✓	✓	✗	✗	✗	✗			Trailer	✓	✓	✓	✓	✓	✓	
	Spwanować	✓	✓	✓	✓	✓	✓			Trailerek	✗	✓	✓	✓	✓	✗	
	Zrespawnić	✗	✗	✗	✗	✓	✗			Trailerowy	✗	✗	✓	✓	✓	✓	
	Zrespawnować	✓	✗	✗	✓	✓	✗		Uzywka								2
	Zrespawnowanie	✓	✗	✗	✗	✗	✗			Uzywka	✓	✓	✓	✓	✓	✓	
	Zrespawnowany	✗	✗	✗	✓	✗	✗			Używkowy	✗	✗	✓	✗	✓	✗	
	Respawn	✓	✓	✓	✓	✓	✗		WSAD								3
	Respawner	✓	✗	✗	✗	✗	✗			WASD	✓	✗	✗	✓	✓	✗	
	Respawnować	✓	✓	✓	✓	✓	✗			WSAD	✓	✗	✓	✓	✓	✗	
	Respawnowanie	✓	✗	✗	✓	✓	✗			WSAD-owy	✓	✗	✗	✗	✗	✗	
	Respawnowy	✓	✗	✓	✓	✗	✗		Xbox								8
	Respawnując	✗	✓	✓	✗	✗	✗			Xboks	✓	✓	✓	✓	✓	✓	
	Respawnujący	✓	✓	✓	✓	✓	✗			Xboksiak	✗	✗	✗	✓	✗	✗	
Sprite								2		Xboksowiec	✓	✓	✓	✓	✓	✗	
	Sprajt	✓	✓	✗	✓	✗	✗			Xboksowy	✓	✓	✓	✓	✓	✓	
	Sprite	✓	✓	✓	✓	✓	✓			Xbox	✓	✓	✓	✓	✓	✓	
Stream								13		Xkloc	✗	✓	✓	✓	✓	✗	
	Stream	✓	✓	✓	✓	✓	✓			Xklocek	✓	✓	✓	✓	✓	✗	
	Streamer	✓	✓	✓	✓	✓	✓			Xklockowy	✗	✗	✗	✓	✗	✗	
	Streamerka	✗	✓	✗	✗	✗	✗		Zapychacz								2
	Streamerski	✗	✗	✗	✗	✗	✓			Zapychacz	✓	✓	✓	✓	✓	✓	
	Streaming	✓	✓	✓	✓	✓	✓			Zapychaczowy	✗	✓	✗	✗	✗	✗	
	Streamingowy	✓	✓	✓	✓	✓	✓										

Nazwa pola	Nazwa formy podstawowej	cda	psx	feda	fpsx	ncda	npsx	Złożoność pola	Nazwa pola	Nazwa formy podstawowej	cda	psx	feda	fpsx	ncda	npsx	Złożoność pola
Bumper								1	Przedłużajka								1
	Bumper	✓	✓	✗	✓	✓	✗			Przedłużajka	✓	✗	✗	✗	✗	✗	
Cooldown								1	Przejściówka								1
	Cooldown	✓	✓	✓	✓	✓	✓			Przejściówka	✓	✓	✓	✓	✓	✓	
Creditsy								1	Przeszkadzajka								1
	Creditsy	✓	✓	✓	✓	✓	✓			Przeszkadzajka	✓	✓	✓	✓	✓	✓	
Cross-buy								1	Przetrwaniówka								1
	Cross-buy	✓	✓	✗	✓	✓	✓			Przetrwaniówka	✓	✗	✗	✗	✗	✗	
Cross-chat								1	PvE								1
	Cross-chat	✓	✗	✗	✓	✗	✗			PvE	✓	✓	✓	✓	✓	✓	
Cross-play								1	PvP								1
	Cross-play	✓	✓	✗	✓	✓	✓			PvP	✓	✓	✓	✓	✓	✓	
Cross-save								1	QTE								1
	Cross-save	✗	✓	✗	✓	✓	✓			QTE	✓	✓	✓	✓	✗	✓	
Devkit								1	QWERTY								1
	Devkit	✓	✓	✓	✓	✓	✗			QWERTY	✓	✓	✓	✓	✗	✗	
DLC								1	Reedycja								1
	DLC	✓	✓	✓	✓	✓	✓			Reedycja	✓	✓	✓	✓	✓	✓	
Dokanałówki								1	Router								1
	Dokanałówki	✓	✓	✓	✓	✗	✗			Router	✓	✓	✓	✓	✓	✗	
Eksbokslan								1	RROD								1
	Eksbokslan	✓	✗	✗	✗	✗	✗			RROD	✗	✓	✓	✓	✓	✓	
Flashówka								1	Samochodówka								1
	Flashówka	✓	✗	✓	✓	✓	✗			Samochodówka	✓	✓	✓	✓	✓	✓	
Framerate								1	Shader								1
	Framerate	✓	✓	✓	✓	✓	✓			Shader	✓	✓	✓	✓	✓	✓	
GotY								1	Side-quest								1
	GotY	✓	✓	✓	✓	✓	✓			Side-quest	✓	✓	✓	✓	✓	✗	
Grzybek								1	Soundtrack								1
	Grzybek	✓	✓	✓	✓	✓	✗			Soundtrack	✓	✓	✓	✓	✓	✓	
Jednostrzałowiec								1	Stacjonarka								1
	Jednostrzałowiec	✗	✓	✓	✓	✗	✗			Stacjonarka	✓	✓	✓	✓	✓	✓	
Kartridż								1	Stodwudziestośmiobitowiec								1
	Kartridż	✓	✓	✓	✓	✓	✓			Stodwudziestośmiobitowiec	✓	✗	✗	✗	✗	✗	
Kontroler								1	Stylus								1
	Kontroler	✓	✓	✓	✓	✓	✓			Stylus	✓	✓	✓	✓	✓	✓	
Konwersja								1	Sub-boss								1
	Konwersja	✓	✓	✓	✓	✓	✓			Sub-boss	✗	✓	✗	✓	✗	✓	
Lokalizacja								1	Sub-quest								1
	Lokalizacja	✓	✓	✓	✓	✓	✓			Sub-quest	✗	✓	✓	✓	✗	✗	
LPM								1	Subskrypcja								1
	LPM	✓	✗	✓	✓	✓	✗			Subskrypcja	✓	✓	✓	✓	✓	✓	
Matryca								1	Three-sixty								1
	Matryca	✓	✓	✓	✓	✓	✓			Three-sixty	✓	✗	✗	✗	✗	✗	
Metagra								1	Trigger								1
	Metagra	✓	✗	✗	✗	✓	✗			Trigger	✓	✓	✓	✓	✓	✓	
Multikart								1	Turówka								1
	Multikart	✗	✓	✗	✗	✗	✗			Turówka	✓	✓	✓	✓	✓	✓	
Multitap								1	Twarzowanie								1
	Multitap	✓	✓	✗	✓	✗	✗			Twarzowanie	✗	✓	✗	✗	✗	✗	
Multitasking								1	Unboxing								1
	Multitasking	✓	✓	✓	✓	✓	✓			Unboxing	✓	✓	✓	✓	✓	✓	
Multitouch								1	Wiesiek								1
	Multitouch	✓	✓	✓	✓	✓	✗			Wiesiek	✓	✓	✓	✓	✓	✓	
Must-have								1	Wtyczka								1
	Must-have	✓	✓	✓	✓	✓	✓			Wtyczka	✓	✓	✓	✓	✓	✗	
Niedzielnik								1	YLOD								1
	Niedzielnik	✓	✓	✓	✓	✓	✓			YLOD	✗	✓	✓	✓	✓	✗	
Notebook								1	Zblurowany								1
	Notebook	✓	✓	✓	✓	✓	✓			Zblurowany	✗	✓	✓	✓	✗	✗	
NPC								1	Znajdźka								1
	NPC	✓	✓	✓	✓	✓	✓			Znajdźka	✓	✓	✓	✓	✓	✓	
Odblokowajka								1	Żywotność								1
	Odblokowajka	✓	✗	✗	✗	✗	✗			Żywotność	✓	✓	✓	✓	✓	✓	
One-liner								1									
	One-liner	✓	✓	✓	✓	✓	✓										
Pad								1									
	Pad	✓	✓	✓	✓	✓	✓										
Pay-to-win								1									
	Pay-to-win	✓	✓	✗	✓	✓	✗										
Podstawka								1									
	Podstawka	✓	✓	✓	✓	✓	✓										
PPM								1									
	PPM	✓	✗	✓	✓	✓	✗										

Powyższe zestawienie składa się ze 189 elementów, które odpowiadają polom wyrazowym, które badam w tej pracy. Na przykład dane dla pola *budżetówka* wyglądają następująco:

Tabela 25. Wykaz wystąpień pola *budżetówka* oraz jego form podstawowych w danych częściach korpusu

Nazwa pola	Nazwa formy podstawowej	cda	psx	fed	fpx	neda	npx	Złożoność pola
Budżetówka								2
	Budżetowiec	X	X	✓	✓	X	X	
	Budżetówka	✓	✓	✓	✓	✓	✓	

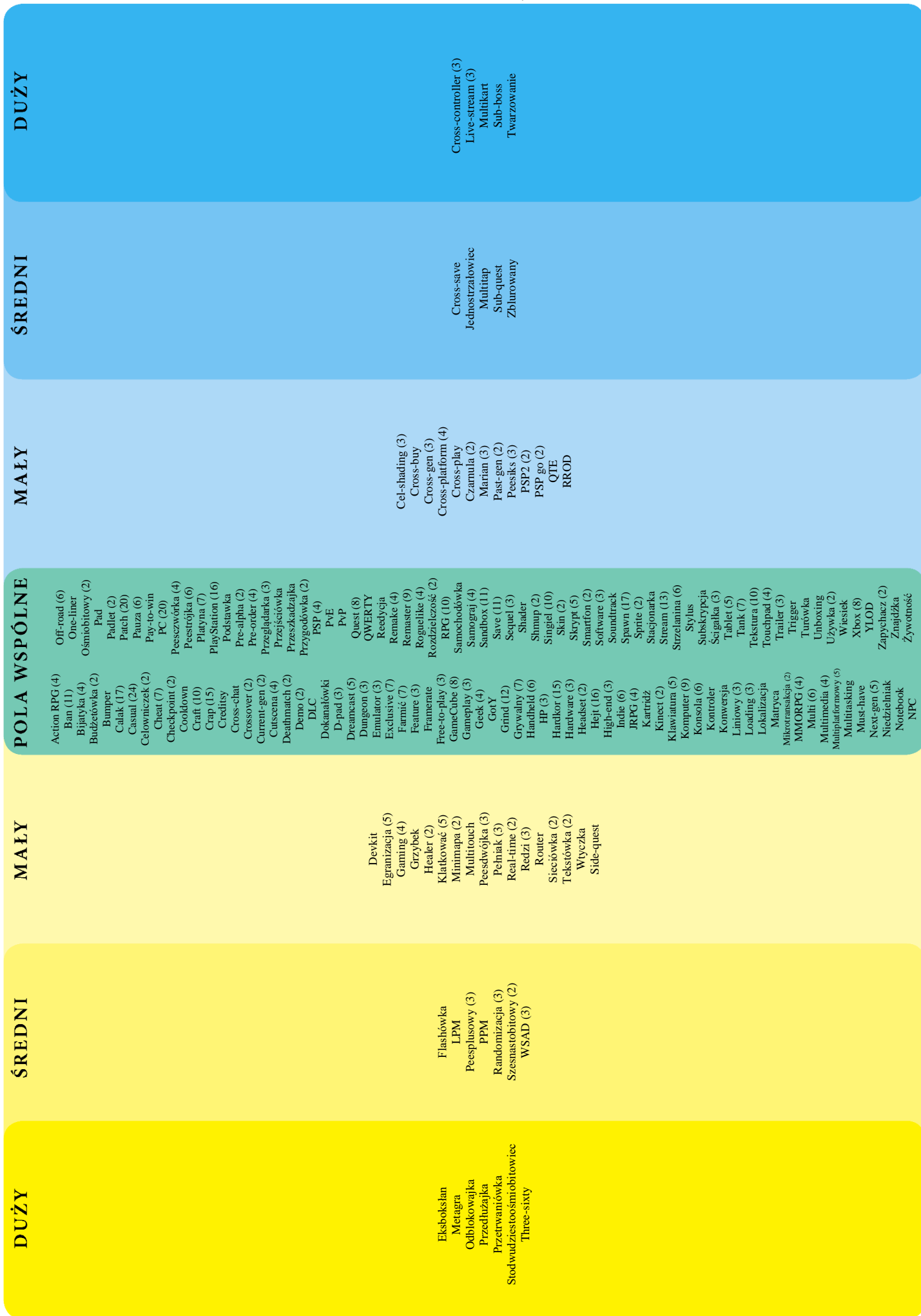
Nazwa powyższego pola, identycznie jak we wcześniejszej wersji tabeli, umieszczona została w kolumnie po lewej stronie – *budżetówka*. W skład pola wchodzi dwie formy podstawowe – *budżetowiec* i *budżetówka*. W tabeli podałem także dane na temat tego, w których korpusach odnotowane zostały poszczególne formy (znaczek ✓), a w których ich zabrakło (X). Natomiast w kolumnie po prawej stronie podałem złożoność danego pola, czyli informacje o tym, ile jest w nim uwzględnionych form podstawowych, w tym przypadku jest ich 2.

Do tej pory, aby stwierdzić czy dane pole wyrazowe związane jest z *konsolowcami* lub *komputerowcami* używałem systemu 1.0. Jednak na tym etapie pracy, trzeba zadać kolejne pytanie badawcze, a mianowicie, czy na podstawie samej obecności pola wyrazowego w danej grupie korpusów związanych z danym środowiskiem, można stwierdzić, że jest ono ważne dla danego środowiska, a dla innego już nie? Otóż wydaje mi się, że stosując powyższy system, który oczywiście dostarczył odpowiedzi zgodnych z rzeczywistością, opartych na zebranych i przeanalizowanych danych, to jednak moim zdaniem nie będzie to odpowiedź pełna. Uważam, że aby dostarczyć bardziej rzetelnej odpowiedzi na pytanie o ważność danego pola w badanej grupie socjolektalnej, trzeba uwzględnić jeszcze jeden aspekt, a mianowicie złożoność pól.

Moim zdaniem im bardziej rozbudowane jest dane pole wyrazowe, tym jest ono ważniejsze dla środowiska. Rozbudowane pole świadczy o tym, że gracze na tyle dobrze przyswoili jego formy podstawowe, że nie stanowi dla nich problemu zabawa nim, a tym samym mogą tworzyć kolejne formy wyrazowe. Dlatego na grafikę przedstawioną na stronie 168, trzeba nałożyć dane dotyczące złożoności pól. Im złożoność będzie większa, tym pole ważniejsze dla graczy, im mniejsza tym mniej ważne.

Poniżej przedstawiam zmodyfikowaną grafikę ze strony 168:

GRAFIKA 10. STOPIEŃ POWIĄZANIA PÓL WYRAZOWYCH Z DANYM ŚRODOWISKIEM (WRAZ Z PODANIEM ZŁOŻONOŚCI)



Grafika została zmodyfikowana w ten sposób, że dodałem po nazwie pola, w nawiasach zwykłych, informację o jego złożoności. Informacje te podałem jedynie w przypadkach, gdy dane pole posiadała w sobie minimum dwie formy podstawowe, natomiast gdy pole było jednoelementowe, to pomijałem tę informację, czyli nie umieszczałem informacji w nawiasach.

W badaniu socjolektu graczy wideo zaobserwować można, że pola najbardziej złożone to te, które zawierają 10 lub więcej form podstawowych. Wyniki pokazują, że wszystkie tego typu pola znalazły się w obszarze zielonym. Jest ich łącznie 18:

casual, patch, PC, calak, spawn, hejt, PlayStation, crap, hardkor, stream, grind, ban, sandbox, save, craft, RPG, singiel, tekstura.

Przyglądając się polom, które zawierają od 5 do 9 form podstawowych, można zauważyć, że większość z nich, bo aż 25 wystąpiła również w obszarze zielonym:

komputer, remaster, GameCube, quest, Xbox, cheat, exclusive, farmić, grywalny, platyna, tank, handheld, indie, konsola, multi, off-road, pauza, Peestrójka, strzelanina, Dreamcast, klawiatura, multiplatformowy, next-gen, skrypt, tablet.

Tylko dwa pola wystąpiły w obszarze żółtym:

egranizacja, klatkować.

Analizując pola, które zawierają od 2 do 4 form podstawowych zauważyć można, że większość z nich także umieszczona została w obszarze zielonym. Odnotowano ich 51:

action RPG, bijatyka, cutscena, gaming, geek, jRPG, MMORPG, multimedia, Pesczwórka, pre-oreder, PSP, remake, roguelike, samograj, touchpad, d-pad, dungeon, emulator, feature, free-to-play, gameplay, HP, hardware, high-end, liniowy, live-stream, loading, przeglądarka, sequel, software, ścigalka, trailer, budżetówka, celowniczek, checkpoint, crossover, current-gen, deathmatch, demo, headset, Kinect, mikrotransakcja, ośmiobitowy, padlet, pre-alpha, przygodówka, rozdzielczość, shmup, skin, smartfon, sprite.

W obszarach żółtych znalazło się 13 takich pól:

peesplusowy, randomizacja, szesnastobitowy, WSAD, gaming, healer, minimapa, Peesdwójka, pełniak, Redzi, real-time, sieciówka, tekstówka,

a w niebieskich tylko 11 pól:

Marian, cel-shading, cross-gen, cross-controller, cross-platform, czarnula, past-gen, PSP go, PSP2, Peesiks, live-stream.

Reasumując, najwięcej rozbudowanych pól wyrazowych znajduje się w obszarze zielonym. Po raz kolejny w tej pracy potwierdza się to, że ważniejsze dla graczy są te pola, które są wspólne zarówno *konsolowcom*, jak i *komputerowcom*.

15. POLA LEKSYKALNE A SUMY WYRAZÓW GRAFICZNYCH

Do tej pory przyglądałem się polom znaczeniowym pod kątem nasycenia wyrazami socjolektalnymi tekstów (czasopism, newsów i forów), a pomijałem bardziej szczegółowe dane liczbowe. Czas więc przyjrzeć się polom bardziej dokładnie, czyli z uwzględnieniem danych na temat użycia wyrazów graficznych, które je budują. Trzeba w tym momencie zmienić skalę postrzegania badanego materiału z binarnego (przydatnego do stwierdzania, czy dane pole wystąpiło w korpusie czy też nie było ono tam obecne) na postrzeganie decymalne – dające większe pole manewru i interpretacji.

15.1. ZESTAWIENIE NAJBOGATSZYCH I NAJUBOŻSZYCH W WYRAZY GRAFICZNE PÓL LEKSYKALNYCH – ANALIZA POSZCZEGÓLNYCH CZĘŚCI KORPUSU PLATFORMY ZIELONEJ

Aby omówić kolejne zagadnienia przewidziane w tej pracy, niezbędne okażą się dane liczbowe, które umieszczone zostały w tabelach (Z1–Z6) w aneksie na stronach 23–28.

Zaprezentowane zostały tam pola z podziałem na dane korpusy, jedna tabela zawiera więc dane na temat wszystkich pól, ale tylko w jednym korpusie, a mówiąc dokładniej, każda z tabel zawiera dokładne dane liczbowe na temat sumy wszystkich wyrazów graficznych, które fizycznie wystąpiły w badanych częściach korpusu. Zostały one ułożone począwszy od tego pola, które posiada największą sumę wyrazów do tego, które posiada ich najmniej.

W części tej przyjrę się wybranym, moim zdaniem najważniejszym aspektom, które wynikają z analizy powyższych tabel.

Analizę rozpocznę od zestawienia 10 pól, które posiadają największą liczbę wystąpień wyrazów graficznych w każdym z korpusów. Do omówienia zagadnienia posłuży mi poniższe zestawienie, które powstało z fragmentów powyżej przedstawionych tabel.

	NPSX	NCDA	FPSX	FCDA	PSX	CDA
1	PLAYSTATION	PLAYSTATION	KONSOLA	KOMPUTER	KONSOLA	PC
2	PC	XBOX	PC	PC	XBOX	KONSOLA
3	XBOX	PC	MULTI	KONSOLA	RPG	KOMPUTER
4	KONSOLA	KOMPUTER	GAMEPLAY	RPG	PC	STRZELANINA
5	PSP	KONSOLA	RPG	PEŁNIAK	MULTI	RPG
6	DLC	DLC	TRAILER	KLAWIATURA	KOMPUTER	TABLET

7	GAMEPLAY	TRAILER	DLC	MULTI	PAD	MULTI
8	TRAILER	GAMEPLAY	DEMO	TABLET	STRZELANINA	SMARTFON
9	KINECT	MULTI	PAD	PAD	KONTROLER	ROZDZIELCZOŚĆ
10	DEMO	RPG	XBOX	DLC	GAMEPLAY	PAD

Tabela 26. Zestawienie pól najbogatszych w wyrazy graficzne (platforma zielona)

Jak widać, we wszystkich korpusach wystąpiły tylko pola *konsola i PC*. Wyciągnąć można wniosek, że są one najważniejsze dla całego środowiska graczy wideo. Pozostałe pola w powyższym zestawieniu nie wystąpiły we wszystkich korpusach.

Dzieląc korpusy na 3 grupy ze względu na sposób publikacji tekstów (newsy, fora i czasopisma), stwierdzić można, że dla newsów charakterystyczne, pomijając oczywiście pola *PC i konsola*, są wyrazy z pól, które odnotowane zostały w obu korpusach z tekstami newsów. Są to pola:

DLC, gameplay, PlayStation, trailer, xbox

dla wpisów na forach internetowych:

DLC, multi, pad, RPG

natomiast dla tekstów z czasopism:

komputer, multi, pad, RPG, strzelanina

Przyglądając się 10 najobszerniejszym polom wyrazowym, przez pryzmat podziału wszystkich korpusów na dwie grupy, czyli na korpusy *konsolowców* (psx, npsx, fpsx) i korpusy *komputerowców* (cda, ncda, fcda), stwierdzić można, że poza polami *konsola i PC*, które występują we wszystkich korpusach, dla środowiska *konsolowców* charakterystyczne są jeszcze pola *gameplay i Xbox*, natomiast dla komputerowców *multi, komputer i RPG*.

Zestawiając dane z powyższej tabeli z informacjami, które zaprezentowane zostały na grafice ze strony 178, zauważyć można, że w tabeli znalazły się prawie wyłącznie pola, które zostały umieszczone w kolumnie zielonej na wspomnianej już grafice, czyli w obszarze *pola wspólne*. Jest tak dlatego, że pola, które występują w obu środowiskach, zarówno konsolowym, jak i komputerowym, są częściej używane, bo i większa jest potencjalna liczba użytkowników tegoż słownictwa. Dlatego polom, które przypisane są do danego środowiska, trudniej jest uzyskać większą lub tak samo dużą popularność jak tym, które znane są wszystkim graczom. Niemniej jednak pole charakterystyczne dla danego środowiska z czasem może rozszerzyć swoje występowanie i stać się równie popularne w obu grupach. W innym scenariuszu może być odwrotnie, czyli pole może jeszcze bardziej zawęzić obszar swojego występowania lub zaniknąć. Dlatego ważne jest, aby obserwować zmiany w zakresie występowania pól i badać je

cyklicznie. Wtedy będą widoczne długofalowe tendencje. Tyczy się to każdego słownictwa środowiskowego (pól znaczeniowych danych środowisk). Dzięki takim obserwacjom można uchwycić moment, gdy pola/wyrazy zwiększają swoją częstotliwość.

Szacuję, że słownictwo graczy będzie zyskiwać na popularności, ze względu na popularyzację gier wśród dzieci i młodzieży. Część badanych przeze mnie pól w przyszłości będzie częścią polszczyzny ogólnej, a największą szansę mają na to pola umieszczone w powyższej tabeli.

Ze wszystkich pól, które wystąpiły w zestawieniu, a było ich łącznie 20 (pomijając te powtarzające się), aż 19 z nich to te z obszaru zielonego na grafice 10 (strona 178), natomiast pozostałe jedno pole – *pełniak*, umieszczone zostało w obszarze jasnożółtym, czyli jest to pole tylko w stopniu *małym* związane z *komputerowcami*. Pole to wystąpiło we wszystkich korpusach z wyjątkiem npsx.

Patrząc na powyższą tabelę zauważyć można jeszcze jedną rzecz, a mianowicie że lista dziesięciu najobszerniejszych pól zawiera po jednym polu w każdym wierszu w tabeli. Nie było w zestawieniu więc takich pól, które mogłyby zająć ex aequo jakieś miejsce na liście, gdyż zawierałyby tyle samo wyrazów graficznych.

Inaczej sprawa wygląda z polami, które znalazły się na liście najuboższych pól wyrazowych, czyli tych które reprezentowane są przez najmniej wyrazów graficznych w analizowanych tekstach. W tym podrozdziale wykorzystam te same tabele (w aneksie na stronach od 23 do 28), które posłużyły do analizy zaprezentowanej w poprzednim podrozdziale.

Poniżej prezentuję zestawienie najuboższych wyrazowo pól:

	NPSX	NCDA	FPSX	FCDA	PSX	CDA
10	UŻYWK CALAK CZARNULA CROSS-SAVE	SHMUP WTYCZKA	ROGUELIKE MULTITAP	PRE-ALPHA	PVE EGRANIZACJA SHMUP REAL-TIME	FRAMERATE PPM
9	HP	MARIAN CHECKPOINT MUST-HAVE PPM FLASHÓWKA PAY-TO-WIN	PPM CROSS-PLATFORM	REAL-TIME RANDOMIZACJA	MINIMAPA OŚMIOBITOWY PAY-TO-WIN UNBOXING	CRAP CREDITSY
8	LINIOWY ŻYWOTNOŚĆ STYLUS CROSS-PLATFORM CROSS-CONTROLLER	CREDITSY LPM GRZYBEK OŚMIOBITOWY	PEESDWÓJKA SUB-BOSS	GEEK CEL-SHADING	ROUTER TANK ROGUELIKE PRE-ALPHA OFF-ROAD MULTITAP CROSS-SAVE	CROSS-OVER NIEDZIELNIAK DOKANAŁÓWKI
7	CRAFT D-PAD STACJONARKA HIGH-END PAST-GEN	PRZEJŚCIÓWKA RROD OFF-ROAD	JEDNOSTRZAŁOWIEC CELOWNICZEK	DREAMCAST MARIAN SIDE-QUEST MUST-HAVE	GRZYBEK COOLDOWN PSP2 ONE-LINER CROSS-CONTROLLER	GOTY PEESTRÓJKA FLASHÓWKA

6	PAUZA GEEK MUST-HAVE QTE	SAMOGRAJ PEESTRÓJKA PADLET MINIMAPA WSAD CROSS-PLAY PEESCZWÓRKA PEESPLUSOWY	LPM	PRZESZKADZAJKA SHMUP	SUB-QUEST	PLATYNA PEESIKS ZAPYCHACZ
5	EMLATOR NOTEBOOK ROGUELIKE SHADER OFF-ROAD	TRIGGER STYLUS HEALER PRZESZKADZAJKA COOLDOWN SZESNASTOBITOWY TEKSTÓWKA	SZESNASTOBITOWY	FREE-TO-PLAY CROSS-PLATFORM OŚMIOBITOWY	MULTITOUCH WTYCZKA FEATURE TEKSTÓWKA BUDŻETÓWKA	CURRENT-GEN CROSS-PLAY PEESCZWÓRKA ODBLOKOWAJKA
4	SAMOCODÓWKA RROD TRIGGER	CZARNULA ZAPYCHACZ PSP GO DEVKIT CROSS-PLATFORM	RANDOMIZACJA MULTITOUCH FLASHÓWKA BUMPER CROSS-SAVE	CELOWNICZEK DEVKIT ONE-LINER	PEESCZWÓRKA	PADLET ONE-LINER
3	GRIND FEATURE PRE-ALPHA SHMUP CHECKPOINT PRZEJŚCIÓWKA CEL-SHADING	REAL-TIME YLOD CROSS-GEN CROSS-SAVE	EGRANIZACJA GRZYBEK PAY-TO-WIN	PEESTRÓJKA MULTITOUCH FLASHÓWKA	KLATKOWAĆ PAST-GEN JEDNOSTRZAŁOWIEC	STACJONARKA SZESNASTOBITOWY RANDOMIZACJA
2	TURÓWKA SPRITE ZNADŹKA CREDITSY SAMOGRAJ PEESTRÓJKA PRZESZKADZAJKA NIEDZIELNIAK	ONE-LINER MULTITOUCH BUMPER	TEKSTÓWKA OŚMIOBITOWY PEESCZWÓRKA CROSS-CHAT LIVE-STREAM	PADLET CURRENT-GEN YLOD OFF-ROAD PAST-GEN DOKANAŁÓWKI PEESDWÓJKA	PEŁNIAK CROSS-GEN PEESPLUSOWY MULTIKART	CROSS-BUY MUST-HAVE BUDŻETÓWKA CROSS-GEN REAL-TIME PEESDWÓJKA PRZETRWANIÓWKA
1	HEJT WIESIEK SPAWN TANK PVE FARMIC MATRYCA BUDŻETÓWKA OŚMIOBITOWY COOLDOWN ZAPYCHACZ CROSS-GEN ONE-LINER MULTITASKING SUB-BOSS	PAST-GEN MULTITASKING SIDE-QUEST NIEDZIELNIAK PEESDWÓJKA RANDOMIZACJA METAGRA	ZBLUROWANY	PEESIKS STYLUS CZARNULA MULTITASKING SUB-QUEST ZBLUROWANY	HEALER REDZI QWERTY DOKANAŁÓWKI PEESDWÓJKA ZBLUROWANY TWARZOWANIE	UNBOXING LPM PEESPLUSOWY DEVKIT MULTITAP CROSS-CHAT EKSBOKSLAN PRZEDŁUŻAJKA STODWUDZIESTOŚMIOBI- TOWIEC THREE-SIXTY

Tabela 27. Zestawienie pól najuboższych w wyrazy graficzne (platforma zielona)

Powyższa tabela jest co najmniej kilkakrotnie większa od tabeli zawierającej najbogatsze pola. Tendencja ta, czyli przewaga wyrazów z pól ubogich wyrazowo nad bogatymi, widoczna była już w pracach J. Sambor. Również w mojej analizie znajduje to potwierdzenie, ponieważ przewaga tego rodzaju jest także widoczna.

Analizując tabelę dostrzec można, że ani jedno pole nie jest wspólne wszystkim korpusom.

Natomiast zestawiając ze sobą korpusy newsów (npsx i ncda), zauważyć już można pola wspólne. W obu korpusach odnotowałem takie pola, jak:

multitasking, must-have, niedzielniak, off-road, one-liner, ośmiobitowy, past-gen, Peestrójka, przejściówka, przeszkadzajka, RROD, samograj, shmup, stylus, trigger, zapychacz, checkpoint, cooldown, creditsy, cross-gen, cross-platform, cross-save, czarnula.

W korpusach zawierających wpisy z forów internetowych (fpsx i fcda) odnotowano następujące pola, które wspólne są dla obu korpusów:

cross-platform, flashówka, multitouch, Peesdwójka, randomizacja, zblurowany.

W wyniku analizy korpusów czasopism (cda i psx) znaleziono takie wspólne pola, jak:

budżetówka, cross-gen, dokanałówki, multimap, one-liner, Peeschwórka, Peesdwójka, peesplusowy, real-time, unboxing.

Przyrównując dane z powyższej tabeli do informacji przedstawionych na grafice 10, widać, że w grupie pól najskromniej reprezentowanych w tekstach (najmniej wystąpień graficznych) znalazły się prawie wszystkie pola z obszarów żółtych i niebieskich, czyli te pola, które związane są z *konsolowcami* lub *komputerowcami*. W tabeli zabrakło tylko dwóch pól z obszaru jasnożółtego – *gaming* i *sieciówka*, oznacza to, że pola mimo delikatnego przechylenia w stronę *komputerowców* są tak często używane, że zabrakło ich w powyższym zestawieniu. Przypuszczać można, że pola te w kolejnych latach zyskają na popularności i zaliczone zostaną do obszaru zielonego. Wystąpiły one we wszystkich korpusach z wyjątkiem npsx. Ich brak w obszarze zielonym (wyrazów wspólnych) można wytłumaczyć także tym, że pola w rzeczywistości występują we wszystkich rodzajach publikacji, ale ze względu na losowy dobór newsów do badań, nie wybrano tych, w których były one obecne. Ale ten pogląd zweryfikują już inne, kolejne badania.

Reasumując, w tabeli z polami ubogimi wyrazowo znalazło się 52 z 54 wszystkich pól, które na grafice zostały zaklasyfikowane jako te, które związane są ze środowiskiem *konsolowców* lub *komputerowców*. Zabrakło tylko pól *gaming* i *sieciówka* z obszaru jasnożółtego.

Poza tymi polami w powyższej tabeli znalazło się także 65 ze 135 pól, które są wspólne obu grupom graczy (wystąpiły one w obszarze zielonym).

Zestawiając razem informacje podane w tabelach z najbogatszymi i najuboższymi polami i odnosząc się do grafiki 10 stwierdzić należy, że pola najobfitsze wyrazowo, to bez mała zawsze te z obszaru zielonego (pola charakterystyczne obu środowiskom graczy wideo), natomiast pola najskromniejsze wyrazowo, to prawie wszystkie z obszarów żółtych i niebieskich (obszarów przypisanych *konsolowcom* lub *komputerowcom*) oraz w przybliżeniu połowa wszystkich pól z obszaru zielonego (pola wspólnego).

Dzięki zestawieniu dostrzec można jeszcze, że zarówno w grupie pól najobszerniejszych, jak i najuboższych wyrazowo wystąpiło jedno pole wspólne tym zestawieniom. Jest to *pełniak* (pole charakterystyczne dla obszaru jasnożółtego). Wynika

z tego, że pole to ma największe wahania pod względem użycia jego wyrazów w tekstach konkretnych korpusów, ponieważ raz jest zaliczane do pól najbogatszych (w korpusie fcda), innym razem do najuboższych (w korpusie psx).

Poprzez to zestawienie widać także, że łącznie w grupie najbogatszych pól, także z tymi powtarzającymi się w różnych korpusach, znalazło się 60 na 1134 możliwe (bo w jednym korpusie może ich być maksymalnie 189 pól, a korpusów jest 6), zaś w tabeli z polami najuboższymi aż 234 (także na 1134 możliwe).

Po usunięciu pól zdublowanych okazało się, że w grupie najbardziej bogatych wyrazowo pól znalazło się ich 21 (na 189 możliwych), natomiast w grupie najbiedniejszych aż 118 na 189 możliwych. Dane te potwierdzają, że pól najbiedniejszych wyrazowo jest najwięcej w zestawieniu wszystkich korpusów.

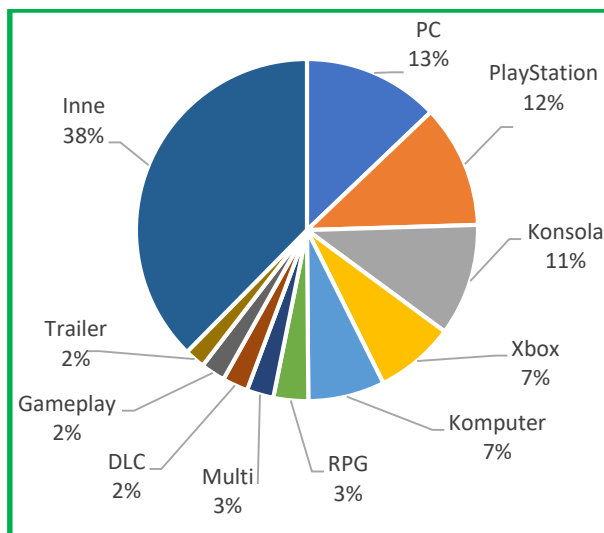
15.2. SUMA NAJBOGATSZYCH I NAJUBOŻSZYCH W WYRAZY GRAFICZNE PÓL LEKSYKALNYCH – ANALIZA CAŁEGO KORPUSU PLATFORMY ZIELONEJ

Do tej pory analizowałem jedynie zestawienia fragmentów tabel odnoszących się do poszczególnych korpusów, w następnej części pracy trzeba więc przyjrzeć się jednej większej tabeli, która powstała w wyniku zsumowania danych ilościowych przedstawionych w tabelach od Z1 do Z6. Tabela ta zaprezentowana została a w aneksie (tabela Z7) na stronie 29.

Tabela została zbudowana identycznie jak te, które dały podstawę do jej powstania, czyli jest to prezentacja pól wyrazowych ułożona od pól najbogatszych do tych, które są najmniej obfite w wyrazy.

Po zsumowaniu ze wszystkich części korpusu danych widać, że pole *PC* jest na pozycji pierwszej, czyli jest to najbogatsze pole pod względem użycia wyrazów graficznych we wszystkich korpusach.

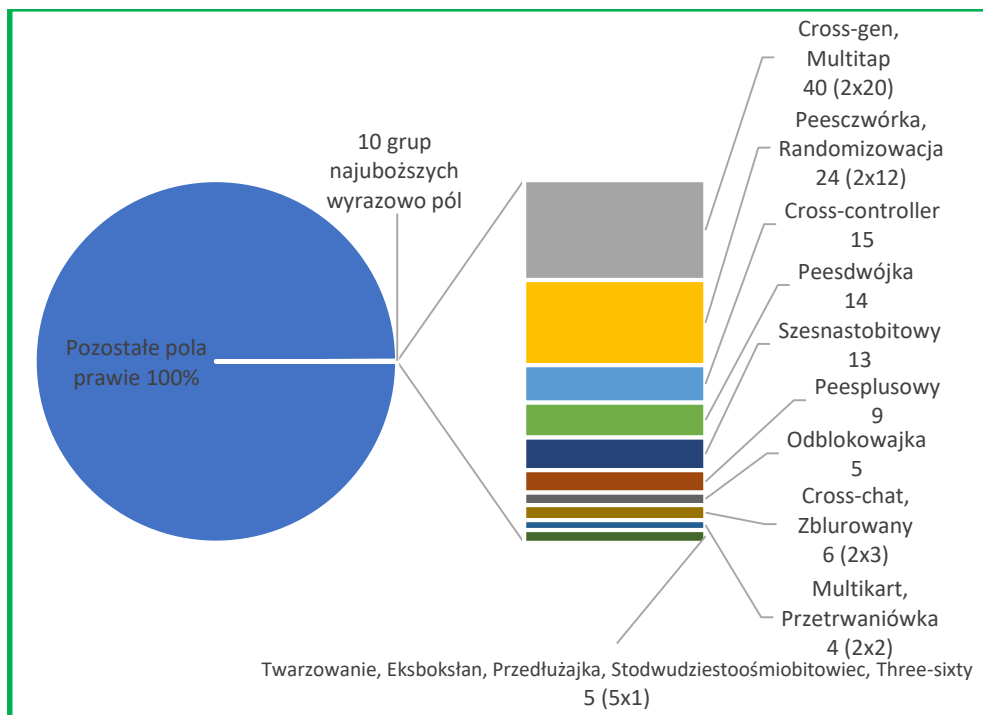
W celu lepszego zobrazowania i pokazania, jak najobszerniejsze wyrazowo pola prezentują się na tle pozostałych załączam poniższy wykres:



Wykres 14. Suma pól najbogatszych w wyrazy graficzne – zestawienie procentowe (platforma zielona)

Z wykresu wynika, że pola najobszerniejsze wyrazowo zajmują prawie 2/3 całego, zbadanego w tej pracy, zasobu wyrazów graficznych. Natomiast pięć pierwszych pól na wykresie to aż połowa wszystkich wyrazów graczy (PC 13% + $PlayStation$ 12% + $konsola$ 11% + $xbox$ 7% + $komputer$ 7%).

Z kolei wykres obrazujący rozłożenie wyrazów z pól najrzadszych wygląda następująco:



Wykres 15. Suma pól najuboższych w wyrazy graficzne – zestawienie procentowe (platforma zielona)

Wyrazy z pól najuboższych, jak widać na powyższym wykresie, nie zajmują w przybliżeniu nawet 1 procenta wszystkich zbadanych wyrazów graczy. Analizując dokładne liczby użyć wyrazów z tych pól widać, że łącznie było ich 135, co stanowi 0,03104% wszystkich zbadanych wyrazów graczy. Dlatego też, aby powyższy wykres był bardziej czytelny, w jego prawej części zrezygnowałem z prezentacji procentowej, ponieważ wszędzie musiałbym wpisać 0%, a zamiast tego podałem realną liczbę wystąpień danych pól z uwzględnieniem tych, które dane miejsce zajmują *ex aequo*. Przed nawiasem podałem łączną liczbę wystąpień wyrazów graficznych z konkretnych pól, natomiast w nawiasach działanie, które doprowadziło do powstania tej liczby, np. na poziomie 10 umieszczone zostały pola *cross-gen* i *multitap*, które łącznie zawierają 40 wyrazów, natomiast pojedynczo po 20.

Przyrównując dane z tabeli do informacji umieszczonych na grafice widać, że wszystkie pola najbogatsze wyrazowo umieszczone zostały w zielonej kolumnie graficznego zestawienia. Informacja ta jest nieco inna od tej, która podana została podczas analizy zestawień tabel z każdego korpusu. Wtedy do zestawienia włączone zostały także dwa pola z obszaru żółtego, natomiast teraz ich nie ma. Wynika z tego, że pola te są popularne tylko w konkretnych korpusach, natomiast po zsumowaniu danych ich znaczenie zmalało i teraz nie stanowią już pól najbogatszych pod względem użycia ich wyrazów.

W przypadku pól najbiedniejszych wyrazowo, tylko dwa z nich znalazły się w obszarze zielonym (wspólnym obu środowiskom) – *peesczwórka* i *cross-chat*, natomiast reszta to pola z obszarów niebieskich (6 pól) i żółtych (10 pól).

Analizując dokładną liczbę użyć wyrazów z danych pól zaprezentowanych na grafice i odnosząc się do danych w zsumowanej tabeli widać, że wyrazy z pól, które znalazły się w żółtych kolumnach wystąpiły łącznie 8.873 razy, co stanowi 2,04% wszystkich zbadanych wyrazów, natomiast te z obszarów niebieskich zostały użyte w tekstach 2.534 razy. Stanowi to 0,58% wszystkich wyrazów graczy. Natomiast wyrazów z obszaru zielonego odnotowano 423.517, co stanowi 97,38% wszystkich wyrazów graczy. Wynika z tego, że wyrazy związane z danym środowiskiem graczy wideo (*konsolowcami* i *komputerowcami*) stanowią tylko niewielki procent wszystkich wyrazów graczy. Dominują pola zawierające wyrazy wspólne tym środowiskom (te z obszaru zielonego).

W celu dopełnienia części pracy, która poświęcona została analizie najobszerniejszych i najuboższych wyrazowo pól platformy zielonej, trzeba jeszcze przyjrzeć się materiałowi, uwzględniając dwa kryteria podziału, czyli kryterium użytkownika (podział na korpusy *konsolowców* i *komputerowców*) oraz kryterium pochodzenia tekstu (podział na korpusy

zawierające teksty newsów, czasopism oraz wpisów na forach internetowych). Najpierw przyjrę się najobszerniejszym polom w grupie korpusów *konsolowców* oraz *komputerowców*.

**15.3. ZESTAWIENIE NAJBOGATSZYCH I NAJUBOŹSZYCH W WYRAZY GRAFICZNE PÓL
LEKSYKALNYCH – ANALIZA Z UWZGLĘDNIENIEM PODZIAŁU NA KORPUSY *KONSOLOWCÓW*
ORAZ *KOMPUTEROWCÓW* PLATFORMY ZIELONEJ**

Do omówienia tytułowego zagadnienia posłużę się tabelami Z8 i Z9, które załączone zostały w aneksie do pracy na stronach 31 i 32. Jej wyimek prezentuję poniżej:

	KONSOLOWCY	KOMPUTEROWCY
1	KONSOLA	PC
2	PLAYSTATION	KOMPUTER
3	PC	PLAYSTATION
4	XBOX	KONSOLA
5	GAMEPLAY	XBOX
6	MULTI	RPG
7	PSP	MULTI
8	DLC	DLC
9	RPG	TABLET
10	TRAILER	STRZELANINA

Tabela 28. Zestawienie pól najbogatszych w wyrazy graficzne w korpusach *konsolowców* oraz *komputerowców* (platforma zielona)

Porównując dwie kolumny powyższej tabeli widać, że listy dziesięciu najobszerniejszych wyrazowo pól w korpusach *konsolowców* oraz *komputerowców* są w dużej mierze ze sobą zbieżne, a wynika to z tego, że aż 7 tych samych pól wystąpiło na obu listach.

W korpusach *konsolowców*, poza siedmioma wspólnymi polami, znalazły się jeszcze takie, jak: *gameplay*, *PSP* i *trailer*, których zabrakło na liście *komputerowców*. Natomiast tylko u *komputerowców* odnotowano pola: *komputer*, *tablet* i *strzelanina*.

Wśród trzech unikalnych pól z listy *konsolowców* najmniej dziwi obecność *PSP*, ponieważ jest to nazwa konsoli, która była bardzo popularna w badanym okresie. Można więc było przypuszczać, że pole związane z tą konsolą znajdzie się na tej liście, czego potwierdzeniem są wyniki badań. Natomiast obecność dwóch pozostałych pól – *gameplay* i *trailer* świadczyć może jedynie o tym, że dla *konsolowców* są one ważnym elementem growej rzeczywistości. Ich duże bogactwo wyrazowe trudno wyjaśnić, odwołując się do rzeczywistości pozajęzykowej.

Przyglądając się unikalnym polom z listy *komputerowców* widać, że bardzo wysoko na liście znalazło się pole *komputer*. Przyrównując ten fakt do tego, że w obu kolumnach tabeli znalazło się pole *konsola*, wysnuć można wniosek, że *komputerowcy* częściej piszą na temat konsol niż *konsolowcy* na temat komputerów, zatem grupie *komputerowców* bliższe są konsole niż *konsolowcom* komputery.

Interesujące jest także wystąpienie pola *strzelanina* wyłącznie w zestawieniu u *komputerowców*. Wynioskować należy, że ten gatunek gier preferowany jest głównie przez tę jedną część środowiska graczy wideo. Znajduje to też swoje odzwierciedlenie w rzeczywistości. Najpopularniejsza *strzelania* wśród graczy, czyli Counter-Strike pierwotnie ukazała się wyłącznie na komputery i głównie z nimi jest kojarzona, ponieważ to na nich rozgrywane są zawody e-sportowe w tę grę. Wybór przez deweloperów komputera jako docelowej platformy dla wydawanych strzelanin podyktowany jest większą funkcjonalnością obsługi komputera dla potrzeb gry.

Przyjrzyjmy się zatem polom najuboższym wyrazowo z uwzględnieniem tytułowego kryterium:

	KONSOLOWCY	KOMPUTEROWCY
10	EGRANIZACJA OŚMIOBITOWY	PADLET DOKANAŁÓWKI JEDNOSTRZAŁOWIEC
9	PAY-TO-WIN GRZYBEK	CROSS-PLAY PEESCZWÓRKA
8	JEDNOSTRZAŁOWIEC MULTITOUCH PPM	ONE-LINER
7	PEESDWÓJKA	CROSS-PLATFORM DEVKIT
6	TEKSTÓWKA	SZESNASTOBITOWY
5	LPM PEESCZWÓRKA	PEESIKS PEESPLUSOWY
4	SZESNASTOBITOWY FLASHÓWKA	CZARNULA YŁOD CROSS-GEN PEESDWÓJKA ODBLOKOWAJKA
3	RANDOMIZACJA	PAST-GEN CROSS-SAVE
2	PEESPLUSOWY CROSS-CHAT ZBLUROWANY MULTIKART	PRZETRWANIÓWKA
1	TWARZOWANIE	SUB-QUEST MULTITAP CROSS-CHAT ZBLUROWANY EKSBOOKSŁAN PRZEDŁUŻAJKA STODWUDZIESTOŚMIOBITOWIEC THREE-SIXTY

Tabela 29. Zestawienie pól najuboższych w wyrazy graficzne w korpusach *konsolowców* oraz *komputerowców* (platforma zielona)

Już po pierwszym spojrzeniu na tabelę widać, że lista najuboższych wyrazowo pól w grupie tekstów *komputerowców* jest znacznie większa niż analogiczne zestawienie u *konsolowców*. Świadczy to o tym, że większą różnorodność pod względem użycia wyrazów graficznych wykazują teksty graczy konsolowych.

W obu kolumnach tabeli, czyli zarówno u *konsolowców*, jak i *komputerowców* odnotowano takie wspólne pola jak: *cross-chat*, *jednostrzałowiec*, *Peeschwórka*, *Peesdwójka*, *peesplusowy*, *szesnastobitowy* oraz *zblurowany*. Są to więc najuboższe wyrazowo pola, które wyłamują się podziałom na dwie grupy graczy wideo. Co się z tym wiąże, pola te posiadają najmniejszą liczbę wyrazów graficznych, więc szansa znalezienia ich w tekstach jest znikoma.

Reszta pól w powyższym zestawieniu tabelarycznym to pola unikalne dla danej grupy graczy (np. *eksbokslan* – pole charakterystyczne dla *komputerowców*, *twarzowanie* pole związane z *konsolowcami*) lub pola bardzo rzadkie w jednej grupy graczy, natomiast obszernie reprezentowane pod względem wyrazów graficznych w drugiej, np. *egranizacja* u *komputerowców* czy *peesiks* u *konsolowców*. Pole *egranizacja* odznacza się małą sumą wyrazów graficznych w korpusach *konsolowców*, dlatego też znalazło się na liście najuboższych pól w tej grupie graczy, natomiast u *komputerowców* jest ono dosyć bogate w wyrazy graficzne. Z analogiczną sytuacją mamy do czynienia w przypadku pola *peesiks*.

Na tym zakończę krótką analizę ilościową z uwzględnieniem kryterium użytkownika, przejdę teraz do także krótkiej analizy z uwzględnieniem kryterium pochodzenia tekstu.

15.4. ZESTAWIENIE NAJBOGATSZYCH I NAJUBOŻSZYCH W WYRAZY GRAFICZNE PÓL LEKSYKALNYCH – ANALIZA Z UWZGLĘDNIENIEM PODZIAŁU NA KORPUSY ZAWIERAJĄCE TEKSTY CZASOPISM, NEWSÓW ORAZ WPISÓW NA FORACH INTERNETOWYCH PLATFORMY ZIELONEJ I ŻÓLTEJ

Podczas poniższej analizy wykorzystałem fragmenty tabel od Z10 do Z12 (oraz Ż2), które zostały umieszczone w aneksie na stronach od 33 do 36 (oraz 38). Warto zaznaczyć, że w poniższym zestawianiu przytoczyłem także wynik dla platformy żółtej, w której zrównane zostały dane w korpusach NPSX i NCDA. Dodanie tej tabeli pozwoli nieco poszerzyć spektrum analizy. Wyimki z tabel prezentują się następująco:

CZASOPISMA		FORA	NEWSY	NEWSY	
1	KONSOLA	KONSOLA	PLAYSTATION	1	PLAYSTATION
2	PC	PC	XBOX	2	PC
3	KOMPUTER	KOMPUTER	PC	3	XBOX
4	RPG	RPG	KONSOLA	4	KONSOLA
5	XBOX	MULTI	KOMPUTER	5	DLC
6	STRZELANINA	GAMEPLAY	DLC	6	PSP
7	MULTI	DLC	GAMEPLAY	7	GAMEPLAY
8	PAD	PAD	TRAILER	8	KOMPUTER
9	SMARTFON	DEMO	PSP	9	TRAILER
10	ROZDZIELCZOŚĆ	TRAILER	MULTI	10	MULTI

Tabela 30. Zestawienie pól najbogatszych w wyrazy graficzne w korpusach tekstów pozyskanych z czasopism, newsów oraz wpisów na forach internetowych (platforma zielona i żółta)

Uwzględniając wyłącznie powyżej zaprezentowaną listę najobszerniejszych pól wyrazowych, które wystąpiły w tekstach czasopism zauważyć można, że znalazły się na niej takie unikalne pola, jak: *rozdzielczość* i *smartfon*.

Czasopisma zazwyczaj piszą o grach w sposób bardziej profesjonalny niż ma to miejsce w tekstach newsów lub na forach internetowych. Z faktem tym łączy się częstsze używanie większej liczby wyrazów z pola *rozdzielczość*, ponieważ to właśnie rozdzielczość w grach jest bardzo ważna dla współczesnego czytelnika/gracza. Dla wielu graczy sprawą drugorzędną jest fabuła gry, a najważniejsza dla nich jest oprawa wizualna danej produkcji. Grafika w grze musi być idealnie dopracowana, natomiast rozdzielczość powinna być jak najwyższa, aby dało się zobaczyć wszystkie szczegóły świata gry na dużym telewizorze/ekranie LCD.

Dziennikarze potrafią wyzyskać słownictwo pola *rozdzielczość* i dzięki temu pisać teksty na ciekawe tematy, które trafiają do graczy. Jednym z nich jest właśnie temat rozdzielczości w grach wideo.

Na podstawie powyższego zestawienia widać, że czasopisma często podejmują także temat smartfonów, który również należy do tych bardziej interesujących. Zwłaszcza czasopismo „CD-Action” dużo pisze o smartfonach, podając przy tym specyfikę danego sprzętu oraz zaznaczając, jakie gry uda się na nim uruchomić.

Natomiast w tekstach opublikowanych na forach internetowych odnotowano, w odróżnieniu od innych sposobów publikacji, wystąpienie pola *demo*. Dema są ważnym elementem świata graczy, ponieważ pozwalają zapoznać się z produkcją przed jej premierą oraz zdecydować o zakupie jej finalnej wersji. Gracze często dyskutują na forach o grach, zwłaszcza tych najbardziej popularnych. Pojawienie się dema takiej produkcji i możliwość zagrania w nie jest czymś ważnym dla graczy, dlatego później opisują i dzielą się wrażeniami właśnie na łamach forów internetowych.

W newsach, jak wynika z powyższego zestawienia, często podejmuje się tematykę konsol, a konkretniej mówiąc, dużo pisze na temat tych, które związane są z firmą Sony. Dlatego też na liście najobszerniejszych wyrazowo pól znalazły się *PlayStation* oraz *PSP*. Potwierdzenie tego widać w powyższych tabelach – żółtej oraz zielonej. W odniesieniu do danych w tabeli żółtej zauważyć można, że pole *PSP* znalazło się wyżej niż na platformie zielonej, to jeszcze dobitniej potwierdza, że pole to często występuje w tekstach newsów.

Do analizy najuboższych pól użyję tych samych tabel z aneksu, które wykorzystałem w powyżej analizie, ale oczywiście skorzystam z innych fragmentów:

CZASOOISMA		FORA	NEWSY	NEWSY	
10	UNBOXING DOKANAŁÓWKI ONE-LINER	CROSS-PLAY RANDOMIZACJA	PRZEJŚCIÓWKA WTYCZKA CREDITSY	10	OŚMIOBITOWY WTYCZKA
9	MULTITAP	CROSS-GEN	TRIGGER OŚMIOBITOWY PAY-TO-WIN PPM FLASHÓWKA	9	FLASHÓWKA NIEDZIELNIAK PAY-TO-WIN PPM
8	FLASHÓWKA PEESZWÓRKA CROSS-SAVE	CELOWNICZEK	GRZYBEK SAMOGRAJ PEESTRÓJKA CROSS-CONTROLLER PAST-GEN LPM	8	GRZYBEK LPM
7	BUDŻETÓWKA CROSS-CONTROLLER	PEESDWÓJKA MULTITAP	PRZESZKADZAJKA	7	COOLDOWN ZAPYCHACZ
6	SUB-QUEST	SUB-BOSS	QTE WSAD COOLDOWN MINIMAPA PEESZWÓRKA PEESPLUSOWY	6	CROSS-GEN MINIMAPA ONE-LINER PEESZWÓRKA PEESPLUSOWY WSAD
5	ODBLOKOWAJKA	OŚMIOBITOWY FLASHÓWKA MULTITOUCH	HEALER ZAPYCHACZ TEKSTÓWKA SZESNASTOBITOWY	5	HEALER SZESNASTOBITOWY TEKSTÓWKA
4	CROSS-GEN	SZESNASTOBITOWY	DEVKIT CROSS-GEN	4	DEVKIT MULTITASKING SUB-BOSS
3	RANDOMIZACJA JEDNOSTRZAŁOWIEC SZESNASTOBITOWY PEESPLUSOWY PEESDWÓJKA PAST-GEN	CROSS-SAVE BUMPER	CEL-SHADING NIEDZIELNIAK YLOD REAL-TIME ONE-LINER	3	REAL-TIME YLOD
2	PRZETRWANIÓWKA MULTIKART	PAY-TO-WIN	MULTITOUCH BUMPER MULTITASKING	2	BUMPER MULTITOUCH
1	LPM CROSS-CHAT ZBLUROWANY EKSBOKSŁAN PRZEDŁUŻAJKA STODWUDZIEŚTOOŚMIOBITOWIEC THREE-SIXTY TWARZOWANIE	LIVE-STREAM PEESZWÓRKA CROSS-CHAT ZBLUROWANY	SIDE-QUEST METAGRA SUB-BOSS RANDOMIZACJA PEESDWÓJKA	1	METAGRA PEESDWÓJKA RANDOMIZACJA SIDE-QUEST

Tabela 31. Zestawienie pól najuboższych w wyrazy graficzne w korpusach tekstów pozyskanych z czasopism, newsów oraz wpisów na forach internetowych (platforma zielona i żółta)

Widać, że najbogatsza w najuboższe pola wyrazowe jest lista pozyskana w wyniku analizy tekstów newsów. Przewagę nad pozostałymi listami widać na przykładzie nie tylko platformy zielonej, lecz także żółtej. Najmniej najuboższych wyrazowo pól odnotowano natomiast w tekstach pozyskanych z wpisów na forach internetowych. Zatem można stwierdzić, że autorzy tekstów newsów eksperymentują i bawią się słownictwem. Często tworzą nowe, kreatywne formy wyrazowe, które zostaną zaaprobowane przez środowisko i rozszerzą zakres swojego użycia, lub przypadną i pozostaną tylko wyrazami graficznymi z pól najuboższych wyrazowo. Pola zaprezentowane powyżej mogą zatem być zaczątkiem powstawania nowych lub efektem nieudolnie stworzonych słowoform.

Polami wyrazowymi, które istnieją ponad podziałami ze względu na rodzaj publikacji, są *cross-gen*, *flashówka*, *Peesczwórka*, *Peesdwójka*, *randomizacja* i *szesnastobitowy*. Zatem te pola można najrzadziej spotkać w różnych tekstach.

Do tej pory przyglądałem się wyłącznie polom znaczeniowym w odniesieniu do płaszczyzny zielonej (i żółtej), warto jeszcze przyjrzeć się płaszczyźnie czerwonej, czyli tej, której dane liczbowe pozyskane z korpusów zrównane zostały do jednakowej wielkości wyjściowej.

15.5. ZESTAWIENIE NAJBOGATSZYCH I NAJUBOŻSZYCH W WYRAZY GRAFICZNE PÓL LEKSYKALNYCH – ANALIZA POSZCZEGÓLNYCH CZĘŚCI KORPUSU PLATFORMY CZERWONEJ

W tej części analizy materiału niezbędne będą dane z tabel od C1 do C6 umieszczone w aneksie na stronach 39–44.

Z tabel wywnioskować można, że zestawienie dziesięciu najbogatszych pól leksykalnych we wszystkich korpusach pozostało bez zmian w odniesieniu do tabel z płaszczyzny zielonej. Tylko w jednym przypadku, w korpusie npsx, ich kolejność uległa zmianie – pole *Kinect* zamieniło miejsce z polem *demo*. W pozostałych korpusach kolejność pól pozostała taka sama, jak w przypadku płaszczyzny zielonej. Jedyne co się zmieniło, to liczba wyrazów zaliczanych do danych pól, ale to nie dziwi, bo po zastosowaniu mnożników normalne jest, że liczba wyrazów zwiększy się, oczywiście z wyjątkiem korpusów zbudowanych z wpisów na forach internetowych, bo to do nich zrównane były pozostałe korpusy. Brak zmian w stosunku do płaszczyzny zielonej wiąże się także z tym, że lista pól

wspólnych danym korpusom, czy też ich przynależności do danych środowisk również nie uległa modyfikacjom.

Nieco inaczej sprawa wygląda w przypadku pól najuboższych.

	NPSX	NCDA	FPSX →	FCDA →	PSX	CDA
1	PRE-ALPHA ↑	OFF-ROAD ↑	ROGUELIKE MULTITAP	PRE-ALPHA	TEKSTÓWKA ↑	PADLET ↑
2	RROD ↑	RROD ↑ SAMOGRAJ ↑ PEESTRÓJKA ↑ MINIMAPA ↑ WSAD ↑ PEESZWÓRKA ↑ PEESPLUSOWY ↑	PPM CROSS-PLATFORM	REAL-TIME RANDOMIZACJA	WTYCZKA ↑	ONE-LINER ↑
3	TURÓWKA ↑ SAMOGRAJ ↑ NIEDZIELNIAK ↑	PADLET ↑ STYLUS ↑ TRIGGER ↑ COOLDOWN ↑	PEESDWÓJKA SUB-BOSS	GEEK CEL-SHADING	PEESZWÓRKA ↑	SZESNASTOBITOWY ↑
4	PRZESZKADZAJKA ↑	CROSS-PLAY ↑ PRZESZKADZAJKA ↑ HEALER ↑ SZESNASTOBITOWY ↑	JEDNOSTRZAŁOWIEC CELOWNICZEK	DREAMCAST MARIAN SIDE-QUEST MUST-HAVE	PAST-GEN ↑	STACJONARKA ↑
5	MATRYCA ↑ CREDITS ↑ PEESTRÓJKA ↑	CZARNULA ↑ ZAPYCHACZ ↑ DEVKIT ↑ TEKSTÓWKA ↑	LPM	PRZESZKADZAJKA SHMUP	KLATKOWAĆ ↑	RANDOMIZACJA ↑
6	BUDŻETÓWKA ↑ SPRITE ↑	CROSS-PLATFORM →	SZESNASTOBITOWY	FREE-TO-PLAY CROSS-PLATFORM OŚMIOBITOWY	JEDNOSTRZAŁOWIEC ↑	BUDŻETÓWKA ↑ REAL-TIME ↑
7	WIESIEK ↑ ZNAJDŹKA ↑	PSP GO → CROSS-GEN ↑ REAL-TIME ↑ YLOD ↑	RANDOMIZACJA MULTITOUCH FLASHÓWKA BUMPER CROSS-SAVE	CELOWNICZEK DEVKIT ONE-LINER	PEŁNIAK ↑ CROSS-GEN ↑ PEESPLUSOWY ↑	MUST-HAVE ↑ PRZETRWAŃCÓWKA ↑
8	TANK ↑ FARMIC ↑ ONE-LINER ↑ SUB-BOSS ↑	CROSS-SAVE →	EGRANIZACJA GRZYBEK PAY-TO-WIN	PEESTRÓJKA MULTITOUCH FLASHÓWKA	MULTIKART ↑	CROSS-BUY ↑ CROSS-GEN ↑ PEESDWÓJKA ↑
9	HEJT ↑ SPAWN ↑ ZAPYCHACZ ↑ CROSS-GEN ↑ MULTITASKING ↑	ONE-LINER → BUMPER → MULTITOUCH →	TEKSTÓWKA OŚMIOBITOWY PEESZWÓRKA CROSS-CHAT LIVE-STREAM	PADLET CURRENT-GEN YLOD OFFROAD PAST-GEN DOKANAŁÓWKI PEESDWÓJKA	HEALER ↑ QWERTY ↑ REDZI ↑ DOKANAŁÓWKI ↑	DEVKIT ↑ MULTITAP ↑ EKSBOKSLAN ↑

10	PVE → OŚMIOBITOWY → COOLDOWN →	PAST-GEN → NIEDZIELNIAK → MULTITASKING → SIDE-QUEST → PEESDWÓJKA → RANDOMIZACJA → METAGRA →	ZBLUROWANY	PEESIKS STYLUS CZARNULA MULTITASKING SUB-QUEST ZBLUROWANY	PEESDWÓJKA → ZBLUROWANY → TWARZOWANIE →	UNBOXING → LPM → PEESPLUSOWY → CROSS-CHAT → PRZEDŁUŻAJKA → STODWUDZIESTO- OŚMIOBITOWIEC → THREE-SIXTY →

Tabela 32. Zestawienie pól najuboższych w wyrazy graficzne (platforma czerwona)

W powyższej tabeli po nazwie danego pola wyrazowego umieściłem małą strzałeczkę. Może ona być zielona i skierowana do góry, co oznacza, że pole w tym zestawieniu uplasowało się na wyższej pozycji niż w identycznym zestawieniu płaszczyzny zielonej (patrz strony 182–183). Strzałka może też być żółta i skierowana w prawą stronę, oznacza to, że pole to wystąpiło na identycznej pozycji w porównaniu z tabelą płaszczyzny zielonej.

Większość strzałek umieszczona została po nazwach pól, ale dwie z nich znalazły się już w główce tabeli, przy skrótowych nazwach korpusów fpsx i fcda, oznacza to, że pola wyrazów odnotowane w tych korpusach nie zmieniły swoich pozycji, a wynika to z tego, że w odniesieniu do nich nie były stosowane żadne mnożniki.

Pierwszą rzeczą, jaką można stwierdzić, patrząc na powyższą tabelę jest to, że uwzględnia ona znacznie mniej pól wyrazowych niż tabela płaszczyzny zielonej. W powyższej tabeli odnotowanych zostało 157 pól, a pomijając powtórzenia 89. Natomiast w tabeli zielonej było ich aż 234, a bez powtórzeń 118. Po zastosowaniu mnożników wiele pól przestało kwalifikować się do kategorii dziesięciu najuboższych, zresztą przyglądając się wyłącznie powyższemu zestawieniu w tabeli czerwonej widać, że raptem, nie licząc tych z korpusów fpsx i fcda, 26 pól (bez powtórzeń 25) pozostało na dawnych pozycjach (te oznaczone żółtą strzałką), reszta natomiast znalazła się na poziomach wyższych (te oznaczone zieloną strzałką), dlatego te, które w zielonym zestawieniu były w górnych wierszach tabeli, jeszcze bardziej zwiększyły swoją pozycję i przestały być obecne w czerwonym zestawieniu.

W wyniku zmniejszenia liczby pól umieszczonych w powyższym zestawieniu, zmianie uległa np. lista pól wspólnych wszystkim korpusom, ale także ta, która uwzględniała różne sposoby publikowania tekstów oraz ta, która dzieliła pola na bliższe danym środowiskom graczy wideo.

W przypadku pól najskromniejszych wyrazowo (uwzględnionych w czerwonej tabeli), ani jedno z nich nie było wspólne wszystkim korpusom.

Jeżeli podzieli się pola na grupy, uwzględniając sposoby publikacji tekstów, to w przypadku czasopism wspólne są takie pola, jak:

cross-gen, Peesdwójka, peesplusowy.

Jeśli uwzględnimy teksty z forów internetowych, to zestaw pól wspólnych jest identyczny, jak w przypadku platformy zielonej, natomiast w przypadku newsów lista ta wygląda następująco:

cooldown, cross-gen, multitasking, niedzielniak, one-liner, Peestrójka, przeszkadzajka, RROD, samograj, zapychacz.

Dokonując podziału pól na grupy związane z danym środowiskiem, nie odnotowałem wśród korpusów związanych z *konsolowcami* żadnego wspólnego pola, zaś dla *komputerowców* lista ta wygląda następująco:

devkit, one-liner, padlet, Peesdwójka, randomizacja, real-time.

Po przyrównaniu informacji z powyższej tabeli do danych z grafiki 10 widać, że pola z powyższego zestawienia to głównie te z obszaru zielonego. W powyższej tabeli znalazły się (pomijając powtórzenia) 42 pola z tego obszaru, 27 z obszaru żółtego i 20 z pól z niebieskiego.

W ogólnym porównaniu zestawień pól z grupami najuboższymi i najobszerniejszymi wyrazowo na platformie czerwonej widać, że identycznie, jak w przypadku platformy zielonej, tylko 1 pole wystąpiło w jednym i drugim zestawieniu: *pełniak*.

Warto także część pracy poświęcić na tabelę, która powstała w wyniku zsumowania wcześniej analizowanych tabel platformy czerwonej, umieszczona ona została w aneksie na stronie 45.

15.6. SUMA NAJBOGATSZYCH I NAJUBOŻSZYCH W WYRAZY GRAFICZNE PÓL LEKSYKALNYCH – ANALIZA CAŁEGO KORPUSU PLATFORMY CZERWONEJ

Dokonując analizy pól najbogatszych z uwzględnieniem danych z powyższej tabeli, już na pierwszy rzut oka można zauważyć, że lista dziesięciu pól uległa zmianie w porównaniu z analogicznymi danymi na płaszczyźnie zielonej:

1	PC →
2	PLAYSTATION →
3	KONSOLA →
4	XBOX →

5	KOMPUTER →
6	RPG →
7	DLC ↑
8	GAMEPLAY ↑
9	MULTI ↓
10	PSP ↑

Tabela 33. Suma pól najbogatszych w wyrazy graficzne (platforma czerwona)

W powyższej tabeli również zdecydowałem się zastosować symbole strzałek, ale w tym zestawieniu musiałem użyć dodatkowo jeszcze jednej, mianowicie czerwonej skierowanej w dół, która oznacza, że dane pole znalazło się na niższej pozycji w czerwonym zestawieniu, niż w analogicznym zestawieniu na płaszczyźnie zielonej.

Jak wynika z powyższej tabeli, aż 6 pól zachowało identyczną pozycję w porównaniu z zieloną płaszczyzną. Są to:

PC, PlayStation, konsola, xbox, komputer, RPG.

Jedno z pól spadło w ogólnej klasyfikacji – *multi* (w zielonej tabeli pole to było na pozycji 7), natomiast 3 pola podwyższyły swoją pozycję: *DLC* (z pozycji 3), *gameplay* (z pozycji 9) oraz *PSP*, które wcześniej nie znalazło się w zestawieniu zielonej tabeli. Pole zajmowało wtedy pozycję 11, czyli poza zestawieniem. Zastąpiło ono w tym zestawieniu pole *trailer*, które w czerwonej, zsumowanej tabeli zajmuje obecnie 11 miejsce.

W przypadku pól z najrzadszymi wyrazami platformy czerwonej po zestawieniu jej z analogiczną listą zieloną widać, że ta pierwsza uwzględnia znacznie mniej pól wyrazowych od tej drugiej. Zamiast 18 pól, które wystąpiły na płaszczyźnie zielonej, w tym zestawieniu (na płaszczyźnie czerwonej) jest ich tylko 15:

10	JEDNOSTRZAŁOWIEC ↓
9	SZESNASTOBITOWY ↑
8	RANDOMIZACJA ↓
7	PEESPLUSOWY ↑
6	PEESDWÓJKA ↓
5	ODBLOKOWAJKA ↑
4	MULTIKART ↑ PRZETRWANIÓWKA ↑

3	CROSS-CHAT ↑
	ZBLUROWANY ↑
2	EKSBOOKSŁAN ↑
1	TWARZOWANIE →
	PRZEDŁUŻAJKA →
	STODWUDZIESTOOŚMIOBITOWIEC →
	THREE-SIXTY →

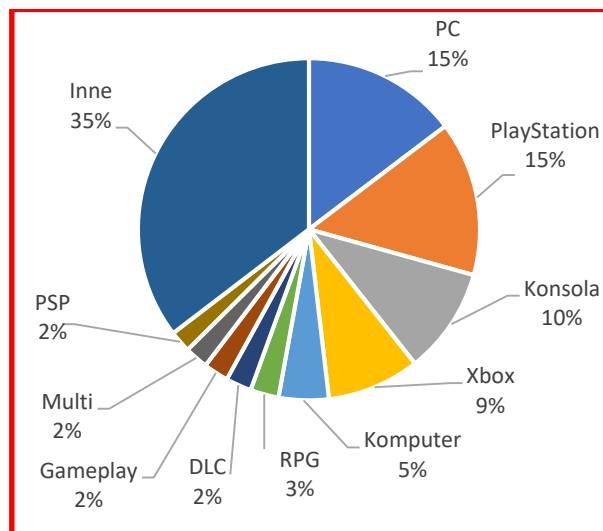
Tabela 34. Suma pól najuboższych w wyrazy graficzne (platforma zielona)

Jak widać, tylko cztery pola, w porównaniu z analogiczną tabelą zieloną, utrzymały swoją dawną pozycję. Są to:

twarzowanie , przedłużajka, stodwudziestośmiobitowiec i Three-sixty,

natomiast reszta pól wyrazowych podniosła lub obniżyła swoją pozycję w powyższej tabeli. Obniżenie to związane jest z tym, że inne pola zwiększyły swoją liczbę wyrazów graficznych (dzięki zastosowaniu mnożników), natomiast inne nie (np. te, które wystąpiły w korpusach fcda lub fpsx), tym samym musiały zająć niższą pozycję w zestawieniu.

Przyglądając się zestawieniom procentowym w stosunku do pól najobszerniejszych i najuboższych widać, że nie różnią się one znacznie w stosunku do swoich odpowiedników na płaszczyźnie zielonej:

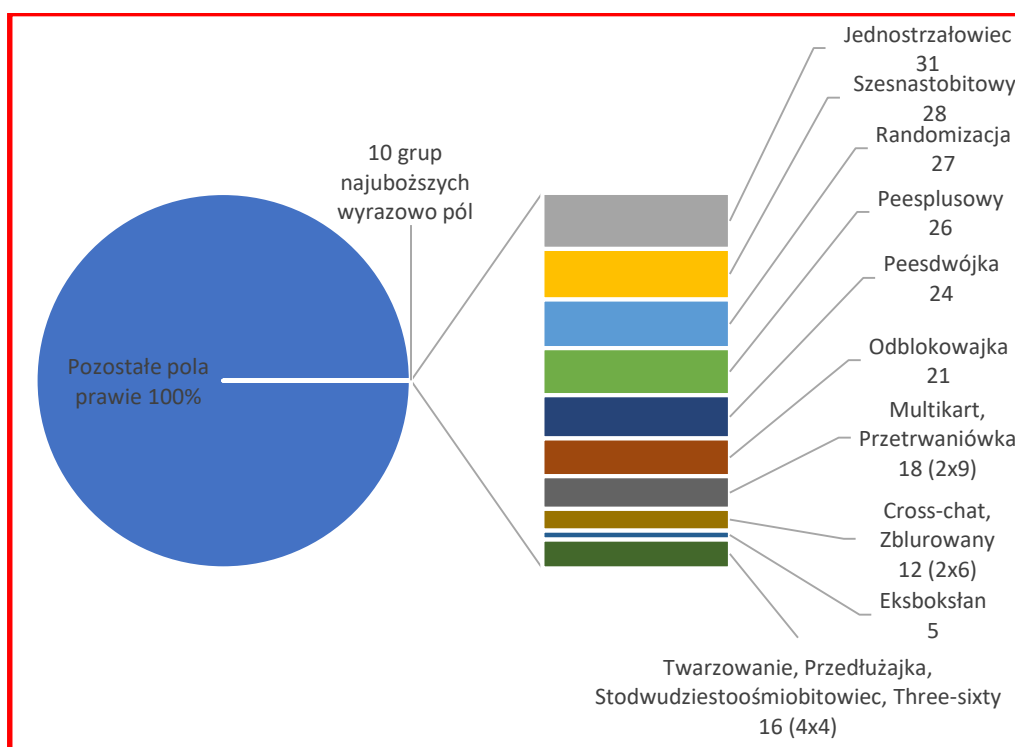


Wykres 16. Suma pól najbogatszych w wyrazy graficzne – zestawienie procentowe (platforma czerwona)

Na wykresie widać, że cztery pola wyrazowe w zestawieniu z płaszczyzną zieloną zyskały po dwa punkty procentowe, są to: *PC*, *PlayStation*, *Xbox* i *komputer*. W trzech polach

wynik procentowy nie uległ zmianie – *RPG*, *DLC* i *gameplay*, natomiast dwa pola straciły po jednym punkcie procentowym – *multi* i *konsola*. Ponadto na wykresie odnotowano nowe pole, które nie wystąpiło w zielonym zestawieniu, czyli *PSP*. Zestawiając wyniki procentowe najobszerniejszych wyrazowo pól widać, że stanowią one jeszcze większy procent wszystkich wyrazów graficznych graczy, niż na platformie zielonej. Wynik ten poprawił się o 3 punkty procentowe.

Poniższy wykres pokazuje, jak wygląda sytuacja dziesięciu najuboższych pól.



Wykres 17. Suma pól najuboższych w wyrazy graficzne – zestawienie procentowe (platforma czerwona)

Tworząc powyższy wykres, podobnie jak w przypadku wykresu zielonego, zdecydowałem się na podanie w zestawieniu po prawej stronie realnej liczby wystąpień, zamiast wyniku procentowego, ponieważ inaczej wszędzie musiałbym wpisać zero.

Identycznie, jak w zielonym wykresie, także tutaj dziesięć najuboższych wyrazowo pól nie stanowi nawet 1 procenta wszystkich wyrazów graczy. Łącznie w polach najmniej obfitych odnotowano 208 użyć jednostek graczy, co stanowi 0,0181853% wszystkich wyrazów graczy zgromadzonych w korpusach platformy czerwonej. Porównując tylko wyniki procentowe platformy czerwonej i zielonej, widać, że w przypadku tej pierwszej wyrazy pól najuboższych,

w odniesieniu do całego zbioru jednostek graczy, stanowią mniej niż połowę tego, co wykazano badając pola platformy zielonej.

Jeśli zestawia się powyższe dane z informacjami przedstawionymi na grafice 10, to w przypadku najbogatszych pól wszystkie one zaklasyfikowane zostały jako pola z obszaru zielonego, natomiast w przypadku pól najuboższych wyrazowo, to 10 z nich wystąpiło w żółtych obszarach (polach *komputerowców*), 3 w obszarach niebieskich (*konsolowców*), a 2 w obszarze zielonym (wspólnym).

15.7. ZESTAWIENIE NAJBOGATSZYCH I NAJUBOŻSZYCH W WYRAZY GRAFICZNE PÓL LEKSYKALNYCH – ANALIZA Z UWZGLĘDNIENIEM PODZIAŁU NA KORPUSY *KONSOLOWCÓW* ORAZ *KOMPUTEROWCÓW* PLATFORMY CZERWONEJ

Materiał do analizy przedstawionej w tej części pochodzi z tabel od C8 do C9, które umieściłem na stronach 47 i 48 w aneksie do tej pracy.

	KONSOLOWCY	KOMPUTEROWCY
1	PLAYSTATION ↑	PC →
2	PC ↑	PLAYSTATION ↑
3	KONSOLA ↓	KOMPUTER ↓
4	XBOX →	KONSOLA →
5	PSP ↑	XBOX →
6	GAMEPLAY ↓	RPG →
7	DLC ↑	STRZELANINA ↑
8	RPG ↑	MULTI ↓
9	MULTI ↓	TABLET →
10	TRAILER →	DLC ↓

Tabela 35. Zestawienie pól najbogatszych w wyrazy graficzne w korpusach *konsolowców* oraz *komputerowców* (platforma czerwona)

W porównaniu z analogiczną tabelą platformy zielonej widać, że w obu listach zestaw pól leksykalnych pozostał ten sam, zmieniła się tylko ich kolejność.

W przypadku *konsolowców* nastąpiła bardzo ciekawa zamiana miejscami pewnych pól. Na platformie zielonej na pierwszym miejscu znajdowało się pole *konsola*, natomiast teraz jego

miejsce zajęło pole *PlayStation*. Zatem obecnie mamy do czynienia z sytuacją, w której *PlayStation* (rodzaj konkretnego sprzętu – konsoli), zajmuje pierwsze miejsce w zestawieniu u *konsolowców*, natomiast u komputerowców – *PC* (rodzaj konkretnego sprzętu – komputera). Mamy więc na tym samym poziomie równowagę pod względem swoistej kategorii „rodzaj konkretnego sprzętu”.

Po zamianie miejscami pewnych pól równowagę dostrzec można także na pozycji trzeciej, bo zarówno u *konsolowców* jest tam umieszczone pole reprezentujące kategorię „ogólna nazwa sprzętu” – *konsola*, jak i u *komputerowców* też jest tam przedstawiciel kategorii „ogólna nazwa sprzętu” – *komputer*.

Można więc stwierdzić, że na trzecim poziomie tabeli mamy do czynienia ze swoistym wyrównaniem hiperonimów – *konsola* i *komputer*, a na poziomie pierwszym hiponimów – *PlayStation* i *PC*. A. Markowski zaznacza, że relacje hiperonimii oraz homonimii dotyczą leksemów [2012: 111], natomiast moim zdaniem można ją przenieść także na grunt pól leksykalnych. Starłem się to pokazać, analizując powyższe przykłady.

Przejdźmy zatem do analizy zestawów pól najuboższych:

	KONSOLOWCY	KOMPUTEROWCY
10	PEESZWÓRKA ↑	ODBLOKOWAJKA ↑
9	JEDNOSTRZAŁOWIEC ↑	DEVKIT ↑
8	WSAD ↓	PEESPLUSOWY ↑
7	PEESDWÓJKA →	CROSS-GEN ↑
6	PEESPLUSOWY ↑	CROSS-PLATFORM ↓ JEDNOSTRZAŁOWIEC ↓ PEESDWÓJKA ↑
5	MULTIKART ↑ PPM ↓	CZARNULA ↑ PRZETRWANIÓWKA ↑
4	LPM ↓	YLOD →
3	SZESNASTOBITOWY ↓ ZBLUROWANY ↑	CROSS-SAVE → MULTITAP ↑ EKSBOKSŁAN ↑
2	FLASHÓWKA ↓ RANDOMIZACJA ↓ TWARZOWANIE ↑	PAST-GEN ↓ CROSS-CHAT ↑ PRZEDŁUŻAJKA ↑ STODWUDZIESTOOŚMIOBITOWIEC ↑ THREE-SIXTY ↑
1	CROSS-CHAT ↓	SUB-QUEST → ZBLUROWANY →

Tabela 36. Zestawienie pól najuboższych w wyrazy graficzne w korpusach *konsolowców* oraz *komputerowców* (platforma czerwona)

Analizując powyższe zestawienie i przyrównując je do platformy zielonej widać, że zmniejszona została liczba pól, które wystąpiły na liście zarówno u *konsolowców*, jak i u *komputerowców*. Podczas analizy platformy zielonej odnotowałem takich pól 7, natomiast teraz jest ich tylko 5. Pola, które w zielonym badaniu zaliczone zostały do tych, które posiadają najmniej wyrazów graficznych (*szesnastobitowy* i *peeszwórka*) przestały takimi być, zatem po zrównaniu danych widać, że wyrazy graficzne pól, których nie odnotowano w zestawieniu czerwonym nie były tak bardzo rzadkie, jakby mogło się początkowo wydawać.. Rzadkość użycia wyrazów graficznych potwierdzona natomiast została w przypadku takich pól, jak: *cross-chat*, *jednostrzałowiec*, *Peesdwójka*, *peesplusowy* i *zblurowany*.

15.8. ZESTAWIENIE NAJBOGATSZYCH I NAJUBOŻSZYCH W WYRAZY GRAFICZNE PÓL LEKSYKALNYCH – ANALIZA Z UWZGLĘDNIENIEM PODZIAŁU NA KORPUSY ZAWIERAJĄCE TEKSTY CZASOPISM ORAZ NEWSÓW PLATFORMY CZERWONEJ

Do analizy najbogatszych w wyrazy graficzne pól potrzebne jest przywołanie tylko jednej tabeli, która skupia w sobie zestaw tych pól wyrazowych, które pozyskano z tekstów newsów. Wystarczy przywołanie wyłącznie tej jednej listy pól, ponieważ w czerwonych tabelach, które zawierają dane z czasopism i forów internetowych, znajdują się dokładnie te same pola wyrazowe, które wskazałem podczas analizy płaszczyzny zielonej. Ponadto dane zawarte w tabelach czerwonych ulokowane zostały na tych samych pozycjach, co w przypadku platformy zielonej. Zwiększona została jedynie suma ich wyrazów graficznych (dzięki zastosowaniu mnożników). Postanowiłem więc skupić się jedynie na newsach, w których zaszły pewne zmiany:

	NEWSY	A	B
1	PLAYSTATION	➡	➡
2	PC	⬆	➡
3	XBOX	⬆	➡
4	KONSOLA	➡	➡
5	DLC	⬆	➡
6	GAMEPLAY	⬆	⬆
7	PSP	⬆	⬇
8	KOMPUTER	⬇	➡
9	TRAILER	⬇	➡
10	MULTI	➡	➡

Tabela 37. Lista pól najbogatszych w wyrazy graficzne w korpusach tekstów pozyskanych z newsów (platforma czerwona)

W powyższej tabeli obok kolumny zawierającej nazwę pola wyrazowego dodałem dwa dodatkowe elementy – kolumnę A i B, a w nich różnego rodzaju strzałki. Powstanie tych kolumn wymuszone zostało tym, że w przypadku newsów dysponuję poza danymi z platformy zielonej i czerwonej, także tymi z platformy żółtej. Kolumna A odnosi się do wyników pozyskanych z platformy zielonej, zaś B do tych z platformy żółtej. Wyniki te przedstawione zostały na stronie 191.

Przyglądając się polu *PC* widać, że na platformie czerwonej zajmuje ono 2 miejsce pod względem sumy wyrazów graficznych, czyli jest to pozycja wyższa niż pole to zajmowało na platformie zielonej – 3 pozycja. Oznaczeniem tej zmiany jest umieszczenie w kolumnie A symbolu strzałki w górę. Porównując dane z platformy zielonej i żółtej zaobserwować można, że pole *PC* zajmuje identyczne miejsce, jak w przypadku platformy czerwonej, dlatego też w kolumnie B umieszczona została żółta strzałka skierowana w prawo.

Na szczególną uwagę podczas analizy przedstawionej w tej części zasługują pola, które na wszystkich trzech platformach zajmowały tę samą pozycję, a są to: *PlayStation*, *konsola* i *multi*. Zajęcie tej samej pozycji w różnych badaniach jest ich dodatkowym potwierdzeniem oraz utwierdzeniem w wynikach. Zatem niezaprzeczalnie można stwierdzić, że polem najobszerniejszym wyrazowo w newsach jest zawsze *PlayStation*.

Potwierdzenie pozycji oraz wysokiego miejsca na liście pól, uwzględniając ich bogactwo wyrazowe, udowodniono także w przypadku pól *konsola* oraz *multi*. W przypadku pozostałych pól z listy ich pozycje w zależności od przyjętej platformy badawczej są różne, ale jedno jest niezmiennie – pola te zawsze będą wyróżnione jako 10 najbogatszych wyrazowo.

Rozpatrując powyższe wyniki pod nieco innym kątem stwierdzić trzeba, że pod względem tematyki newsów oraz treści w nich zawartych wiele tekstów dotyczyło w jakimś stopniu różnych konsol, a najwięcej jednej z nich, czyli PlayStation oraz usług i gier do niej dostosowanych.

Przejdźmy zatem do list pól najuboższych wyrazowo, w której z kolei pominąłem dane dotyczące forów internetowych na płaszczyźnie czerwonej, ponieważ nie uległy one żadnej zmianie, bo to do nich zrównane zostały wartości innej części korpusu. Także tutaj w przypadku newsów zastosowałem podział na kolumny A i B, które reprezentują platformy zieloną i żółtą. Spójrzmy zatem na listę pól najuboższych wyrazowo:

ZASOPIŚMA		NEWSY		A	B
10	BUDŻETÓWKA ↑	10	ZAPYCHACZ	↑	↑
9	CROSS-CONTROLLER ↑	9	WSAD	↑	↑
8	SUB-QUEST ↑		MINIMAPA	↑	↑
7	ODBLOKOWAJKA ↑		PEESCZWÓRKA	↑	↑
6	CROSS-GEN ↑	8	PEESPLUSOWY	↑	↑
5	SZESNASTOBITOWY ↑	7	CROSS-GEN	↑	↑
	PEESPLUSOWY ↑	6	ONE-LINER	↑	↑
	PAST-GEN ↑	6	HEALER	↑	↑
4	RANDOMIZACJA ↑	5	SZESNASTOBITOWY	↑	↑
	JEDNOSTRZAŁOWIEC ↑		TEKSTÓWKA	→	→
	PEESDWÓJKA ↑	4	DEVKIT	↑	↑
3	PRZETRWANIÓWKA ↑	4	MULTITASKING	↑	→
	MULTIKART ↑	3	YLOD	→	→
2	EKSBOKSLAN ↑		REAL-TIME	→	→
1	LPM →		SUB-BOSS	↑	↑
	CROSS-CHAT →	2	MULTITOUCH	→	→
	PRZEDŁUŻAJKA →	BUMPER	→	→	
	STODWUDZIESTOOŚMIOBITOWIEC →	1	SIDE-QUEST	→	→
	THREE-SIXTY →		METAGRA	→	→
	ZBLUROWANY →		RANDOMIZACJA	→	→
	TWARZOWANIE →		PEESDWÓJKA	→	→

Tabela 38. Zestawienie pól najuboższych w wyrazy graficzne w korpusach tekstów pozyskanych z czasopism oraz newsów (platforma czerwona)

Odnosząc się do list pól platformy zielonej, które przedstawiłem na stronie 192 i zestawiając je z powyższymi danymi platformy czerwonej, która pełni swoistą funkcję kontrolną, dostrzec można, że w przypadku czasopism wszystkie pola z poziomu pierwszego zostały na niezmiennych pozycjach względem platformy zielonej. Zatem w czasopismach znaleźć najmniej można wyrazów z takich pól, jak: *LPM*, *cross-chat*, *przedłużajka*, *zblurowany*, *stodwudziestoośmiobitowiec*, *Three-sixty*, i *twarzowanie*.

Natomiast w przypadku newsów swoje pozycje względem platformy zielonej zachowały (z uwzględnieniem także platformy żółtej) pola z poziomu pierwszego, drugiego oraz częściowo trzeciego i piątego. Są to więc: *side-quest*, *metagra*, *randomizacja*, *Peesdwójka*, *multitouch*, *bumper*, *YLOD*, *real-time* oraz *tekstówka*. Pola te są tymi, które najrzadziej wystąpiły w tekstach newsów.

Na zakończenie tego rozdziału chciałbym zaznaczyć, że analizując pola pod względem sumy wyrazów graficznych w poszczególnych częściach korpusu nie wyczerpałem oczywiście całego spektrum badawczego zebranego materiału, jednak starałem się zwrócić uwagę na moim zdaniem najciekawsze elementy możliwe do uchwycenia w zebranych materiale.

16. ZŁOŻONOŚĆ A SUMA WYRAZÓW GRAFICZNYCH W CAŁYM KORPUSIE

Warto w tym rozdziale na wstępie przytoczyć listę pól z największą liczbą form podstawowych:

	NAZWA POLA	LICZBA FORM PODST.
1	CASUAL	24
2	PATCH	20
	PC	20
3	CALAK	17
	SPAWN	17
4	HEJT	16
	PLAYSTATION	16
5	CRAP	15
	HARDKOR	15
6	STREAM	13
7	GRIND	12
8	BAN	11
	SANDBOX	11
	SAVE	11
9	CRAFT	10
	RPG	10
	SINGIEL	10
	TEKSTURA	10
10	KOMPUTER	9
	REMASTER	9

Tabela 39. Liczba podstawowych form wyrazowych w poszczególnych polach

Przyrównując dane z powyższej tabeli do zestawienia pól z największą sumą wyrazów tekstowych, która przedstawiona została w poprzednim rozdziale (patrz strony 180–181) widać, że duża złożoność pól rzadko łączy się z dużą sumą wyrazów graficznych w nich obecnych. Tylko cztery pola z zestawienia obejmującego najobszerniejsze wyrazowo pola posiada złożoność na poziomie 9 lub więcej form podstawowych. Są to pola: *PlayStation*, *PC*, *RPG* i *komputer*.

Co ciekawe pole, które w zestawieniu na stronach 180–181 wystąpiło we wszystkich kolumnach, czyli suma jego wyrazów graficznych była w nich na tyle wysoka, że zawsze kwalifikowało się do rankingu 10 najobszerniejszych wyrazowo, nie posiada wysokiej złożoności. Mowa oczywiście o polu *konsola*, którego niska złożoność dziwi, gdyż jest to jedno z podstawowych pól środowiska graczy wideo.

Oczywistym, ale wartym odnotowania, jest także to, że w powyższym zestawieniu pól z największą złożonością, nie odnotowano żadnego pola z tych, które w sumie zawierają najmniej wyrazów graficznych, zarówno na płaszczyźnie czerwonej, jak i zielonej.

Można więc wywnioskować, że wysoka złożoność pól sprzyja większej liczbie użyciu ich wyrazów graficznych w tekstach, a tym samym większej popularności całych pól, dzięki czemu znane i używane są w obu środowiskach graczy wideo. Natomiast większa suma wyrazów graficznych danych pól niekoniecznie łączy się z większą złożonością.

Analizując nie pola, a wyłącznie formy podstawowe, zauważyć można, że najwięcej odnotowano ich w korpusie fpsx – 694 (na 709 możliwych), łączy się to bezpośrednio z informacją, którą przedstawiłem na stronie 160, że w tym korpusie odnotowano także obecność największej liczby pól wyrazowych.

Na drugim miejscu uplasował się korpus psx z 687 formami podstawowymi, a następnie cda z 683 formami. Korpusy zawierające teksty z czasopism, także pod względem bogactwa pól wyrazowych uplasowały się na drugiej pozycji (patrz strona 160)

Kolejne miejsca zajęły korpusy: ncda z 680 formami, fcda – 670 oraz npsx – 633.

Analizując dane z podziałem na sposób publikacji można zauważyć, że najbogatsze w formy podstawowe okazały się korpusy zbudowane z tekstów czasopism. Widać, że gracze, którzy są dziennikarzami operują większym bogactwem form podstawowych niż Ci, którzy są nimi nie są i publikują teksty na forach internetowych. Wynika to między innymi z tego, że dziennikarze chcą stworzyć interesujące teksty i przyciągnąć czytelników, a uzyskać to mogą posługując się większą liczbą ciekawych form wyrazowych. Gracze, którzy piszą na forach posługują się zdecydowanie uboższym słownictwem. W korpusach, które uwzględniają teksty z czasopism odnotowano aż 706 form podstawowych (z 709 możliwych). Na drugim miejscu znalazły się teksty z forów internetowych oraz newsów, w których znaleziono po 694 formy podstawowe.

Z kolei dzieląc korpusy na te, które związane są z *konsolowcami* i *komputerowcami*, odnotowano więcej form podstawowych w tym drugim środowisku graczy wideo. To teksty *konsolowców* są bogatsze w formy podstawowe. Odnotowano w nich aż 702 użyte formy, natomiast u *komputerowców* było ich 700.

17. POLA WYRAZOWE A CZĘSTOTLIWOŚĆ WYSTĄPIEŃ WYRAZÓW GRAFICZNYCH

Do tej pory zbadałem pola wyrazowe pod względem ich obfitości w wyrazy graficzne bez rozbijania na kwartały, czyli zsumowane za 5 lat oraz pod względem złożoności, czyli pokazałem, ile podstawowych form wyrazowych w sobie zawierają. W tej części pracy przyjrę się tym samym polom, tylko pod kątem ich częstotliwości z rozbiciem na kwartały. Jednocześnie pokażę, jaka jest suma wystąpień wyrazów graficznych w poszczególnych kwartałach. Łącznie zbadałem 189 pól, które zawierały w sobie 434.924 (w zielonej platformie) i 1.143.781 (w czerwonej platformie) wyrazów graficznych⁷⁸.

Ta część pracy pozwoli uzyskać odpowiedzi na kolejne pytania/założenia przedstawione na początku pracy. Pierwsze z nich to, czy istnieją pola, których wyrazy są znacznie częściej używane w danych kwartałach oraz czy mają one jakieś wyraziste punkty na wykresie graficznym/liniowym, które wskazują diametralną zmianę liczby użyć wyrazów z danego pola wyrazowego. Drugie – czy można wskazać ogólną tendencję całego pola poprzez kwartalną analizę wyrazów tekstowych w danym przedziale czasu. Innymi słowy, czy możliwe jest wskazanie, czy dane pole w przeciągu całego okresu badawczego (5 lat) wykazuje tendencję wzrostową (coraz więcej wyrazów z danego pola jest używanych w tekstach) czy też malejącą. Trzecie pytanie, czy dominują wykresy, których poszczególne punkty symbolizujące liczbę użyć wyrazów w danym kwartale są bardzo od siebie oddalone (wykresy niestabilne), czy też wykresy, w których liczba wyrazów w poszczególnych kwartałach jest zbliżona.

Ze względu na ogrom pozyskanych materiałów zdecydowałem, że w tej części zaprezentuję głównie wykresy dla sumy wyrazów graficznych z poszczególnych kwartałów danych korpusów⁷⁹. Suma ta powstała przez dodanie wyników kwartalnych z korpusów *cda*, *ncda*, *psx* i *npsx* oraz z zaokrąglonej średniej liczby użyć wyrazów w jednym kwartale w korpusach *fcda* i *fpsx*. Poszczególne kroki w celu uzyskania sumy niezbędnej do stworzenia wykresów, zaprezentowałem poniżej na przykładzie pola znaczeniowego – *craft*⁸⁰.

⁷⁸ We fragmencie rozdziału posługuję się także wynikami pozyskanymi na platformie żółtej, która nie obejmuje wszystkich części korpusu, a jedynie wyrównuje te pozyskane z newsów. Jest zatem dobrym uzupełnieniem badań.

⁷⁹ Jednak dwa krótkie rozdział poświęcę na analizę ciekawszych wykresów, które stworzone zostały w oparciu o inny podział materiału.

⁸⁰ Pozyskiwanie materiałów w danej grze, np. rudy żelaza z kopalni.

CRAFT				
KWARTAŁ	CDA	PSX	NCDA	NPSX
1 KW. 2011	5	0	1	0
2 KW. 2011	9	1	2	0
3 KW. 2011	6	1	0	0
4 KW. 2011	7	0	0	0
1 KW. 2012	12	1	0	0
2 KW. 2012	3	3	0	0
3 KW. 2012	13	0	0	0
4 KW. 2012	13	0	2	0
1 KW. 2013	19	1	1	0
2 KW. 2013	7	0	7	0
3 KW. 2013	20	1	2	0
4 KW. 2013	13	3	0	0
1 KW. 2014	16	0	1	0
2 KW. 2014	21	1	1	2
3 KW. 2014	31	2	6	1
4 KW. 2014	25	1	9	0
1 KW. 2015	19	4	3	0
2 KW. 2015	12	1	3	0
3 KW. 2015	22	1	3	1
4 KW. 2015	14	7	0	3

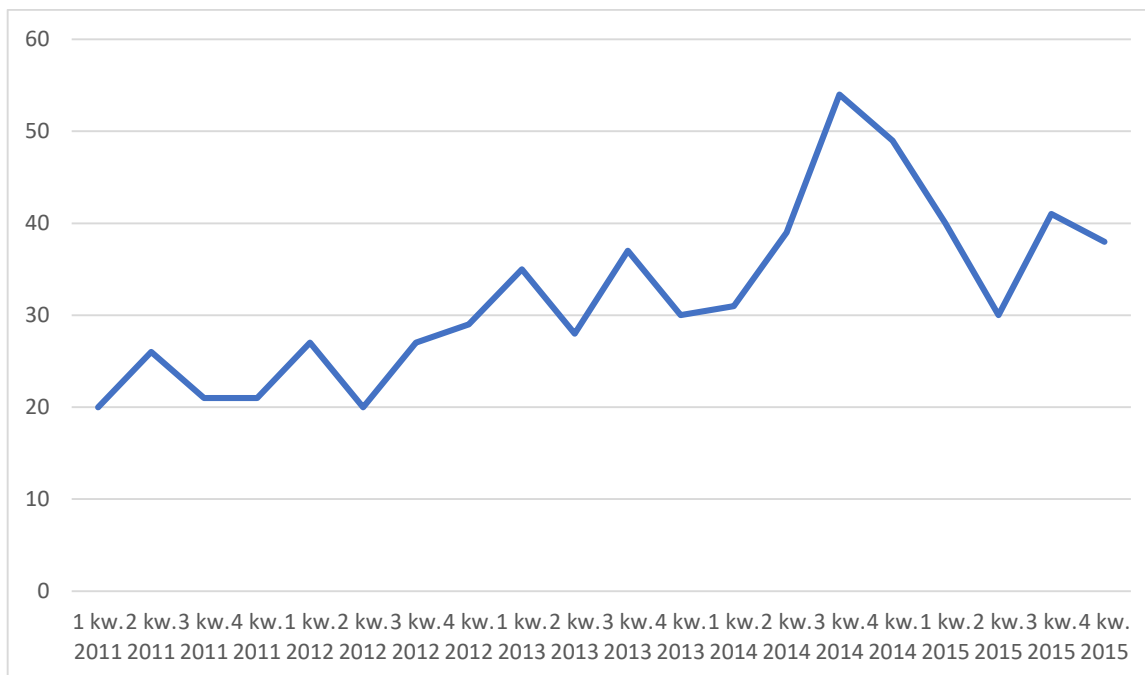
CRAFT	
KWARTAŁ	SUMA
1 KW. 2011	20
2 KW. 2011	26
3 KW. 2011	21
4 KW. 2011	21
1 KW. 2012	27
2 KW. 2012	20
3 KW. 2012	27
4 KW. 2012	29
1 KW. 2013	35
2 KW. 2013	28
3 KW. 2013	37
4 KW. 2013	30
1 KW. 2014	31
2 KW. 2014	39
3 KW. 2014	54
4 KW. 2014	49
1 KW. 2015	40
2 KW. 2015	30
3 KW. 2015	41
4 KW. 2015	38

CRAFT		
	FCDA	FPSX
SUMA	91	184

Tabela 40. Informacje na temat liczebności wyrazowej pola *craft* oraz dane liczbowe wykorzystane do stworzenia wykresu tego pola

W tabeli po lewej stronie umieściłem zebrane dane na temat liczby wystąpień wyrazów pola *craft* w danym kwartale. Uwzględniłem przy tym podział na korpusy: *cda*, *psx*, *ncda* i *npsx*. W środkowej tabeli podałem wyniki pozyskane z korpusów *fcda* i *fpsx*, natomiast w tabeli po prawej, efekt sumowania danych z tabeli środkowej oraz tej po lewej stronie.

Liczby umieszczone w prawej tabeli powstały poprzez zsumowanie liczb z danego kwartału lewej tabeli (np. dla 1 kwartału: 5+0+1+0) oraz dodania zaokrąglonej średniej liczby wystąpień wyrazów w *fcda* i *fpsx*. Powstała ona przez podzielenie danych środkowej tabeli (91 i 184) przez liczbę kwartałów (20). W ten sposób obliczyłem, że średnia dla *fcda* wynosi 4,55 (po zaokrągleniu 5), a dla *fpsx* – 9,2 (po zaokrągleniu 9). Efektem było stworzenie danych bazowych, które były podstawą stworzenia prezentacji graficznej (wykresu), która dla pola *craft* wygląda następująco:

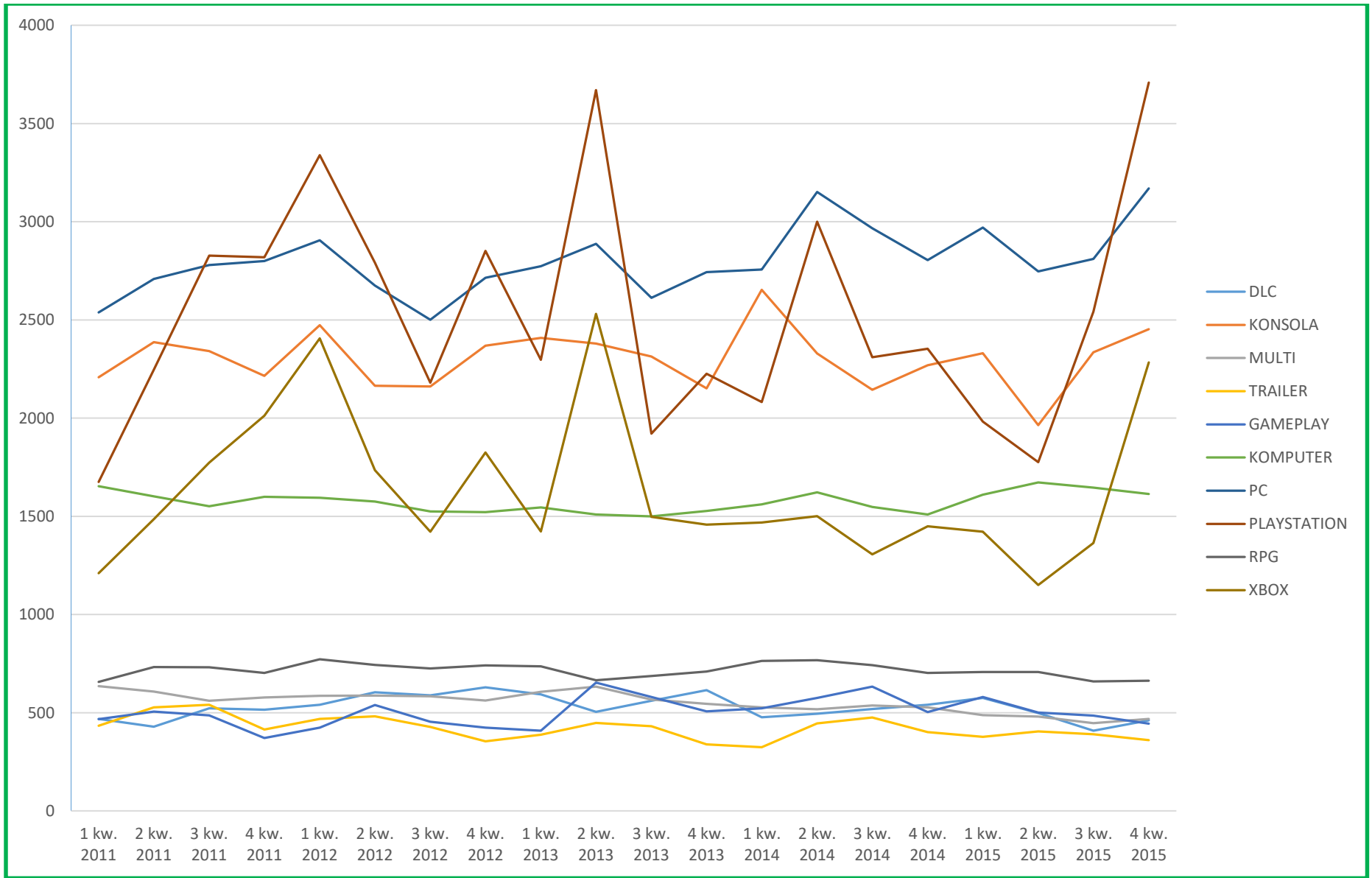


Wykres 18. Graficzna prezentacja częstotliwości pola *craft*

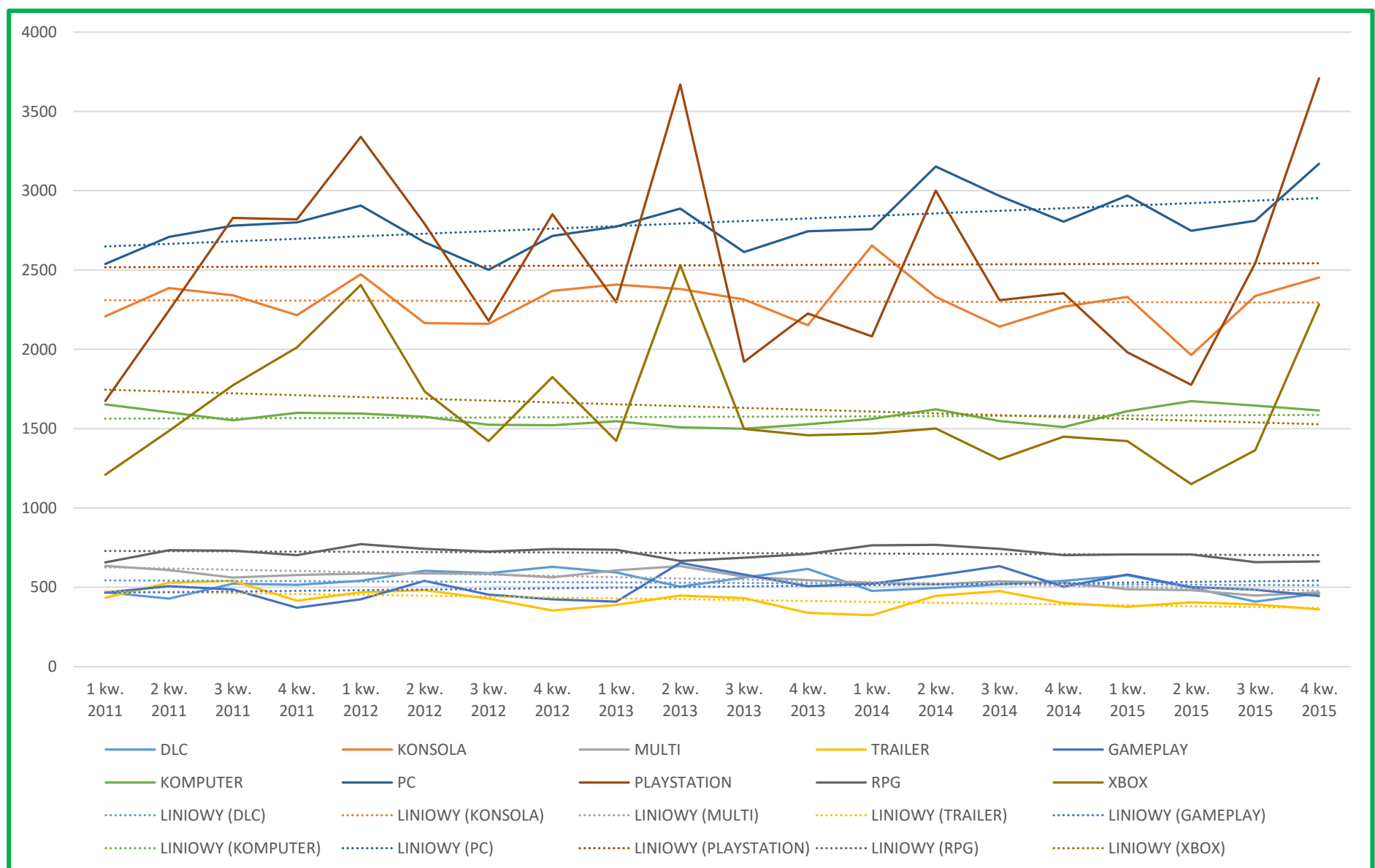
Na osi x umieściłem kwartały, których zaprezentowane dane dotyczą, natomiast na osi y łączną liczbę wystąpień w tekstach wyrazów graficznych. W analogiczny sposób postąpiłem tworząc wykresy dla reszty pól znaczeniowych.

17.1. ANALIZA WYKRESÓW PLATFORMY ZIELONEJ I CZERWONEJ W ODNIESIENIU DO PÓL NAJBOGATSZYCH W WYRAZY GRAFICZNE

Podobnie, jak w przypadku poprzedniej części pracy, gdzie analizowałem sumy wystąpień wyrazów graficznych w korpusach, także tutaj zacznę od analizy wykresów pól, które były najobszerniejsze wyrazowo. Zacznę od wykresu dla płaszczyzny zielonej:



Wykres 19. Graficzne przedstawienie 10 najbogatszych w wyrazy graficzne pól platformy zielonej



Wykres 20. Graficzne przedstawienie 10 najbogatszych w wyrazy graficzne pól platformy zielonej (wraz z liniami trendu)

Wykresy różnią się między sobą przede wszystkim poziomem stabilizacji. Jedne wykazują większe (nawet ponad 1000 jednostek) zróżnicowanie między danymi z różnych kwartałów, a inne mniejsze.

Wykresy bardziej ustabilizowane, czyli takie, które mają niewielkie odstępstwa od średniej liczby użyć wyrazów w danym kwartale, świadczą o tym, że są one reprezentacją tego pola, które w równym lub prawie równym stopniu występuje w każdym z kwartałów. Istnieje więc duże prawdopodobieństwo, że w prawie każdym tekście opublikowanym w czasopiśmie, na stronie internetowej lub forum wyrazy z tego pola będą obecne, bo są ważne dla graczy.

Analizując wykresy mniej stabilne, można wyciągnąć wniosek, że reprezentują one wyrazy z tych pól, z których gracze korzystają często, ale tylko w pewnych sytuacjach. Można więc znaleźć teksty, w których tych wyrazów będzie mniej, ale też takie, w których będzie ich znacznie więcej od innych. Wyrazy te są ważne dla graczy, jednak najczęściej tylko w pewnych sytuacjach.

Czy więc tylko o liczbie wystąpień wyrazów w tekstach informują wykresy częstotliwości? Otóż nie, ponieważ ważne są jeszcze np. wyjątkowo duże zwiększenia lub zmniejszenia użyć wyrazów w danych kwartałach. Mogą one być źródłem informacji o samym środowisku graczy, o ich działalności, zmienności itd.

Patrząc na dwa najbardziej niestabilne wykresy – *Xbox* (nazwa konsol firmy Microsoft) i *PlayStation* (nazwa konsol firmy Sony) – można zauważyć, że w pierwszej części (do 3 kw. 2013 r.) są one do siebie bardzo podobne. Mają w identycznych miejscach wzrost użycia wyrazów i w podobnych spadek. W przypadku bardzo dużego i nagłego wzrostu (lub bardzo dużego i nagłego spadku) użycia wyrazów należy zadać sobie pytanie, z czego one wynikają? Czy jest jakiś powód tych zmian? W przypadku omawianych wykresów zauważyć można wzrost użycia wyrazów w 1 kwartale 2012 r. oraz 2 kwartale 2013 r. Anomalie te są odbiciem rzeczywistości środowiska graczy wideo w ich socjolekcie. W przypadku pierwszego dużego skoku (1 kw. 2012), to dla wykresu *PlayStation* był on spowodowany premierą nowej konsoli Sony, czyli *PlayStation Vita*, która w Europie miała miejsce 22 lutego. Natomiast w przypadku wykresu *Xbox*, wzrost ten był spowodowany tym, że zbliżała się konferencja *Consumer Electronic Show*, podczas której Microsoft miał przedstawić pierwsze informacje na temat nowej konsoli. Ostatecznie nic nie przedstawiono, a opowiedziano tylko o ówczesnej konsoli – *Xboksie 360*.

Odpowiedzią na duży wzrost wykorzystania wyrazów z pól *PlayStation* i *Xbox* w 2 kw. 2013 roku jest to, że wtedy zaprezentowane zostały konsole nowej generacji: *Xbox One*

i Playstation 4. Prezentacja miała miejsce na najważniejszych na świecie targach dla graczy, czyli E3, które odbywały się od 11 do 13 czerwca 2013 roku w Los Angeles.

W tym miejscu trzeba zwrócić uwagę na jedną ważną rzecz związaną z interpretacją danych na wykresach, mianowicie na zmniejszoną skalę. Zaburza ona nieco, w pewnych sytuacjach, postrzeganie informacji tam zamieszczonych, np. czasami trudno jednoznacznie stwierdzić, który z wykresów miał większe wahania, a który mniejsze, ponieważ zmniejszona skala zniekształcała postrzeganie. Jednym z rozwiązań tego problemu byłoby stworzenie olbrzymich wykresów, pozbawionych zmniejszonej skali, gdzie jeden centymetr (lub nawet milimetr) odpowiadałby jednemu wyrazowi graficznemu. Jednak jest to niewykonalne z powodów technicznych. Innym rozwiązaniem jest posłużenie się wzorami matematycznymi. Aby możliwie obiektywnie stwierdzić, który z wykresów był mniej stabilny i wyniki porównać z innymi wykresami, obliczyłem, jakie jest średnie odchylenie sumy wyrazów w danych kwartałach względem średniej dla wszystkich wyrazów. Innymi słowy, posłużyłem się sposobem na wyliczenie odchylenia średniego, którego działanie przedstawię na przykładzie pola *craft*.

Odnoszę się teraz do danych przedstawionych na stronie 208, a konkretnie do tabeli po prawej stronie. Średnia liczba wyrazów graficznych tego pola w jednym kwartale wynosiła 32,15, obliczyłem to dodając wszystkie liczby z kolumny i dzieląc przez 20 (liczba kwartałów). Liczba wyrazów z pierwszego kwartału różni się od średniej o 12,15 ($32,15 - 20$), w drugiego kwartału o 6,15 ($32,15 - 26$), trzeciego kwartału o 11,15 ($32,15 - 21$) itd. Po zsumowaniu wszystkich wartości bezwzględnych, wyszła wartość – 151,6. Ostatnim działaniem było podzielenie uzyskanej wartości przez 20 (liczba kwartałów). W ten sposób otrzymałem – 7,58 (po zaokrągleniu 8). Liczba ta oznacza, że średnio w każdym kwartale jest używanych o 8 więcej lub mniej wyrazów względem średniej. Dane pozyskane w ten sposób odzwierciedlały zmiany w pojedynczych polach. Dla celów porównawczych należało wprowadzić bardziej zaawansowane obliczenia.

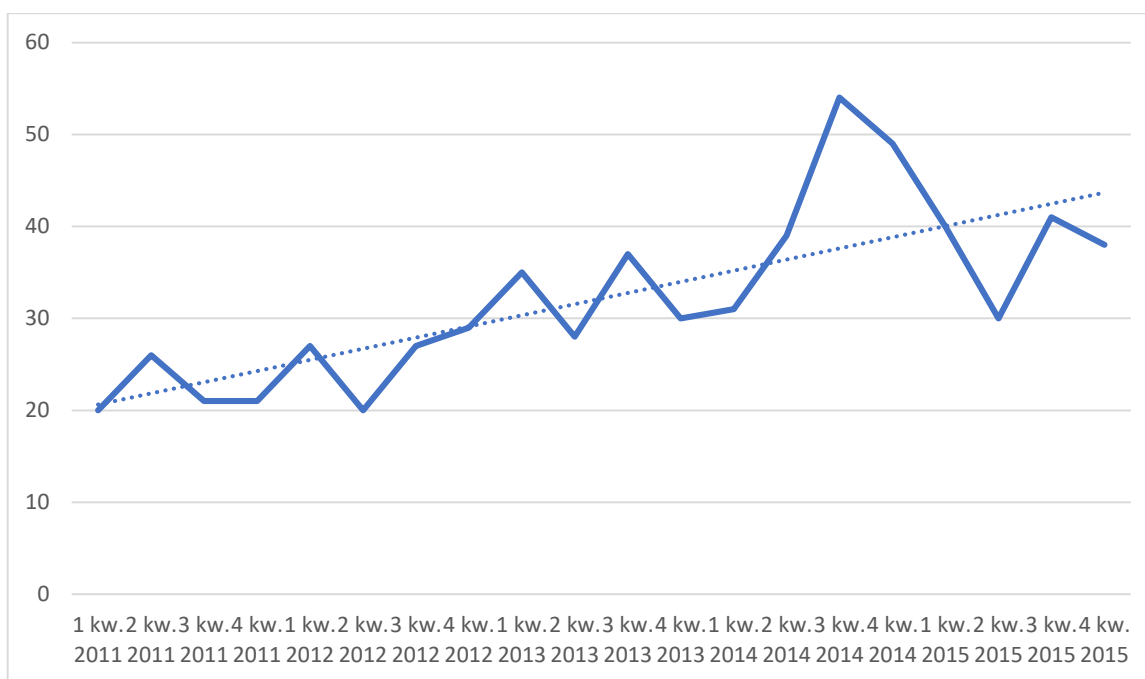
W celu zrównania wyników z różnych pól musiałem dla każdego z nich wyliczyć współczynnik zmienności wykresu. Uczyniłem to, dzieląc wynik odchylenia danego pola, w przypadku pola *craft* było to 7,58, przez średnią dla wszystkich wyrazów w tym polu – 32,15. W ten sposób otrzymałem wynik 0,235. Przemnożyłem go przez 100% i uzyskałem współczynnik zmienności na poziomie 23,5%. Tak uzyskane wyniki zestawilem ze sobą i przeanalizowałem, ale to już zaprezentuję w dalszej części pracy. Teraz natomiast przyjrę się wynikom dla dziesięciu najobszerniejszych wyrazowo pól:

komputer – 2,823%
 RPG – 4,135%
 trailer – 11,000%
 multi – 7,821%
 gameplay – 11,291%
 DLC – 9,517%
 PlayStation – 18,897%
 Xbox – 19,021%
 konsola – 4,967%
 PC – 4,499%

Po przyrównaniu danych procentowych do wykresów graficznych od razu dostrzec można potrzebę stosowania wyliczeń, ponieważ patrząc tylko na samą prezentację graficzną wykresów, można by stwierdzić, że największy współczynnik zmienności ma *PlayStation*, natomiast dane matematyczne pokazują, że nico większy współczynnik ma wykres *Xbox*. Winna temu jest skala, ale rozwiązaniem tego problemu jest właśnie zastosowanie działań matematycznych.

Analiza wykresów dostarcza jeszcze jednej ważnej informacji, a mianowicie takiej, czy wykres ma tendencję rosnącą czy też malejącą.

Odwołując się do wykresu pola *craft*, od razu widać, że jego tendencja jest rosnąca. Potwierdza to także poniższa wizualizacja, w której umieściłem dodatkowo linie trendu.



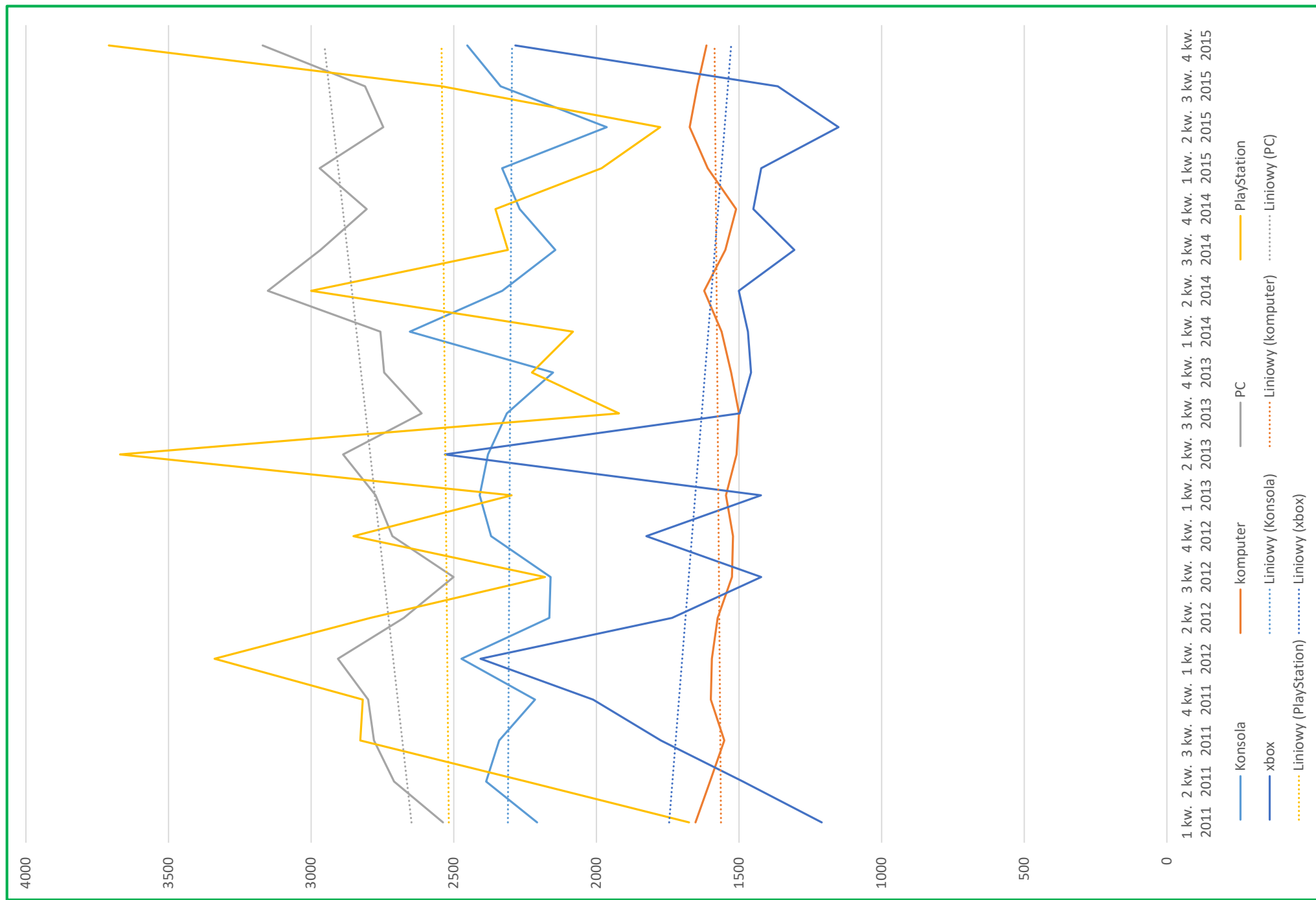
Wykres 21. Linią trendu dla pola wyrazowego *craft*

Widoczność tendencji na wykresach zależy od ich dokładności i skali. Czasem trzeba ją maksymalnie uszczegółowić, aby dostrzec, w którą stronę zmierza. Możliwość dostrzeżenia

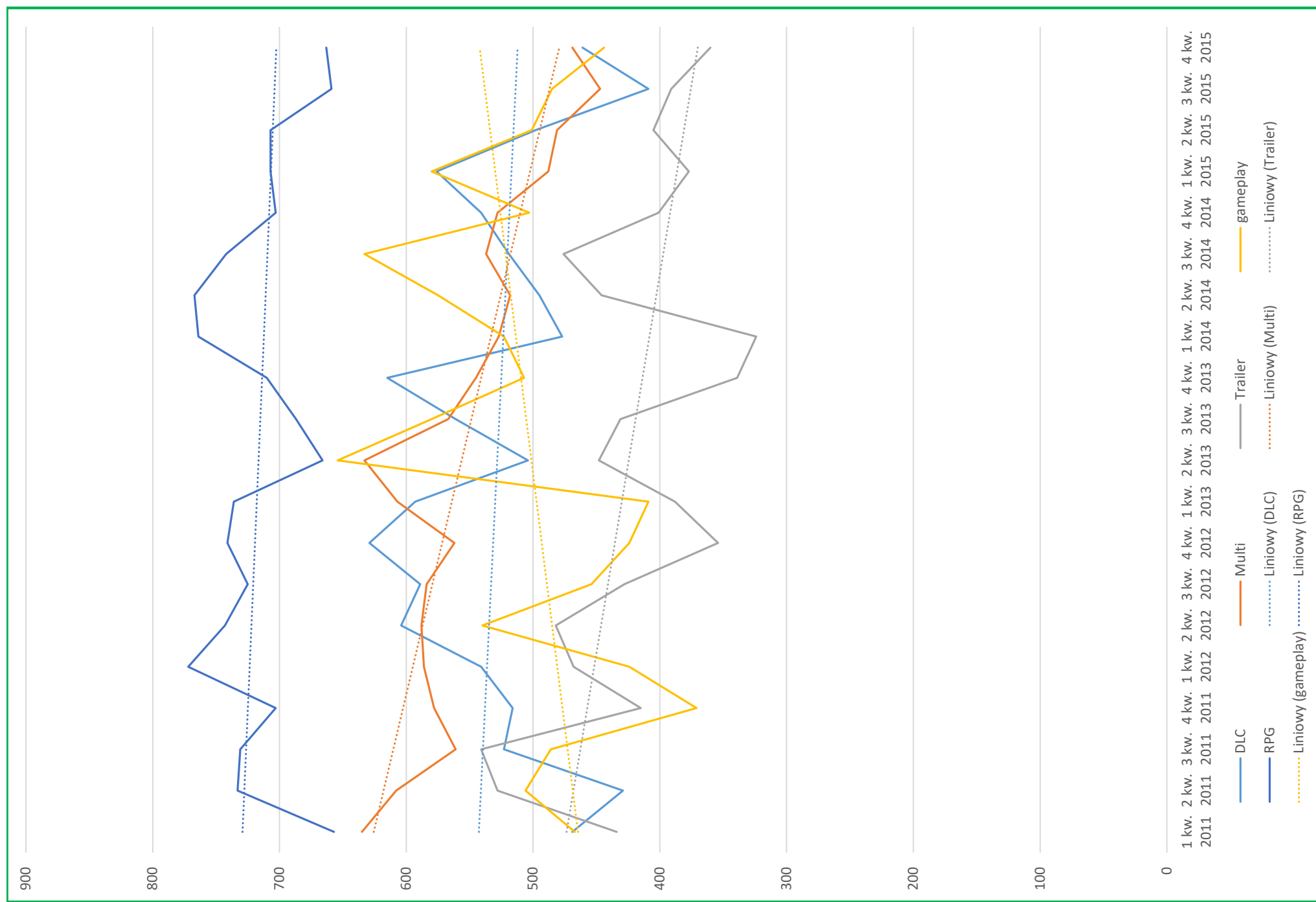
tendencji na wykresach zależy od ich dokładności i skali. Problem uchwycenia tendencji dotyczy także wykresu 19 (oraz 20), gdzie skumulowane zostały dane z 10 pól, które różnią się między sobą sumą wyrazów. Wszystkie one zostały ujęte w jedną prezentację graficzną, dlatego skala powierzchni została dostosowana do wszystkich danych. W efekcie część informacji na temat tendencji konkretnych pól stała się nieczytelna.

Rozwiązaniem tego problemu byłoby pogrupowanie pól na mniejsze zbiory (po 5) i zwężenie obszaru wykresu. Po tym zabiegu większość tendencji stałaby się czytelna (patrz wykresy 22 i 23).

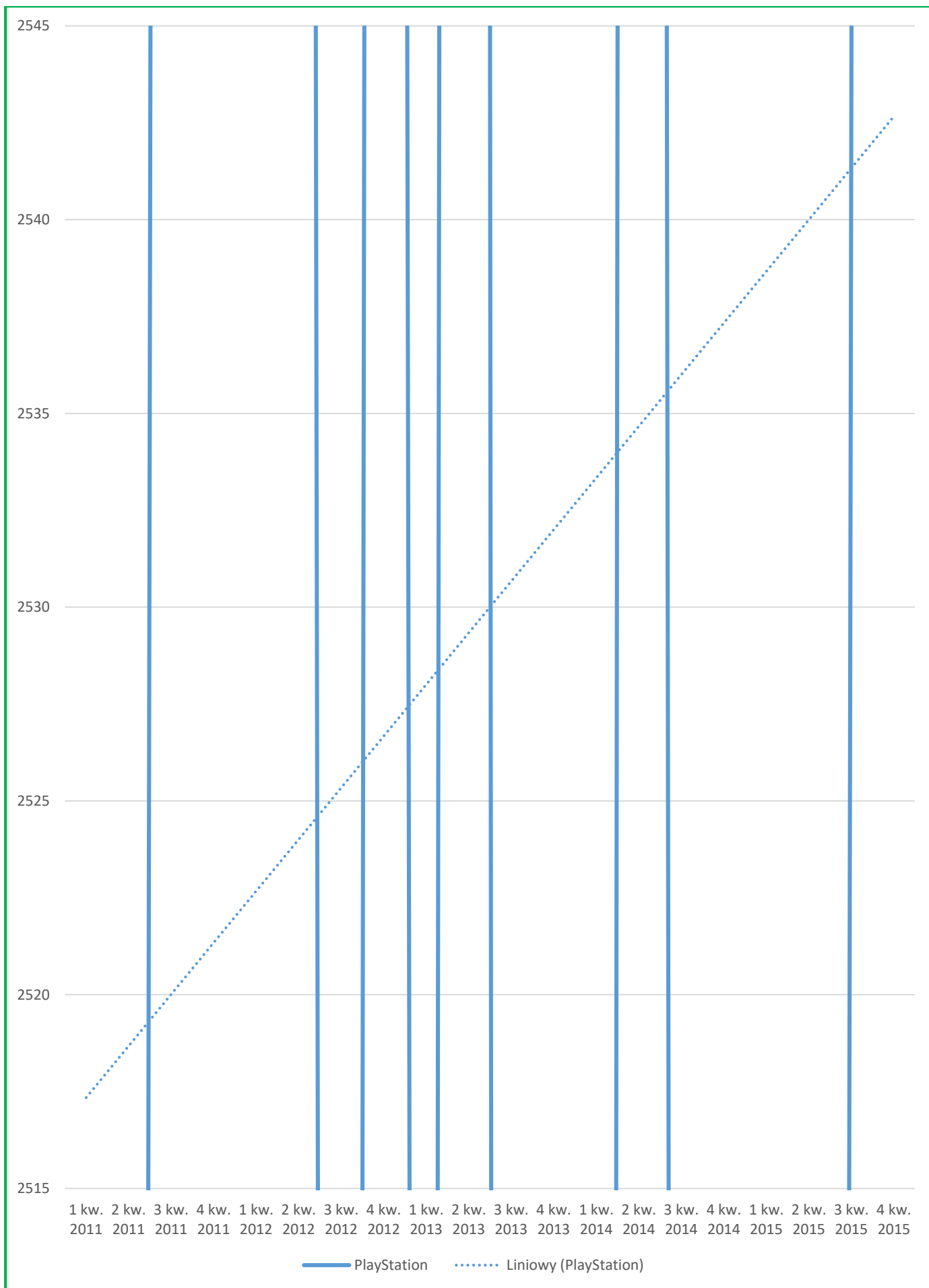
Innym rozwiązaniem problemu byłoby maksymalne powiększeniu fragmentu wykresu danego pola, na którym znajduje się linia trendu. Wtedy bardzo dobrze widać, jaka jest tendencja konkretnego pola (patrz wykres 24).



Wykres 22. Graficzne przedstawienie 10 najbogatszych w wyrazy graficzne pól platformy zielonej (część 1)



Wykres 23. Graficzne przedstawienie 10 najbogatszych w wyrazy graficzne pól platformy zielonej (część 2)



Wykres 24. Powiększeniu fragmentu wykresu pola *PlayStation*

Reasumując, tendencję wzrostową wykazują takie pola, jak: *gameplay*, *komputer*, *PC*, *PlayStation*. Natomiast pozostałe pola z grupy dziesięciu najobszerniejszych wyrazowo wykazują tendencję spadkową.

W kolejnej części analizie zostanie poddanych dziesięć najobszerniejszych pól, ale w odniesieniu do platformy czerwonej:

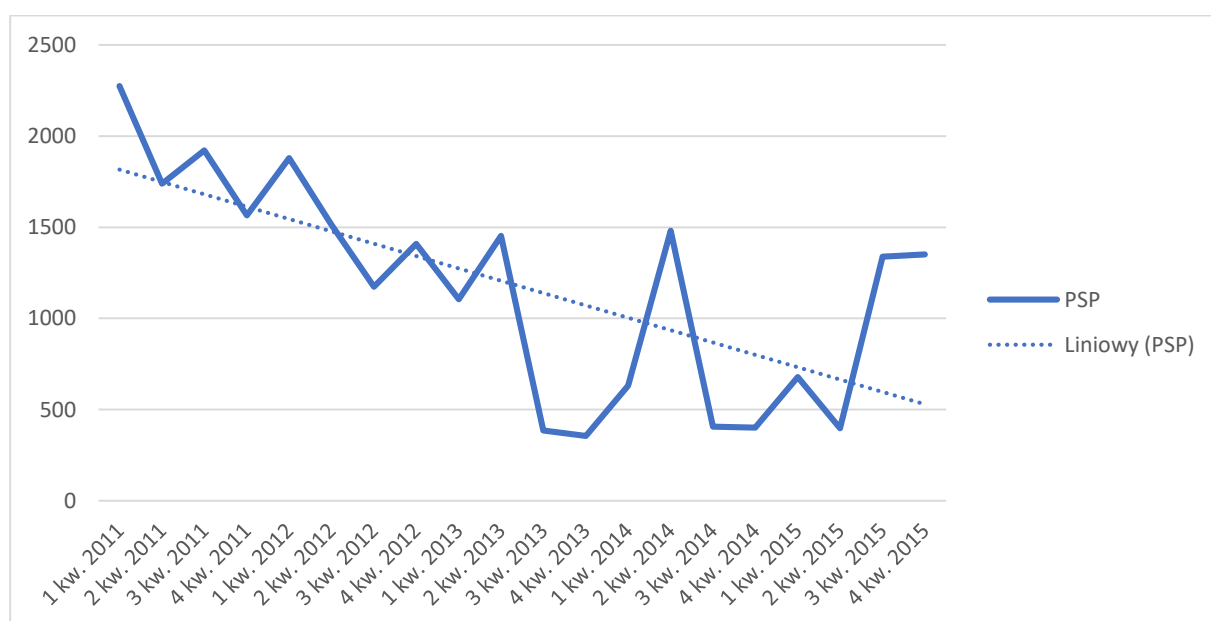
W porównaniu z wykresem platformy zielonej widać, że zestaw dziesięciu najobszerniejszych pól pod względem liczby wyrazów graficznych uległ zmianie. Obecnie zamiast pola *trailer* w zestawieniu jest pole *PSP*. Inne wykresy stały się bardziej niestabilne, tym samym zwiększył się ich współczynnik zmienności:

DLC – 21,034%
 konsola – 6,971%
 multi – 14,391%
 gameplay – 25,382%
 komputer – 6,490%
 PC – 11,882%
 PlayStation – 14,705%
 PSP – 42,829%
 RPG – 7,946%
 Xbox – 16,029%

Zmniejszeniu uległy natomiast tylko te, z powyżej zaprezentowanych współczynników, które dotyczyły pola *PlayStation* (zmiana o 4,192%) oraz *Xbox* (zmiana o 2,992%). Reszta pól zwiększyła swoje współczynniki, co potwierdza większą niestabilność na wykresie graficznym.

Warto zwrócić uwagę, że w tym zestawieniu, w wykresie *PlayStation*, wyeksponowany został kolejny szczytowy punkt, który dotyczy 2 kwartału 2014 roku. W wykresie zielonym punkt ten także posiadał wysoką wartość, ale dopiero tutaj znacznie wyróżnił się na tle innych. Anomalia ta wynika z tego, że w tym kwartale odbyła się kolejna edycja targów E3 (10–12.06.2014), na których gracze spodziewali się prezentacji wielu nowych gier na konsolę *PlayStation 4*, potwierdza tę teorię fragment z czasopisma „PSX Extreme”: „w zeszłym roku poznaliśmy w L.A. (Los Angeles – przyp. R.M.) konsole ósmej generacji. Teraz liczymy na gry. Dużo gier” [PSX, s. 3].

Zdecydowanie na tle innych współczynników wyróżnia się wielkością ten, który dotyczy pola *PSP*. Wykres tego pola po wyodrębnieniu i zwiększeniu skali wygląda następująco:



Wykres 25. Linia trendu dla pola *PSP*

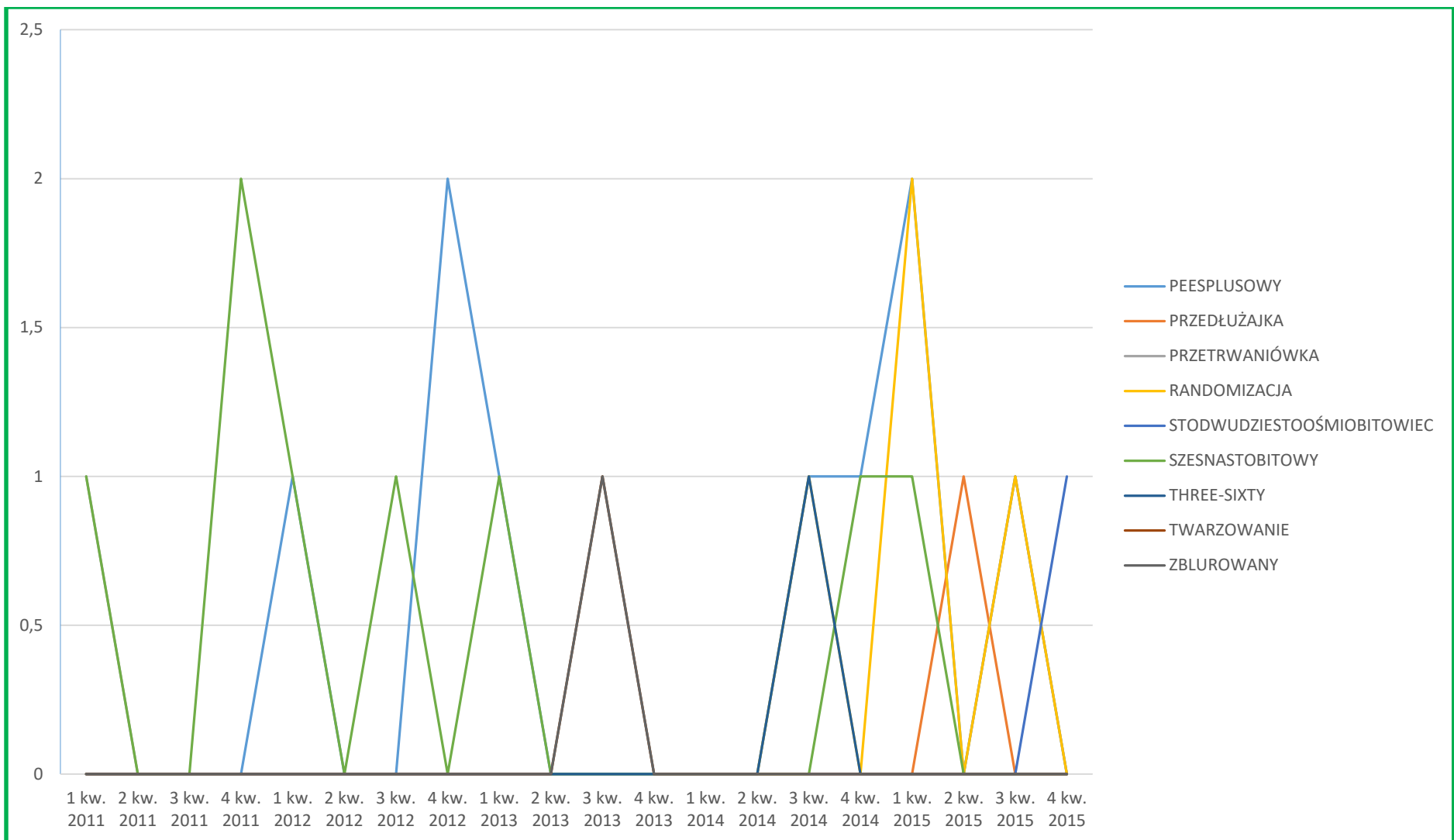
Z wykresu wynika, że pole to ma tendencję spadkową. Połączyć należy tę informację z faktem, że w pierwszym kwartale 2012 roku premierę miała konsola PS Vita, która jest następcą konsoli PSP. Premiera ta przyczyniała się do tego, że coraz częściej pisano o konsoli nowszej, natomiast coraz rzadziej o konsoli PSP, która stopniowo przestała być ważnym tematem w środowisku graczy. Na wykresie widać, że były jeszcze momenty, kiedy więcej używano słownictwa z tego pola, ale ogólna jego tendencja jest spadkowa.

Poza polem *PSP*, tendencję spadkową mają także: *xbox*, *multi* i *DLC*. Reszta ma natomiast tendencję wzrostową.

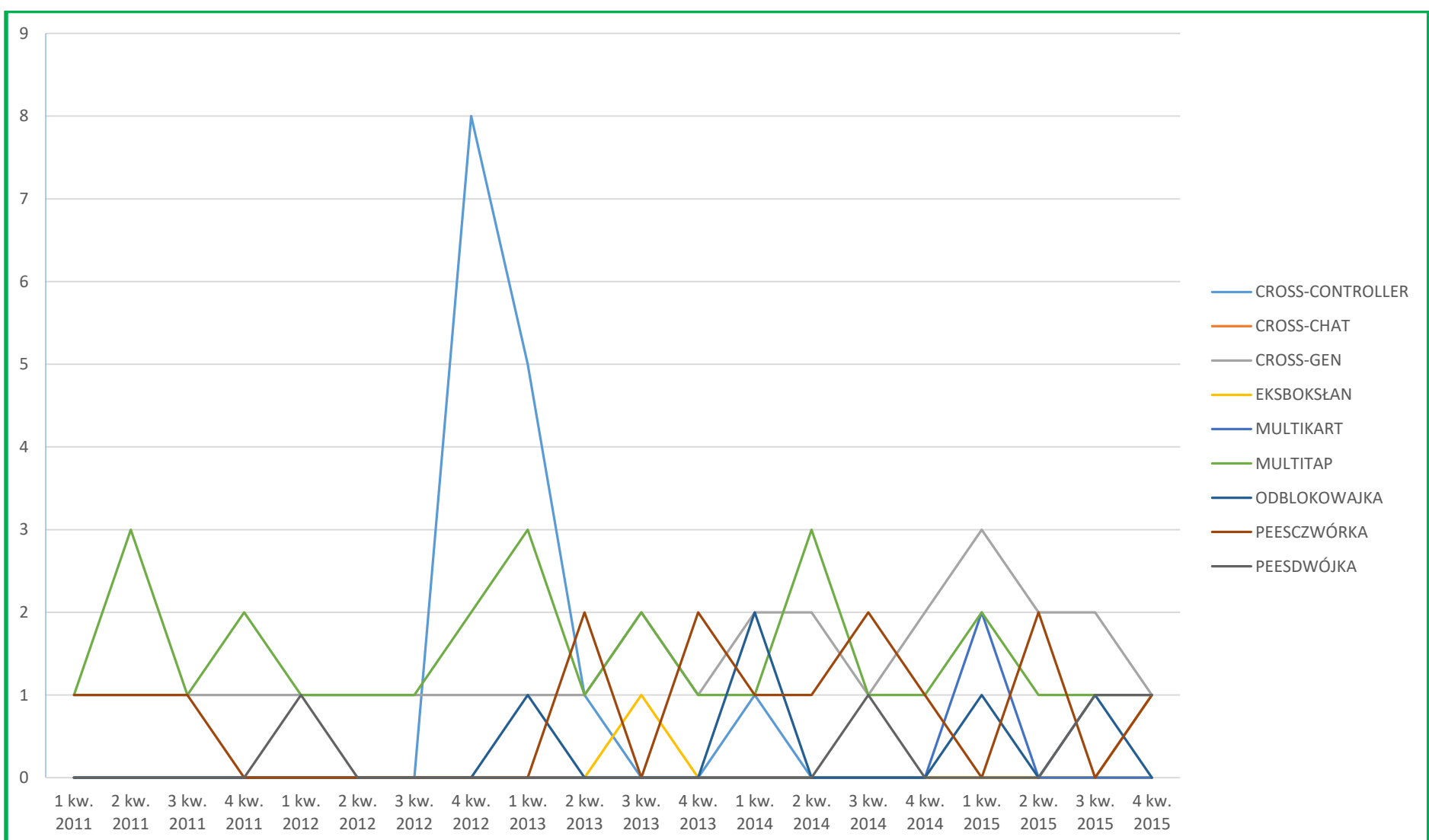
Zestawiając dane z obu płaszczyzn, niezaprzeczalnie można stwierdzić, że trzy pola, te które w obu przypadkach miały linię trendu malejącą – *xbox*, *multi* i *DLC*, rzeczywiście taką tendencję posiadają. Badanie przeprowadzone na dwóch płaszczyznach które, dają podobne wyniki, tylko w tym utwierdza. Natomiast w przypadku tendencji wzrostowej, to badanie na płaszczyźnie czerwonej jeszcze bardziej utwierdziło w przekonaniu, że wyrazy z pól: *gameplay*, *komputer*, *PC*, *PlayStation* rzeczywiście są coraz częściej używane w tekstach graczy.

17.2. ANALIZA WYKRESÓW PLATFORMY ZIELONEJ I CZERWONEJ W ODNIESIENIU DO PÓŁ NAJUBOŻSZYCH W WYRAZY GRAFICZNE

Warto przyjrzeć się jeszcze dziesięciu polom, które są najuboższe wyrazowo. Początkowo przeanalizuję wyniki dla platformy zielonej:



Wykres 26. Graficzne przedstawienie grupy 10 najuboższych w wyrazy graficzne pól platformy zielonej (część 1)



Wykres 27. Graficzne przedstawienie grupy 10 najuboższych w wyrazy graficzne pól platformy zielonej (część 2)

Wykresy tej grupy, zdecydowanie różnią się od tych prezentowanych powyżej. Różnica jest np. w skali, która jest skrajnie zminimalizowana, ponieważ jednemu wyrazowi graficznemu odpowiada na wykresie jedno pole. Problemem jest natomiast wielkość danych tej grupy, bo czy w przypadku tak małej liczby wyrazów w polach, można mówić o jakiegokolwiek częstotliwości? Moim zdaniem tak, ale tylko w bardzo ograniczonym stopniu, ponieważ dane są zwyczajnie zbyt małe. Powyższy wykres informuje jedynie o tym, czy dany wyraz graficzny z danego pola wystąpił w danym okresie lub czy wtedy go nie używano.

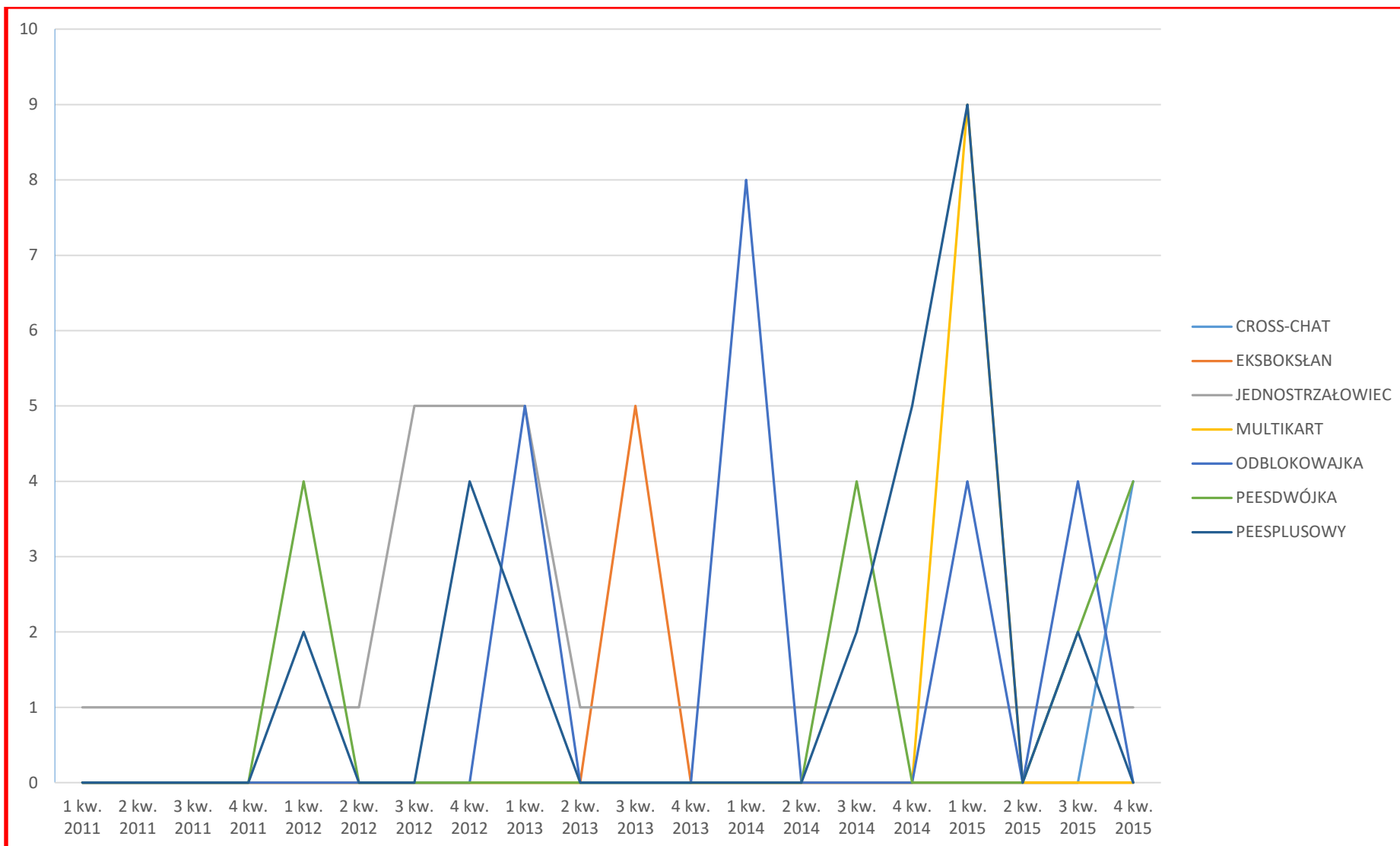
Jednak na tle całości wyróżnia się pole *cross-controller*, którego wyrazy w 4 kwartale 2012 roku wystąpiły aż 8 razy. Wynika to z tego, że w tym okresie ujawniona została informacja, o dostępności opcji *cross-controller* między konsolą PS Vita a PS3⁸¹.

W przypadku tej grupy nie można wyliczyć ani współczynnika zmienności, ani pokazać linii trendu, bo jest ona zbyt mała.

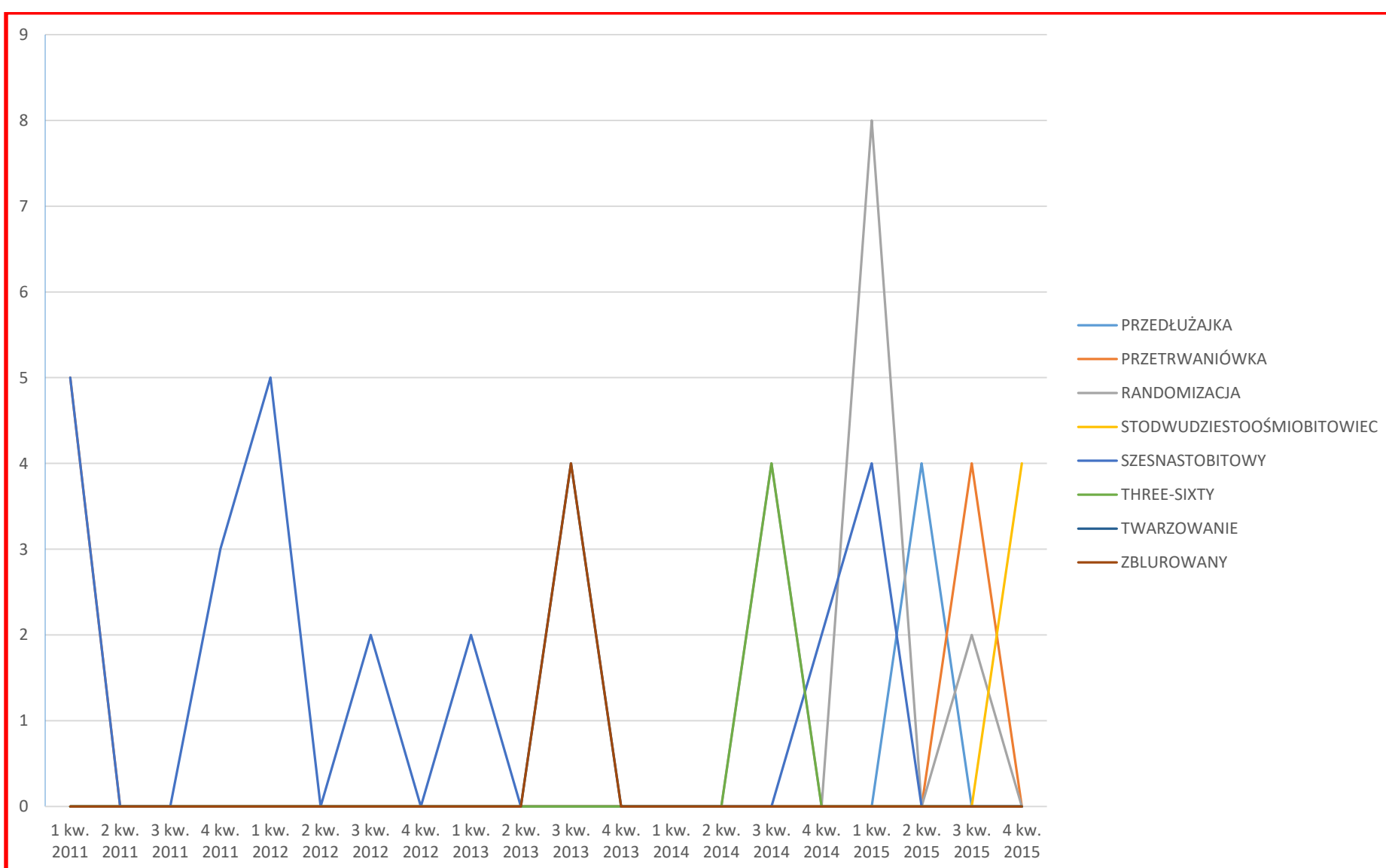
Pola z tej grupy należą do kręgu terminów profesjonalnych (profesjonalizmów), są kreatywnymi tworem środowiska lub fragmentami większych pól znaczeniowych, które nie zostały włączone do badań.

W celu uzupełnienia informacji, przytaczam jeszcze wykres dla płaszczyzny czerwonej:

⁸¹ PS Vita jest konsola przenośną, natomiast PS3 stacjonarną. Opcja cross-controller pozwala na możliwość użycia przenośnej konsoli, jako kontrolera do konsoli stacjonarnej.



Wykres 28. Graficzne przedstawienie grupy 10 najuboższych w wyrazy graficzne pól platformy czerwonej (część 1)



Wykres 29. Graficzne przedstawienie grupy 10 najuboższych w wyrazy graficzne pól platformy czerwonej (część 2)

Po zastosowaniu mnożników zrównujących do płaszczyzny czerwonej, pole *cross-controller* przestało być uwzględniane w tej grupie, czyli w grupie najuboższej wyrazowo i dlatego też nie ma jego wykresu w powyższym zestawieniu.

Na uwagę zasługuje wykres pola *jednostrzałowiec*⁸², którego przebieg ani razu nie osiągał wartości zerowej. Pole *jednostrzałowiec* w 3 kwartałach odnotowało wyższe użycie, natomiast w pozostałych, używane było jednokrotnie.

17.3. ANALIZA WYBRANYCH WYKRESÓW PLATFORMY ZIELONEJ I CZERWONEJ W ODNIESIENIU DO PÓL, KTÓRE OSIĄGAJĄ ŚREDNIE WARTOŚCI NASYCENIA WYRAZAMI GRAFICZNYMI (NIE SĄ ANI NAJUBOŻSZE, ANI NAJBOGATSZE)

W tej części pracy przyjrę się wykresom pozostałych pól na płaszczyźnie zielonej i czerwonej oraz omówię bardziej interesujące przykłady. W tym celu odwoływać się będę do wykresów przedstawionych w aneksie na stronach 55–62 i 115–123, czyli do wykresów od Z(A) do Z(AF) i od C(A) do C(AG). Zbadane zostaną wykresy nietypowe, wyróżniające się przebiegiem na tle innych.

W pierwszej kolejności na uwagę zasługuje wykres pola *padlet*⁸³ (wykres Z(T)). Widać, że duży skok wykorzystania wyrazów tego pola zarejestrowany został w 1 kwartale 2013 roku. Skok ten wiąże się z tym, że pod koniec ostatniego kwartału 2012 r. premierę miała konsola Wii U, której kontrolerem było urządzenie nazywane przez graczy *padletem*. Patrząc na ten wykres, zaobserwować można, że wyrazów z tego pola zaczęto używać dopiero w 3 kwartale 2012 roku, więc widać, że dopiero wtedy to pole powstało – w okolicy premiery konsoli. Z wydarzeniem, jakim była premiera konsoli Wii U powiązać należy także wzrost użycia wyrazów z pola *multitouch*⁸⁴ (wykres Z(U)). *Padlet* w swojej konstrukcji nie posiadał funkcji *multitouch*, co bardzo nie spodobało się graczom, mówiąc kolokwialnie – narzekali oni, że jej nie ma. Raptowne zwiększenie użycia wyrazów graficznych z pola *multitouch*, odnotowane zostało w tym samym okresie, co wzrost w polu *padlet* (1 kwartał 2013).

Kontroler (*padlet*), który niestety nie posiadał funkcji *multitouch*, był głównym, najbardziej atrakcyjnym, elementem konsoli, dzięki któremu wyróżniała się ona na tle innych urządzeń wyposażonych w zwykłe pady. Niestety, pomysł z dołączeniem do konsoli *padletu*

⁸² Gra „na jeden raz”. Produkcja najczęściej krótka i łatwa, którą można ukończyć bardzo szybko.

⁸³ Kontroler do gier konsoli Wii U (produkcji firmy Nintendo). Jest on połączeniem tabletu (posiada ekran dotykowy) oraz pada (przyciski sterujące).

⁸⁴ Funkcja ekranów dotykowych, która pozwala na obsługiwanie ich wieloma palcami, ponieważ ekran odczytuje wiele punktów dotyku, a nie tylko jeden.

zamiast zwykłego *pada* okazał się fiaskiem finansowym i firma produkująca konsolę przegrała batalię z konsolami Sony i Microsoftu.

Z premierami nowych konsol wiąże się także większe używanie wyrazów z pola *next-gen* ‘następna generacja’ (wykres Z(F)). Na polskim rynku gier konsolowych liczą się tak naprawdę konsole dwóch firm – Sony i Microsoftu. Konsole firmy Nintendo stanowią niszę. Gdy zbliża się premiera lub wydarzenie związane z prezentacją nowego sprzętu do gier, to okres ten jest świętem dla graczy. Odzwierciedlenie tego widać także w języku, a co się z tym wiąże także na wykresie *next-gen*. Wykres ten stopniowo rośnie, do momentu premiery konsoli, po czym stopniowo spada, czyli zmniejsza się użycie wyrazów z tego pola. Warto zauważyć, że wyrazy z tego pola są swoistym wyrazami wytrychami. Gracze zazwyczaj nie znają oficjalnej nazwy nowego sprzętu, korzystają więc wtedy z wyrazów pola *next-gen*. Po premierze, gdy jest już znana oficjalna nazwa sprzętu, przestają ich stopniowo używać.

Interesujący jest także przebieg wykresu *remaster* (wykres Z(I)), który pokazuje, jak zmienia się rynek gier. Niegdyś bardzo dużo było na rynku nowości growych, nowych kreatywnych tytułów opartych na przemyślanych scenariuszach, natomiast obecnie coraz więcej jest gier odświeżanych, czyli takich, które pierwotnie wydane zostały kilka/kilkanaście lat temu na sprzęt starszego typu, natomiast teraz ponownie są wydawane, tylko delikatnie zmienione np. poprawiona została w nich grafika. Gier tego typu jest coraz więcej na rynku, a gracze potocznie nazywają je m.in. *remasterami* lub *odgrzewanymi kotletami*, ponieważ dostają jeszcze raz tę samą grę, tylko nieco zmienioną lub uwspółcześnioną. Wykres *remaster* dobrze pokazuje, że gry tego typu są coraz bardziej popularne, co też ma swoje odzwierciedlenie w języku.

Wykresem, na którym, podobnie jak w przypadku powyższym, widać stopniowy wzrost, ale już z pewnymi wahaniami, jest ten, który dotyczy pola *mikrotransakcja* (wykres Z(M)). Wzrost świadczy to o tym, że firmy coraz częściej decydują się na wydanie gry, która teoretycznie jest darmo, ale żeby osiągnąć lepszy wynik lub uzyskać pomoc w rozgrywce, trzeba zapłacić realnymi pieniędzmi. Ten sposób płatności nazywany jest *mikrotransakcjami*. W branży gier wideo jest to coraz popularniejszy sposób pozyskiwania od graczy pieniędzy.

Podobnym wzrostem odznacza się wykres pola *kartridż* (wykres Z(O)). Widać, że tak naprawdę od pierwszego kwartału 2015 r. wyrazy z tego pola coraz częściej są w użyciu. Niegdyś popularne *kartridże* były nośnikami danych z grą jeszcze na początku lat 90-tych, później, w przypadku konsol stacjonarnych, zastąpione zostały przez płyty CD, następnie DVD i Blu-ray. *Kartridż* jako nośnik danych zachował się natomiast w konsolach przenośnych. O grach wydanych na *kartridżach* pisano zdecydowanie rzadziej niż o tych, które wydane

zostały na „duże”, stacjonarne platformy. Dlatego też wyrazy z pola *kartridż* używane były rzadko. Sytuacja zmieniła się, gdy ogłoszono, że nowa „duża” konsola firmy Nintendo, zamiast płyt będzie korzystała z *kartridży*. Przecieki na ten temat zaczęły się pojawiać właśnie w 2015 roku, czyli wtedy, kiedy odnotowano wzrost w polu *kartridż*. Im bliżej premiery konsoli, tym coraz częściej używano wyrazów z tego pola.

W późniejszym okresie, którego już moje badania nie dotyczą, zapewne jeszcze częściej używano wyrazów tego z pola, ponieważ firma Nintendo zdecydowała się na pokrycie nośników gier substancją zabezpieczającą przed spożyciem ich przez dzieci. Gdyby dziecko chciało zjeść *kartridż*, to po wsadzeniu go do ust, poczułoby bardzo nieprzyjemny smak i go wypluło. Starsi gracze, już po premierze konsoli, zainteresowani pomysłem firmy, nie dowierając producentowi, sami lizali *kartridże* i sprawdzali ich smak. W tym kontekście myślę, że wyrazy z tego pola były także często używane.

Na uwagę zasługuje także wykres pola *PSP2* (wykres $Z(X)$). W świecie graczy, jak już wspomniałem, świętem jest prezentacja nowych konsol. Często zdarza się, że ich nazwy nie są wcześniej znane, dlatego gracze albo używają określenia *next-gen*, albo sami tworzą nazwę dla nowego sprzętu. Nazwa ta może powstać, jak w tym przypadku, przez dodanie lub zmianę cyfry po nazwie konsoli obecnej generacji, której następcą dopiero zostanie zaprezentowany. Obecną generacją była wtedy konsolka PlayStation Portable, która w skrócie nazywana była PSP. Gracze, chcąc nawiązać do następcy tej konsolki, dodawali dwójkę po jej nazwie, podkreślając, że jest to druga konsolka z serii PSP. W ten sposób powstała robocza nazwa następcy konsoli PSP, czyli PSP2. Patrząc na wykres tego pola, widać, że jest bardzo często używane na początku 2011 roku, a przez następne trzy kwartały wykorzystanie jego wyrazów spada do zera, ponieważ okazało się, że następcą konsolki PSP, nie jest PSP2, ale PS Vita. Z premierą konsoli PS Vita, wiąże się także spadek użycia wyrazów w polu *PSP GO* (konsola do gier, inna wersja konsoli PSP) (wykres $Z(AA)$). Spadek dodatkowo powiązać można także z tym, że w 2011 roku przestano produkować konsolki PSP GO.

Spadek, co prawda mniej gwałtowny i nie do zera, zaobserwować można także w przypadku pola *Kinect* (wykres $Z(C)$), czyli urządzenia firmy Microsoft, które było kontrolerem ruchu gracza. Urządzenie to było bardzo popularne w 2011 roku (premierę miało w ostatnim kwartale 2010 roku), po czym stopniowo zainteresowanie nim spadało. Winna temu częściowo była firma Microsoft, która wydawała mało gier wykorzystujących urządzenia oraz sami gracze, którzy woleli niestety grać, leżąc na kanapie z padem w dłoni, niż ruszać się przed telewizorem.

Na uwagę zasługuje jeszcze jeden wykres, w którym odnotowano punktowe wzrosty w kilku kwartałach. Jest to pole *używka* ‘używana gra’ (wykres Z(Ł)). Wyrazy tego pola szczególnie często wykorzystywane były w 4 kwartale 2011, 1 kwartale 2012 oraz 1 i 2 kwartale 2013 r., najczęściej natomiast w 3 kwartale 2013 r. Wzrost użycia jest pokłosiem informacji, którą podał Microsoft na targach E3, że aby uruchomić gry używane (*używki*) na nowej konsoli – XBoksie One, trzeba będzie dodatkowo zapłacić. Gracze byli oburzeni, więc często ten temat podejmowali, aż do momentu, w którym Microsoft, widząc reakcję na informację, wycofał się z tego pomysłu.

Chciałbym teraz przyjrzeć się polom z platformy czerwonej, których wykresy dodane zostały w aneksie na strona 115–123.

Po zrównaniu liczby wyrazów do platformy czerwonej, podobnie, jak podczas analizy dziesięciu najobszerniejszych wyrazowo pól tematycznych, także i w tym przypadku wykresy są zdecydowanie mniej stabilne.

Na platformie czerwonej uwydatniły się pewne zmiany w kilku wykresach, które wcześniej były mniej widoczne, lub nie aż tak wyróżniające się na tle innych. Na uwagę zasługuje wykres pola *exclusive* ‘gry na wyłączność dla danej platformy’ (wykres C(D)), w którym zaobserwować można dwa duże wzrosty użycia wyrazów tego pola. Pierwszy z nich miał miejsce w 3 kwartale 2013 roku, a drugi rok później, czyli w 3 kwartale 2014 roku. W światku graczy, głównie konsolowych, ale w mniejszym stopniu także i komputerowych, wśród gier najbardziej liczą się *exclusive*’y. To ich obecność na danym sprzęcie do gier wpływa na decyzję o zakupie tego, a nie innego sprzętu do gier.

Wzrost użycia wyrazów z pola *exclusive* odnotowany został bezpośrednio po targach E3 z 2013 i 2014 roku. W tych kwartałach często pisano o grach pokazanych na targach (łączy się to bezpośrednio z informacją, którą podałem omawiając jedno z dziesięciu najobszerniejszych wyrazowo pól – *PlayStation*, strona 211).

Wyróżnia się także wykres pola *Redzi* (wykres C(AA)). Jest to środowiskowa nazwa pracowników firmy growej CD Project Red, którzy w swoim dorobku mają najlepszą polską grę wideo, czyli *Wiedźmina*. W 2015 roku odbyła się premiera 3 części cyklu, dlatego też na wykresie pola *Redzi* widać znaczny wzrost używania jego wyrazów. Z tym wydarzeniami łączy się także wzrost użycia wyrazów w polu *Wiesiek*⁸⁵ (wykres C(O)). Zwiększone użycie wyrazów

⁸⁵ Polscy gracze często zamieniają oryginalną nazwę gry (i jednocześnie profesję głównego bohatera) *Wiedźmin*, na polskiego *Wieska*. Zamiast powiedzieć „chcesz pograć w Wiedźmina?” powiedzą „chcesz pograć w Wieska?”.

z pola *Wiesiek*, widać także w kwartałach 2011 roku, kiedy to premierę miała 2 część gry o wiedźminie.

17.4. WYBRANE WYKRESY WSZYSTKICH PLATFORM W ODNIESIENIU DO KORPUSÓW *KONSOLOWCÓW ORAZ KOMPUTEROWCÓW*

W tej części pracy dokonam analizy wykresów z uwzględnieniem kryterium użytkownika, czyli z podziałem na *konsolowców* oraz *komputerowców*. Stworzyłem wykresy w odniesieniu do tych grup, ponieważ podobnie jak w przypadku wykresów powstałych w oparciu o dane sumaryczne, także tutaj zaobserwować można pewne anomalie pod względem użycia wyrazów graficznych w danych kwartałach.

Ponadto dzięki zastosowaniu podziału na wykresy *konsolowców* oraz *komputerowców* można je porównać ze sobą, a dzięki temu zobaczyć, jakie występują podobieństwa oraz jakie różnice zachodzą między nimi. Innymi słowy sprawdzić, czy identyczne pola, ale w różnych grupach korpusów, mają podobny przebieg kwartalny, czy też są one diametralnie inne.

Dane wyjściowe do stworzenia konsolowych wykresów pozyskałem przez zsumowanie wartości kwartalnych trzech części korpusu: PSX, NPSX oraz FCDA. W przypadku ostatniej z części musiałem matematycznie wyliczyć rozkład wyrazowy w danych kwartałach, co zrobiłem w identyczny sposób, jak to przedstawiłem na stronie 208.

Do stworzenia wykresów *komputerowców* wykorzystałem i połączyłem dane z takich części, jak: CDA, NCDA i FCDA (wyliczając podział kwartalny).

Podkreślę w tym miejscu jeszcze raz, że omówię w tej części jedynie kilka moim zdaniem interesujących przykładów konkretnych wykresów pól wyrazowych. Wszystkie wykresy dodałem w aneksie, włącznie z tymi, których omówienie pominąłem.

Na początku warto przyjrzeć się wykresom, które dotyczą pola wyrazowego *Marian* ‘bohater gry Super Mario Bros’ (wykresy Z(AR), C(AQ), Z(CT) i C(CW)).

Po zestawieniu ze sobą wykresów platformy zielonej zarówno *komputerowców* (Z(CT)), jak i *konsolowców* (Z(AR)) widać, że na wykresie odnoszącym się do tekstów *komputerowców* wyrazy pola *Marian* są bardzo rzadko w nich wykorzystywane. Widać, że są nawet takie kwartały, w których ani razu nie użyto wyrazów tego pola, a gdy już doszło do użycia, to były to wystąpienia jednostkowe w wybranych kwartałach.

Zgoła odmienna sytuacja prezentuje się na wykresie *konsolowców*, w którym wyrazy pola *Marian* można znaleźć w każdym badanym kwartale, a co warto podkreślić, użycie tych

wyrazów nie spada poniżej 15 jednostek w ani jednym kwartale. Ponadto na wykresie *konsolowców* zauważyć można, że w 4 kwartale 2012 roku nastąpił gwałtowny wzrost użycia jednostek tego pola. Wzrost ten co prawda widać już na wykresie sumarycznym (Z(O)), ale dopiero po wydzieleniu wykresów konsolowych i komputerowych można stwierdzić, że wzrost użycia nastąpił wyłącznie w tekstach *konsolowców*.

Wzrost użycia wyrazów w tym kwartale łączy się z premierą konsoli Wii U (18.11.2012 w Ameryce Północnej, 30.11.2012 w Europie oraz 8.12.2012 w Japonii), która zadebiutowała wraz z różnymi tytułami startowymi ‘gry, które dostępne są już w dniu premiery nowej platformy’, wśród których znajdował się New Super Mario Bros.U, czyli kolejna gra z serii, w której bohaterem jest Mario/Marian.

Platforma czerwona *konsolowców* z wykresem pola *Marian* (wykres C(AQ)) uwypukla jeszcze jeden wzrost użycia jego wyrazów, który nie był widoczny na platformie zielonej, a miał on miejsce w 1 i 2 kwartale 2015 roku. Ten z kolei wzrost związany jest ze zbliżającą się wtedy premierą gry Super Mario Maker, która to miała miejsce 10.09.2015 roku. Jednak już na targach E3 (16.06.2015–18.06.2015) można było zapoznać się i zagrać w niewydaną jeszcze produkcję.

Interesujący jest także jednostkowy/kwartalny wzrost użycia wyrazów innego pola, które być może nie jest związane bezpośrednio, ale na pewno pośrednio z pewną grą wideo, czyli z MineCraftem. Mowa tutaj oczywiście o polu *craft* (jego wykresy: Z(A Y), Z(CA), C(BB), C(BZ)).

Zarówno na platformie zielonej, jak i czerwonej w odniesieniu do wykresów konsolowych widać, że rozkład wyrazów graficznych w poszczególnych kwartałach jest mniej więcej podobny. W przypadku analizy wykresów komputerowych (czerwonego i zielonego) można natomiast dostrzec, że w 3 kw. 2013 roku użyto najwięcej wyrazów graficznych w odniesieniu do pięciu badanych lat. Jak już wspomniałem wzrost ten wiąże się z grą – głównie komputerową – MineCraft, w której *craft/crafting* jest mechaniką dominującą i na niej w zasadzie opiera się cała gra.

15 września 2014 roku miało miejsce ważne wydarzenie związane z firmą (Mojang) odpowiedzialną za produkcję tej gry. Mojang został wykupiony przez giganta branży gier wideo – Microsoft, za około 2,5 miliarda dolarów. W tamtym okresie był to najważniejszy temat w całej branży gier wideo, dlatego też dużo pisano o samej grze, a co się z tym wiąże, używano wielu wyrazów z pola *craft*.

Innym ciekawym wykresem pola z podziałem na kwartały jest *headset* ‘zestaw słuchawkowy połączony z mikrofonem’, reprezentujący pole z kręgu „sprzęt graczy wideo” (jego wykresy to Z(A_Y), Z(C_L), C(A_U) i C(C_L)).

Podobnie jak w przypadku pierwszego omówionego w tej części pola (*Marian*), także i tutaj wyrazy pola *headset* częściej występują w tekstach *konsolowców* niż *komputerowców*. U *konsolowców* (Z(A_Y)) użycie kwartały wyrazów graficznych nie spada poniżej 6 jednostek, uwzględniając platformę zieloną, natomiast u *konsolowców* (Z(C_L)) zazwyczaj jest to wartość dużo niższa.

Podczas analizy wykresu pola *headset* widać także, że u *konsolowców* są wyraźnie zarysowane punkty, w których widoczny jest znaczny wzrost użycia wyrazów graficznych. Punkty te dotyczą 3 i 4 kwartału 2013 roku oraz 1 kwartału roku 2014. U komputerowców żadnego spektakularnego wzrostu nie odnotowano. Poszczególne anomalie w przypadku wykresów *konsolowców* należy powiązać z premierą konsol PlayStation 4 (Ameryka Północna – 15.11.2013, Europa – 29.11.2013) oraz Xbox One (22.11.2013).

Headset jest obecnie dla wielu graczy nieodzownym elementem konsoli, czymś bez czego nie da się grać oraz sprzętem, który ułatwia komunikację w grach w trybie multi. Wzrost użycia wyrazów tego pola jeszcze wyraźniej widać na platformie czerwonej *konsolowców* (C(A_U)).

Pozostając w kręgu tematycznym „sprzęt graczy”, trzeba przyjrzeć się wykresom pola *smartfon* (Z(B_R), Z(A_M), C(A_L), C(B_R)) oraz *notebook* (wykresy Z(B_H), Z(B_Y), C(B_H), C(B_X)) w odniesieniu do obu grup graczy wideo.

W przypadku zielonego wykresu *konsolowców* zaobserwować można, że radykalny wzrost użycia wyrazów pola *smartfon* nastąpił w 1 kwartale 2012 roku. Analizowany wykres można więc podzielić na dwie części – przed tym kwartałem i po nim. W części przed granicznym kwartałem odnotowano znacznie mniejsze kwartalne użycie wyrazów w tekstach *konsolowców*, natomiast po tym okresie ich liczebność kwartalna wzrosła i praktycznie już nie powróciła do stanu przed 1 kwartałem 2012 roku.

Platforma czerwona (C(A_L)) również pokazuje, że nastąpiła znaczna zmiana w liczbie użycia wyrazów.

Na wykresach *komputerowców* także widoczny jest wzrost użycia wyrazów tego pola w poszczególnych kwartałach, z tym że był on stopniowy, więc ani na platformie zielonej, ani czerwonej nie widać żadnego punktu przełomowego.

Wzrost użycia wyrazów badanego pola we wszystkich wykresach wytłumaczyć należy tym, że zwiększa się popularność oraz liczba urządzeń elektronicznych typu smartfon w grupach graczy wideo.

Nieco inaczej sprawa wygląda w przypadku pola *notebook*. W odniesieniu do wykresów platformy zielonej i czerwonej *konsolowców* widać, że pole to posiada znikomą liczbę użyć wyrazów graficznych w danych kwartałach, a jeśli już one występują to jednostkowo. Zupełnie inaczej sprawa przedstawia się podczas analizy wykresów *komputerowców*. Przyglądając się wykresowi zielonemu można zauważyć, że pole *notebook* w 1 kwartale 2011 roku posiadało dużą liczbę użyć wyrazów, która wynosi ponad 80 jednostek. Po tym kwartale natomiast widać stopniowy, z nielicznymi wahaniami, spadek użycia wyrazów. Co warto podkreślić to fakt, że pole *notebook* w badanym okresie już nigdy nie osiągnęło pułapu liczby wyrazów z pierwszego kwartału 2011 roku. Stopniowy spadek użycia wyrazów pola *notebook* widać także na platformie czerwonej.

Na zakończenie tej części pragnę zwrócić jeszcze uwagę na zielone wykresy pola *free-to-play* (Z(AX), Z(CB)). Zestawiając je ze sobą zaobserwować można, że największy skok użycia wyrazów pola *free-to-play* odnotowano w 1 kwartale 2013 roku. Widać go na wykresie *komputerowców*. U *konsolowców* także widać pewne punktowe wzrosty, ale żaden z nich nie może równać się z tym jednym u *komputerowców*. Wzrost ten spowodowany zapewne był tym, że pod koniec roku 2012 roku bardzo popularna gra internetowa – *Star Wars: The Old Republic* przeszła z modelu płatności abonamentowej na model płatności *free-to-play*.

Powyżej zaprezentowałem wybrane komentarze dotyczące kilku ciekawszych wykresów. W następnym podrozdziale przejdę do analizy, która oparta została na danych pozyskanych z tekstów czasopism oraz newsów.

17.5. WYBRANE WYKRESY WSZYSTKICH PLATFORM W ODNIESIENIU DO KORPUSÓW CZASOPISM ORAZ NEWSÓW

Niniejszy podrozdział poświęcony został analizie wykresów konkretnych pól w odniesieniu do tekstów czasopism oraz newsów. Wykresy podzielone zostały na te, które powstały w wyniku wyszukania wyrazów danych pól w tekstach czasopism, czyli są efektem połączenia danych liczbowych z części korpusu: CDA i PSX oraz te, które uzyskano ze scalenia NCDA i NPSX.

W przypadku wykresów pól, które dotyczą newsów, będą one rozpatrywane w odniesieniu do trzech platform: zielonej, żółtej i czerwonej. Wykresy pól pozyskanych

z czasopism będą natomiast analizowane tylko na platformach zielonej i czerwonej. Przed przystąpieniem do analizy trzeba jeszcze wyjaśnić, dlaczego pominąłem i nie stworzyłem wykresów dla danych pozyskanych z forów internetowych. Uczyniłem tak, ponieważ podczas zbierania materiału nie było możliwe, żeby w pojedynkę dokonać podziału całego materiału na kwartały (dokładniej pisałem o tym na stronie 98).

Na początku analizy w tej części pracy przyjrę się wykresom pola *demo* (Z(DD), Z(EŁ), Ż(C), C(DE), C(EN)).

Patrząc wyłącznie na wykresy newsów (zielony, żółty i czerwony) można dostrzec, że ogólna tendencja użycia wyrazów w kolejnych kwartałach jest spadkowa. Analizując natomiast wykresy pola *demo* w odniesieniu do czasopism można zauważyć, że na platformie zielonej liczba użyć wyrazów jest delikatnie rosnąca, zaś na platformie czerwonej delikatnie malejąca. Zatem wysnuć należy wniosek, że użycie wyrazów w poszczególnych kwartałach jest mniej więcej równe. Ponadto, przyglądając się wyłącznie wykresom czasopism, dostrzec można jeden kwartał wyróżniający się pod względem użycia wyrazów graficznych – 3 kwartał 2013.

Jest on spowodowany tym, że dziennikarze, którzy uczestniczyli w targach E3 (10–12.06.2014) i mieli dostęp do różnego rodzaju gier w wersji *demo*, opisali je w jednym z numerów pisma. Warto podkreślić, że nie były to dema na zwykłą platformę do gier, ale dema gier na konsole nowej generacji, a wszystko związane z kolejną generacją jest warte opisanie, bo tego typu tematy są bardzo poczytne w środowisku. Zatem dlaczego wzrost wystąpił tylko w czasopismach? Moim zdaniem jest tak, dlatego że czasopismo ma określony cykl wydawniczy i materiały, które są publikowane w kolejnych numerach zawsze dotyczą poprzedniego miesiąca. Dlatego w numerze czasopisma nastąpiła kumulacja opisów wszystkich dem zaprezentowanych na targach. Natomiast w przypadku newsów nie ma cyklu wydawniczego, więc informację można dodać dowolnego dnia, a co więcej jest ona natychmiast dostępna czytelnikom. Przed targami E3 zawsze są jakieś przecieki, dlatego część dem jest dostępna jeszcze przed ich rozpoczęciem, więc gdy redaktor newsów zdobędzie takie *demo* i zagra w nie, od razu może opisać swoje wrażenia i opublikować je jako news. Dlatego w przypadku newsów informacje o demach rozłożone zostały na dwa kwartały – 2 i 3 roku 2014, a w przypadku czasopism musiały zostać skumulowane w jednym.

Ciekawę są także wykresy uwzględniające wyrazy z pola *patch* (Z(DE), Z(EŁ), Ż(C), C(DF), C(EN)). We wszystkich wykresach (na wszystkich platformach) widać, że pole to wykazuje tendencję wzrostową. Szczególnie duży wzrost użycia wyrazów tego pola odnotowany został w 1 kwartale 2015 roku na wykresie, który stworzony został w oparciu o teksty czasopism (Z(DE)). Wzrost ten można wytłumaczyć tym, że jest to okres bezpośrednio

po premierze kolejnej części wielkiej growej serii – Assassin Creed, która została zatytułowana Unity. Ukazała się ona 11 listopada 2014 roku. W momencie premiery była jednak tak niedopracowana, że praktycznie nie dało się w nią grać. Jej błędy były wychwytywane przez graczy i wyśmiewane. Niektórzy nawet nazywali ją najbardziej zabugowaną (zwierającą błędy) grą⁸⁶. Przykład jednego z błędów pokazałem na poniższym zdjęciu.



Zdjęcie 36. Postaci w grze Assassin Creed: Unity, która w wyniku błędów straciła skórę twarzy

Źródło: www.arstechnica.com/gaming/2014/11/ubisoft-working-to-fix-game-breaking-assassins-creed-unity-bugs/

Autorzy gry, aby ratować sytuację, zaczęli naprawiać błędy i publikować co kilka dni kolejne *patche*. Podobnie jak w przypadku poprzednim (pole *demo*) także i tutaj w przypadku newsów informacje o *patchach* rozłożone zostały w czasie, natomiast czasopisma skumulowały je, czego efektem jest gwałtowny wzrost użycia wyrazów w jednym kwartale.

Na zakończenie warto zwrócić uwagę na fakt, który znalazł wyraźne odzwierciedlenie w wykresach (zamieszczonych w aneksie), że często wzrost użycia wyrazów z różnych pól leksykalnych jest odbiciem konkretnych i możliwych do odtworzenia sytuacji pozajęzykowych. W przypadku środowiska graczy wideo największy wpływ na wzrost użycia wyrazów pewnych pól ma premiera nowego sprzętu do gier, najczęściej konsol. Widać, że z ich premierą nie łączy się jedynie wzrost użycia wyrazów z pola *konsola*, ale także np.

⁸⁶ Pisano tak o niej na stronie <https://naekranie.pl/aktualnosci/assassin-s-creed-unity-najbardziej-zabugowana-gra-zobacz-wideo> [dostęp: 28.07.2018].

rozdzielczość, next-gen, pad itd., czyli tych pól, które są w jakiś sposób związane z nowym sprzętem.

Reasumując, materiał ilościowy można analizować pod różnym kątem i z różnym stopniem uszczegółowienia. Ogólne analizy dotyczą łącznej sumy wystąpień wyrazów graficznych w całym analizowanym okresie. Dokładniejsze wyniki daje rozłożenie tych wystąpień na kwartały. Poza tym można badać materiał ze względu na typ tekstu (czasopismo, news, fora) lub typ użytkownika (*komputerowiec, konsolowiec*).

Powyższe badania potwierdzają wyraźną zbieżność między częstotliwością określonej leksyki a wydarzeniami z nią związanymi.

Dotychczasowy opis różnych ujęć związanych z analizą częstotliwości występowania wyrazów w badanych polach znaczeniowych miał charakter osobnych ustaleń. Pojawia się w tym momencie potrzeba syntezy różnych aspektów w celu wykrycia ich wzajemnej zależności. Odpowiedzią na to zadanie badawcze będzie system 2.0, który moim zdaniem pozwoli dobrze uporządkować zebrany materiał i wyciągnąć interesujące wnioski na temat polskiego socjolektu graczy wideo.

18. SYSTEM 2.0

18.1. SYSTEM 2.0 W ODNIESIENIU DO SUMY WYRAZÓW GRAFICZNYCH POSZCZEGÓLNYCH PÓL W CAŁYM KORPUSIE – PLATFORMA ZIELONA I CZERWONA

W jednej z poprzednich części pracy, na stronie 165, przedstawiłem zasady działania systemu 1.0, który okazał się pomocny przy dzieleniu pól leksykalnych, na te, które bliższe są *konsolowcom* oraz te, które ważniejsze są dla *komputerowców*. Przy jego pomocy udało mi się ustalić także listę pól wspólnych obu grupom. Następnie stworzyłem listy pól posiadających najwięcej i najmniej wyrazów graficznych, a później, podczas krótkiego omówienia złożoności badanych pól, wspomniałem, że będzie ona jednym z elementów systemu 2.0. Po zbadaniu częstotliwości występowania wyrazów i wyliczeniu współczynników zmienności dla danych pól oraz tendencji z nimi związanych, dysponuję już wszystkimi pięcioma elementami (trzema głównymi i dwoma dodatkowymi) potrzebnymi, aby móc wprowadzić system 2.0. Zanim jednak go omówię, to najpierw zaprezentuję wyniki, ponieważ łatwiej będzie opisać jego działanie dysponując już tymi danymi. Początkowo zaprezentuję jedynie wyniki w odniesieniu do zielonej platformy:

TABELA 41.
SYSTEM 2.0 - PLATFORMA ZIELONA

A	B	C	D	E	F	G	H	I	J
PC	56013	20	4,499		↑	4	4	4	12
PLAYSTATION	50603	16	18,897		↑	4	4	4	12
RPG	14304	10	4,135		↓	4	4	4	12
KONSOLA	46065	6	4,968		↓	3	4	4	11
XBOX	32729	8	19,021		↓	3	4	4	11
KOMPUTER	31500	9	2,823		↑	3	4	4	11
MULTI	11049	6	7,821		↓	3	4	4	11
PATCH	5402	20	12,113		↑	4	3	4	11
SINGIEL	4662	10	8,331		↓	4	3	4	11
GAMEPLAY	10058	3	11,291		↑	2	4	4	10
TRAILER	8426	3	11		↓	2	4	4	10
STRZELANINA	5898	6	9,813		↓	3	3	4	10
QUEST	3832	8	5,514		↓	3	3	4	10
TEKSTURA	2185	10	9,559		↑	4	2	4	10
SAVE	2177	11	8,084		↓	4	2	4	10
HARDKOR	2046	15	11,024		↓	4	2	4	10
BAN	1665	11	6,826		↑	4	2	4	10
HEJT	1338	16	18,209		↑	4	2	4	10
CASUAL	1312	24	15,115		↓	4	2	4	10
SANDBOX	1306	11	17,376		↑	4	2	4	10
CRAP	1195	15	7,483		↑	4	2	4	10
CALAK	853	17	18,129		↑	4	2	4	10
GRIND	828	12	17,601		↑	4	2	4	10
SPAWN	817	17	11,376		↑	4	2	4	10
DLC	10550	1	9,517		↓	1	4	4	9
TABLET	4630	5	12,568		↓	2	3	4	9
KLAWIATURA	3834	5	7,176		↓	2	3	4	9
PEŁNIAK	3288	3	5,203		↑	2	3	4	9
BIJATYKA	3256	4	8,863		↓	2	3	4	9
GRYWALNY	2910	7	6,162		↑	3	2	4	9
EXCLUSIVE	2842	7	12,107		↑	3	2	4	9
STREAM	1783	13	32,137		↑	4	2	3	9
HANDHELD	1582	6	16,222		↓	3	2	4	9
PLATYNA	1007	7	7,966		↑	3	2	4	9
GAMECUBE	938	8	19,405		↓	3	2	4	9
CRAFT	638	10	23,577		↑	4	2	3	9
PAUZA	519	6	14,528		↑	3	2	4	9
FARMIC	462	7	10,508		↑	3	2	4	9
TANK	460	7	15,532		↑	3	2	4	9
CHEAT	370	7	18,378		↓	3	2	4	9
PSP	7593	4	33,86		↓	2	3	3	8
PAD	7410	1	6,649		↑	1	3	4	8
DEMO	6730	2	11,264		↓	1	3	4	8
ROZDZIELCZOŚĆ	4300	2	10,288		↑	1	3	4	8
SMARTFON	3835	2	11,417		↑	1	3	4	8
SEQUEL	2920	3	14,021		↓	2	2	4	8
REMAKE	2881	4	9,583		↑	2	2	4	8

JRPG	2334	4	10,383	↑	2	2	4	8
GAMING	2234	4	12,064	↑	2	2	4	8
INDIE	1911	6	35,675	↑	3	2	3	8
MULTIMEDIA	1613	4	13,156	↓	2	2	4	8
PRZEGLĄDARKA	1583	3	13,257	↓	2	2	4	8
SKRYPT	1493	5	10,535	↓	2	2	4	8
REMASTER	1273	9	37,972	↑	3	2	3	8
HARDWARE	1025	3	14,631	↓	2	2	4	8
LINIOWY	930	3	12,386	↓	2	2	4	8
DUNGEON	884	3	6,663	↑	2	2	4	8
LOADING	830	3	13,637	↓	2	2	4	8
ŚCIGAŁKA	587	3	12,526	↓	2	2	4	8
D-PAD	558	3	17,07	↓	2	2	4	8
HP	540	3	8,614	↑	2	2	4	8
MARIAN	484	3	19,658	↓	2	2	4	8
SOFTWARE	413	3	19,268	↓	2	2	4	8
MULTIPLATFORMOWY	339	5	18,38	↓	2	2	4	8
PEESIKS	329	3	18,929	↑	2	2	4	8
FEATURE	313	3	14,153	↑	2	2	4	8
KONTROLER	2679	1	18,833	↓	1	2	4	7
PRZYGODÓWKA	2453	2	8,359	↑	1	2	4	7
SOUNDTRACK	1910	1	11,686	↑	1	2	4	7
PODSTAWKA	1458	1	10,302	↑	1	2	4	7
NPC	1428	1	6,585	↓	1	2	4	7
ROUTER	1300	1	9,16	↑	1	2	4	7
LOKALIZACJA	1220	1	16,942	↓	1	2	4	7
GOTY	1170	1	11,19	↑	1	2	4	7
KONWERSJA	1155	1	9,843	↓	1	2	4	7
MMORPG	1121	4	20,441	↓	2	2	3	7
PVP	1065	1	12,744	↑	1	2	4	7
EMULATOR	1009	3	25,085	↑	2	2	3	7
SKIN	848	2	18,807	↑	1	2	4	7
CUTSCENA	798	4	22,166	↑	2	2	3	7
WIESIEK	761	1	15,789	↓	1	2	4	7
PRE-ORDER	716	4	21,018	↑	2	2	3	7
FREE-TO-PLAY	597	3	39,658	↑	2	2	3	7
DREAMCAST	592	5	24,431	↑	2	2	3	7
QTE	556	1	17,143	↓	1	2	4	7
FRAMERATE	503	1	14,286	↑	1	2	4	7
PRZEJŚCIÓWKA	458	1	11,345	↑	1	2	4	7
CHECKPOINT	396	2	14,646	↓	1	2	4	7
ŻYWOTNOŚĆ	382	1	14,668	↓	1	2	4	7
STACJONARKA	350	1	12,353	↑	1	2	4	7
ACTION RPG	343	4	24,719	↓	2	2	3	7
SAMOCHODÓWKA	305	1	18,519	↓	1	2	4	7
TURÓWKA	249	1	16,863	↑	1	2	4	7
EGRANIZACJA	188	5	30,526	↓	2	2	3	7
SAMOGRAJ	167	4	21,788	↑	2	2	3	7
ZAPYCHACZ	151	2	12,911	↑	1	2	4	7
REDZI	139	3	29,329	↑	2	2	3	7

WSAD	136	3	32,5	↑	2	2	3	7
HIGH-END	126	3	36,875	↑	2	2	3	7
KLATKOWAĆ	116	5	32,8	↑	2	2	3	7
KINECT	3334	2	53,954	↓	1	3	2	6
NEXT-GEN	1718	5	56,177	↑	2	2	2	6
MATRYCA	1126	1	29,075	↑	1	2	3	6
UŻYWKA	762	2	22,92	↓	1	2	3	6
NOTEBOOK	687	1	30,315	↓	1	2	3	6
SUBSKRYPCJA	629	1	26,908	↑	1	2	3	6
MIKROTRANSAKCJA	568	2	30,87	↑	1	2	3	6
ZNAJDŹKA	477	1	32,134	↑	1	2	3	6
REEDYCJA	439	1	22,242	↑	1	2	3	6
DEATHMATCH	364	2	35,028	↓	1	2	3	6
HEADSET	347	2	30,304	↑	1	2	3	6
CROSSOVER	293	2	23,421	↑	1	2	3	6
SPRITE	286	2	26,62	↑	1	2	3	6
TRIGGER	259	1	22,593	↓	1	2	3	6
PVE	235	1	20,165	↑	1	2	3	6
SHADER	186	1	21,209	↓	1	2	3	6
TOUCHPAD	172	4	41,637	↑	2	2	2	6
CURRENT-GEN	168	2	38,012	↓	1	2	3	6
STYLUS	147	1	33,475	↓	1	2	3	6
HEALER	147	2	22,581	↓	1	2	3	6
GEEK	136	4	50	↑	2	2	2	6
CZARNULA	126	2	21,935	↓	1	2	3	6
GRZYBEK	117	1	38,333	↑	1	2	3	6
MINIMAPA	113	2	29,63	↓	1	2	3	6
PRZESZKADZAJKA	112	1	36,733	↓	1	2	3	6
CREDITSY	108	1	20,385	↓	1	2	3	6
CEL-SHADING	104	3	45	↑	2	2	2	6
KARTRIDŹ	439	1	44,477	↑	1	2	2	5
SIECIÓWKA	303	2	40,266	↓	1	2	2	5
ROGUELIKE	194	4	66,193	↑	2	2	1	5
WTYCZKA	169	1	54,294	↑	1	2	2	5
SHMUP	156	2	55,033	↑	1	2	2	5
COOLDOWN	110	1	42,414	↑	1	2	2	5
UNBOXING	107	1	52,673	↑	1	2	2	5
PEESTRÓJKA	83	6			3	1	1	5
OFF-ROAD	50	6			3	1	1	5
CROSS-BUY	286	1	85,517	↑	1	2	1	4
PADLET	199	2	85	↓	1	2	1	4
MULTITOUCH	186	1	62,961	↑	1	2	1	4
PSP2	139	2	150	↓	1	2	1	4
CELOWNICZEK	124	2	63,894	↓	1	2	1	4
CROSS-PLATFORM	58	4			2	1	1	4
LIVE-STREAM	37	3			2	1	1	4
CROSS-GEN	20	3			2	1	1	4
RANDOMIZACJA	17	3			2	1	1	4
PEESZCZÓRKA	17	4			2	1	1	4
CROSS-CONTROLLER	15	3			2	1	1	4

PEESDWÓJKA	14	3		2	1	1	4
PEESPLUSOWY	9	3		2	1	1	4
PSP GO	95	2		1	1	1	3
PRE-ALPHA	92	2		1	1	1	3
BUDŻETÓWKA	86	2		1	1	1	3
MUST-HAVE	85	1		1	1	1	3
PPM	84	1		1	1	1	3
OŚMIOBITOWY	77	2		1	1	1	3
RROD	70	1		1	1	1	3
SIDE-QUEST	62	1		1	1	1	3
QWERTY	57	1		1	1	1	3
REAL-TIME	57	2		1	1	1	3
CROSS-PLAY	56	1		1	1	1	3
MULTITASKING	54	1		1	1	1	3
BUMPER	52	1		1	1	1	3
NIEDZIELNIAK	52	1		1	1	1	3
ONE-LINER	50	1		1	1	1	3
PAY-TO-WIN	47	1		1	1	1	3
YLOD	43	1		1	1	1	3
DEVKIT	41	1		1	1	1	3
TEKSTÓWKA	40	2		1	1	1	3
DOKANAŁÓWKI	38	1		1	1	1	3
LPM	38	1		1	1	1	3
PAST-GEN	33	2		1	1	1	3
SUB-QUEST	30	1		1	1	1	3
CROSS-SAVE	26	1		1	1	1	3
METAGRA	26	1		1	1	1	3
FLASHÓWKA	25	1		1	1	1	3
SUB-BOSS	22	1		1	1	1	3
JEDNOSTRZAŁOWIEC	22	1		1	1	1	3
MULTITAP	20	1		1	1	1	3
SZESNASTOBITOWY	13	2		1	1	1	3
ODBLOKOWAJKA	5	1		1	1	1	3
CROSS-CHAT	3	1		1	1	1	3
ZBLUROWANY	3	1		1	1	1	3
MULTIKART	2	1		1	1	1	3
PRZETRWANIÓWKA	2	1		1	1	1	3
EKSBOOKSŁAN	1	1		1	1	1	3
PRZEDŁUŻAJKA	1	1		1	1	1	3
STODWUDZIESTOÓSMIOBITOWIEC	1	1		1	1	1	3
THREE-SIXTY	1	1		1	1	1	3
TWARZOWANIE	1	1		1	1	1	3

Jak już wspominałem, system 2.0 składa się z 5 elementów:

- 1) sumy wyrazów graficznych w danym polu (kolumna B),
- 2) poziomu złożoności danego pola (kolumna C),
- 3) współczynnika zmienności wykresów odnoszących się do danego pola (kolumna D),
oraz dodatkowo:
- 4) wyników zastosowania systemu 1.0 (kolumna E),
- 5) ogólnej tendencji wyrazów z danego pola (kolumna F).

Pierwsze 3 elementy są elementami głównymi, natomiast wyniki systemu 1.0 i tendencja występowania wyrazów to elementy dodatkowe. Zostały one zaklasyfikowane jako takie elementy, ponieważ ich wyniki nie stanowią konkretnych liczb, jak ma to miejsce w przypadku trzech elementów głównych. Są one tylko wskazaniem, do której grupy przynależy dane pole znaczeniowe oraz jaką ma ono tendencję. Mimo że są to elementy dodatkowe, to dużo mówią na temat badanego materiału, dlatego nie zrezygnowałem z ich umieszczenia w powyższej tabeli.

System powstał w celu uporządkowania danych, które posłużą do wyciągnięcia kolejnych wniosków na temat socjolektu graczy wideo. Aby łatwiej było pracować z tym systemem oraz umieszczonymi w nim danymi, wprowadziłem klasyfikację punktową, która obejmowała informacje umieszczone w trzech głównych kolumnach (B, C, D).

Kolumna B zawiera dane na temat sumy wyrazów danego pola znaczeniowego.

W wyniku obserwacji i analizy danych pod kątem sumy ich wyrazów graficznych, wydzieliłem cztery kategorie punktowe. Pola które zawierają:

- ponad 8001 wyrazów otrzymały 4 punkty
- od 3001 do 8000 wyrazów otrzymały 3 punkty
- od 101 do 3000 wyrazów otrzymały 2 punkty
- i poniżej 100 otrzymały 1 punkt.

Wyniki punktowe zaprezentowałem w kolumnie H.

Analizując poziom złożoności danego pola (kolumna C), zaobserwowałem, że wiele z nich posiada w sobie tylko jedną formę podstawową, którą jest najczęściej rzeczownik. Wśród pól zawierających 2 formy podstawowe dominują te z rzeczownikiem i przymiotnikiem od niego utworzonym, ewentualnie dwa warianty rzeczownikowych form podstawowych, które różnią się zapisem. Pola z większą liczbą form podstawowych zawierały przysłówki oraz czasowniki, a te jeszcze większe - różne imiesłowy, zaś najbardziej rozbudowane pola to bogactwo kreatywnych modyfikacji różnych wyrazów. Na podstawie tego zdecydowałem, że punktację rozdzielię następująco:

10 i więcej form podstawowych – 4 punkty

od 6 do 9 – 3 punkty

od 3 do 5 – 2 punkty

i od 1 do 2 form – 1 punkt.

Wyniki punktowe zaprezentowałem w kolumnie G.

Natomiast z klasyfikacją wyników współczynników przyszła mi z pomocą jedna ze stron internetowych, gdzie współczynniki zmienności interpretowano następująco:

„0 - 20% - Małe zróżnicowanie.

20 - 40% - Średnie zróżnicowanie.

40 - 60% - Duże zróżnicowanie.

Powyżej 60% - Bardzo duże zróżnicowanie.” [Statystyka.cba, dostęp: 20.06.2018]

Polom, które miały małe zróżnicowane przyznałem 4 punkty, średnie 3 punkty, duże 2 punkty i bardzo duże 1 punkt.

Wyniki zaprezentowałem w kolumnie I.

Z tym jednak zastrzeżeniem, że pola, które liczyły łącznie mniej niż 100 wyrazów tekstowych, automatycznie dostawały 1 punkt, ponieważ jest to zbyt mała liczba do badań nad częstotliwością wystąpień wyrazów na przestrzeni poszczególnych kwartałów.

Wyniki systemu 1.0 przedstawiłem w tabeli (kolumna E) z wykorzystaniem kolorów. Tych samych, których użyłem w grafice na stronie 178. Użyłem żółtego na oznaczenie pola związanego z grupą *komputerowców*, niebieskiego na oznaczenie pola, które związane jest z *konsolowcami* oraz zielonego dla pól, które w równym stopniu używane są przez oba środowiska graczy wideo.

Podając tendencję danego pola (kolumna F) zastosowałem prezentację graficzną – strzałki. Strzałka w górę oznacza tendencję rosnącą, natomiast strzałka w dół malejącą.

W tabeli, która jest wynikiem zastosowania systemu 2.0, pola zostały ułożone od tych, które uzyskały największą sumę punktów (pokazanych w kolumnie J), do tych, które uzyskały ich najmniej. Przejdźmy zatem do analizy wyników.

Jak widać po wynikach, najważniejszymi polami dla graczy wideo są *PC*, *PlayStation* oraz *RPG*. Wszystkie one uzyskały maksymalną możliwą liczbę punktów, czyli 12. Wszystkie one także zaklasyfikowane zostały jako przynależące do pola zielonego, czyli oznacza to, że są używane zarówno przez *konsolowców*, jak i *komputerowców*. Dwa z nich (*PlayStation* i *PC*) mają tendencję wzrostową, natomiast jedno (*RPG*) malejącą. Z wyników badań wywnioskować należy, że w Polsce graczami wideo są najczęściej osoby, które grają na *pecetach* i/lub konsolach *PlayStation*. Wyniki te odzwierciedlenie znajdują w rzeczywistości, ponieważ

w Polsce najczęściej jest *pecetowców* oraz *playstationowców*. *Pecetowców*, ponieważ w przypadku gier wydawanych na *PC*, niestety popularny jest proceder nielegalnego ich ściągania. Natomiast *playstationowców*, ponieważ pierwsza z konsol z serii *PlayStation* pojawiła się w Polsce już w latach 90, i w tamtym czasie była to jedyna nowoczesna platforma do gier, poza oczywiście *pecetem*. W tamtym okresie wychowywali się gracze, którzy obecnie mają 25–30 lat. Przyzwyczajeni za młodu do konsol *PlayStation*, obecnie także są do nich przywiązani i kupują jej następne wersje, aby pograć w ulubione gry, których kontynuację, w niektórych przypadkach, wydawane są do dzisiaj, w wersjach *ekskluzywnych* dla *PlayStation*. Wybierając grę, polscy gracze najczęściej sięgają po tytuły RPG.

Patrząc na grupę pól, które uplasowały się na drugim miejscu w zestawieniu, są to: *konsola*, *xbox*, *komputer*, *multi*, *patch* i *singiel*. Podobnie, jak w przypadku pól z pierwszego miejsca w tabeli, także tutaj są to pola z grupy zielonej – ogólnej. Większość pól ma tendencję spadkową, a tylko *patch* i *komputer* rosną.

Wysokie miejsce pola *patch* w powyższym zestawieniu świadczy o tym, że jest ono ważne w socjolekcie graczy wideo, natomiast tendencja rosnąca świadczy o tym, że jeszcze bardziej zyskuje ono na znaczeniu. Wynika to z faktu, że coraz częściej gry są niedopracowane i wydawane z błędami. Twórcy mogą sobie pozwolić na wydawanie takich gier, dzięki czemu oszczędzają na czasochłonnych testach, ponieważ teraz prawie każdy gracz ma wszystkie sprzęty do gier połączone z internetem, a dzięki temu twórcy mogą poprawić każdy błąd w grze, udostępniając drogą internetową *patch*. Dawniej, np. w latach 90-tych, twórcy nie mogli sobie na to pozwolić, ponieważ internet nie był tak rozpowszechniony, więc gra musiała być od początku dopracowana, bo nie było łatwej możliwości naprawienia jej.

Poza *patchem* na drugim miejscu w zestawieniu znalazły się także pola odnoszące się do trybu rozgrywki – *multi* i *singiel* oraz pole nazywające serie konsol Microsoft – *xbox*. *Xbox* jest drugą najpopularniejszą konsolą w Polsce.

Patrząc całościowo na dwie pierwsze grupy pól, stwierdzić trzeba, że dominują pola, które zaliczone zostały, już na początku pracy, do kręgu *sprzęt gracza* (patrz strona 101). W grupie dziewięciu pól na samej górze zestawienia aż pięć z nich pochodzi właśnie z tego kręgu. Widać więc, że to nie pola opisujące elementy gry, jakby można było przypuszczać, stanowią trzon najważniejszych pól, ale właśnie te związane ze sprzętem.

Zwykło się także twierdzić, że jeżeli danych wyrazów jest dużo w tekście, to pole je zawierające jest bardzo ważne. Badania pokazały, że nie zawsze tak jest. Patrząc na pierwszą dziesiątkę w zestawieniu, widać, że np. pole *RPG*, które zawiera znacznie mniej wyrazów niż *konsola* znalazło się wyżej od niej. Dlatego nie tylko suma liczby wyrazów w tekście powinna

stanowić o ich ważności, ale także inne czynniki np. częstotliwość wystąpień. Ciekawym spostrzeżeniem, związanym z częstotliwością, jest to, że na pierwszej stronie zestawienia znalazły się tylko 3 pola, które nie otrzymały 4 punktów w systemie 2.0. Resztę stanowią pola z niskim współczynnikiem zmienności.

W przypadku pól przynależących do środowiska *komputerowców*, najważniejsze okazało się pole *pełniak*⁸⁷. Wynika to zapewne z tego, że czytelnicy czasopisma CD-Action są bardzo ciekawi gier w pełnych wersjach, które dodane zostaną do kolejnych numerów pisma. Często o tym dyskutują, np. na forach internetowych. Natomiast najwyżej umieszczonym w zestawieniu polem ze środowiska *konsolowców* jest *Marian* ‘nazwa bohatera gier Super Mario’ i *Peesiks* ‘środowiskowa nazwa konsoli PlayStation’. *Marian* jest swoista maskotką gier konsolowych, natomiast *Peesiks* to konsola bardzo popularna w latach 90-tych. Wielu graczy wciąż lubi na niej pograć.

Na samym dole wykresu znalazły się pola, których wyrazy były bardzo rzadko używane. Są to fragmenty większych pól znaczeniowych – *Threesixty* (fragment pola – *Xbox 360* ‘nazwa konsoli firmy Microsoft’), albo pola zawierające wyrazy bardzo specjalistyczne, jak np. *metagra* ‘gra ponad grą’, *devkit* ‘wersja sprzętu do gier udostępniana deweloperom’ czy *multitasking* ‘wielozadaniowość’. Znajdują się tam także pola z wyrazami już nieco zapomnianymi, np. *multitap* ‘urządzenie pozwalające podpiąć do konsoli dodatkowe kontrolery’. Wyrazy z tego pola były często używane w czasach, gdy nie było kontrolerów bezprzewodowych, a konsole posiadały najczęściej tylko dwa gniazda na kontrolery, wtedy trzeba było skorzystać z tego swoistego rozgałęźnika. Poza tymi grupami pól są także te, które zawierają wyrazy kreatywne lub okazjonalne – *odblokowajka* ‘element do odblokowania w grze’ czy *twarzowanie* ‘udostępnianie wyglądu twarzy w celu umieszczenia jej w grze, czyli stworzenia postaci zawierającej daną twarz’.

Na zakończenie warto przyjrzeć się jeszcze wynikom uzyskanym przez zastosowanie systemu 2.0, ale w odniesieniu do platformy kontrolnej – czerwonej:

⁸⁷ *Komputerowcy* pełne wersje gier zwykli nazywać *pełniakami*. Te wersje gry stanowią przeciwieństwo *dem*, gdzie graczom udostępniony zostaje tylko fragment rozgrywki.

TABELA 42.
SYSTEM 2.0 - CZERWONA PLATFORMA

A	B	C	D	E	F	G	H	I	J
PC	168027	20	11,882	↑	↑	4	4	4	12
PLAYSTATION	167315	16	14,705	↑	↑	4	4	4	12
RPG	30714	10	7,946	↑	↑	4	4	4	12
KONSOLA	114966	6	6,971	↑	↑	3	4	4	11
XBOX	100159	8	16,029	↓	↓	3	4	4	11
KOMPUTER	54142	9	6,49	↑	↑	3	4	4	11
MULTI	25255	6	14,391	↓	↓	3	4	4	11
STRZELANINA	18867	6	12,784	↓	↓	3	4	4	11
SINGIEL	9524	10	15,558	↓	↓	4	3	4	11
HARDKOR	6614	15	10,121	↓	↓	4	3	4	11
PATCH	11594	20	24,809	↑	↑	4	3	3	10
GRYWALNY	7959	7	11,556	↓	↓	3	3	4	10
QUEST	7180	8	12,509	↓	↓	3	3	4	10
TEKSTURA	5703	10	15,149	↑	↑	4	2	4	10
SAVE	3970	11	18,214	↓	↓	4	2	4	10
BAN	2280	11	17,155	↑	↑	4	2	4	10
CRAP	1941	15	15,356	↑	↑	4	2	4	10
SPAWN	1877	17	17,4	↓	↓	4	2	4	10
GAMEPLAY	27114	3	25,382	↑	↑	2	4	3	9
TRAILER	19952	3	22,253	↓	↓	2	4	3	9
PAD	16480	1	12,109	↓	↓	1	4	4	9
BIJATYKA	9431	4	12,98	↓	↓	2	3	4	9
KLAWIATURA	7821	5	16,907	↓	↓	2	3	4	9
EXCLUSIVE	7280	7	27,329	↑	↑	3	3	3	9
REMAKE	6816	4	19,329	↑	↑	2	3	4	9
SANDBOX	3819	11	26,196	↑	↑	4	2	3	9
CASUAL	2822	24	23,972	↓	↓	4	2	3	9
HEJT	2245	16	33,899	↑	↑	4	2	3	9
CALAK	1938	17	34,433	↑	↑	4	2	3	9
CRAFT	1750	10	33,533	↑	↑	4	2	3	9
GRIND	1581	12	37,962	↑	↑	4	2	3	9
DLC	27852	1	21,034	↓	↓	1	4	3	8
PSP	23459	4	42,829	↓	↓	2	4	2	8
ROZDZIELCZOŚĆ	11935	2	13,403	↑	↑	1	3	4	8
TABLET	11580	5	20,581	↓	↓	2	3	3	8
SMARTFON	11561	2	16,688	↑	↑	1	3	4	8
SEQUEL	9183	3	22,431	↑	↑	2	3	3	8
PRZYGODÓWKA	6760	2	14,269	↑	↑	1	3	4	8
GAMING	5698	4	19,081	↑	↑	2	2	4	8
HANDHELD	5072	6	28,862	↓	↓	3	2	3	8
STREAM	3978	13	50,783	↑	↑	4	2	2	8
PEŁNIAK	3845	3	9,761	↑	↑	2	2	4	8
LINIOWY	2722	3	18,718	↓	↓	2	2	4	8
GAMAECUBE	2330	8	39,538	↓	↓	3	2	3	8
DUNGEON	1817	3	14,673	↑	↑	2	2	4	8
PLATYNA	1444	7	23,472	↑	↑	3	2	3	8
PAUZA	1255	6	25,055	↑	↑	3	2	3	8

TANK	1051	7	30,057	↑	3	2	3	8
HP	921	3	20	↑	2	2	4	8
CHEAT	704	7	34,83	↓	3	2	3	8
FARMIC	664	7	27,626	↑	3	2	3	8
DEMO	14199	2	23,88	↓	1	3	3	7
KONTROLER	9401	1	20,044	↓	1	3	3	7
INDIE	5937	6	56,163	↑	3	2	2	7
JRPG	4563	4	24,248	↑	2	2	3	7
PRZEGLĄDARKA	3806	3	23,136	↓	2	2	3	7
MULTIMEDIA	3659	4	23,083	↓	2	2	3	7
KONWERSJA	3631	1	15,837	↑	1	2	4	7
SKRYPT	3413	5	22,646	↓	2	2	3	7
MMORPG	3325	4	29,739	↓	2	2	3	7
NPC	3065	1	15,234	↓	1	2	4	7
PODSTAWKA	2499	1	17,799	↑	1	2	4	7
CUTSCENA	2365	4	30,714	↑	2	2	3	7
HARDWARE	2145	3	35,948	↓	2	2	3	7
LOADING	2094	3	26,805	↓	2	2	3	7
PRE-ORDER	1661	4	33,947	↑	2	2	3	7
DREAMCAST	1583	5	37,215	↑	2	2	3	7
D-PAD	1486	3	30,312	↓	2	2	3	7
ŚCIGAŁKA	1425	3	20,226	↓	2	2	3	7
MARIAN	1338	3	33,162	↑	2	2	3	7
SOFTWARE	1100	3	35,46	↓	2	2	3	7
ACTION RPG	1068	4	34,894	↓	2	2	3	7
MULTIPLATFORMOWY	927	5	28,943	↓	2	2	3	7
PEESIKS	828	3	35,928	↑	2	2	3	7
EGRANIZACJA	528	5	35,849	↓	2	2	3	7
FEATURE	525	3	28,381	↑	2	2	3	7
PEESTRÓJKA	213	6	43,962	↓	3	2	2	7
SOUNDTRACK	3741	1	21,909	↑	1	2	3	6
REMASTER	3301	9	66,042	↑	3	2	1	6
LOKALIZACJA	3285	1	23,71	↓	1	2	3	6
PVP	2510	1	23,437	↑	1	2	3	6
FREE-TO-PLAY	2300	3	47,442	↑	2	2	2	6
ROUTER	2191	1	22,364	↓	1	2	3	6
SKIN	1825	2	32,397	↑	1	2	3	6
QTE	1775	1	22,884	↓	1	2	3	6
EMULATOR	1721	3	48,921	↑	2	2	2	6
GOTY	1602	1	22,687	↑	1	2	3	6
REEDYCJA	1217	1	31,373	↑	1	2	3	6
CHECKPOINT	1106	2	25,298	↓	1	2	3	6
FRAMERATE	1105	1	38,598	↑	1	2	3	6
WIESIEK	1088	1	26,283	↓	1	2	3	6
PRZEJŚCIÓWKA	1009	1	23,34	↓	1	2	3	6
ŻYWOTNOŚĆ	965	1	23,542	↓	1	2	3	6
TRIGGER	822	1	32,377	↓	1	2	3	6
SAMOCHODÓWKA	729	1	29,265	↓	1	2	3	6
SIECIÓWKA	645	2	34,992	↓	1	2	3	6
TOUCHPAD	606	4	49,669	↑	2	2	2	6

PVE	591	1	29,699	↑	1	2	3	6
STACJONARKA	578	1	27,183	↑	1	2	3	6
TURÓWKA	453	1	34,619	↑	1	2	3	6
WSAD	405	3	47,046	↑	2	2	2	6
HIGH-END	391	3	47,252	↑	2	2	2	6
CEL-SHADING	389	3	53,676	↑	2	2	2	6
SAMOGRAJ	356	4	46,196	↑	2	2	2	6
ZAPYCHACZ	273	2	30,714	↑	1	2	3	6
CREDITSY	240	1	36,271	↓	1	2	3	6
REDZI	236	3	54,146	↑	2	2	2	6
KINECT	11347	2	60,444	↓	1	3	1	5
NEXT-GEN	5143	5	82,114	↑	2	2	1	5
MATRYCA	3598	1	40,3	↑	1	2	2	5
NOTEBOOK	2235	1	41,425	↓	1	2	2	5
MIKROTRANSAKCJA	1779	2	45,308	↑	1	2	2	5
KARTRIDŹ	1697	1	51,306	↑	1	2	2	5
SUBSKRYPCJA	1598	1	44,528	↑	1	2	2	5
ZNAJDŹKA	1572	1	42,606	↑	1	2	2	5
DEATHMATCH	1253	2	41,11	↓	1	2	2	5
UŻYWKWA	1252	2	44,535	↓	1	2	2	5
HEADSET	1082	2	54,286	↑	1	2	2	5
SPRITE	730	2	41,648	↑	1	2	2	5
CROSSOVER	677	2	40,349	↓	1	2	2	5
ROGUELIKE	625	4	79,236	↑	2	2	1	5
STYLUS	498	1	48,78	↓	1	2	2	5
GEEK	446	4	62,697	↑	2	2	1	5
GRZYBEK	434	1	43,638	↑	1	2	2	5
PRZESZKADZAJKA	374	1	44,518	↓	1	2	2	5
SHADER	358	1	40,169	↓	1	2	2	5
COOLDOWN	346	1	56,364	↑	1	2	2	5
MINIMAPA	329	2	41,481	↓	1	2	2	5
HEALER	323	2	44,411	↓	1	2	2	5
CZARNULA	290	2	47,778	↓	1	2	2	5
LIVE-STREAM	249	3	112,47	↑	2	2	1	5
CROSS-PLAY	218	1	57,556	↑	1	2	2	5
CROSS-PLATFORM	213	4	119,849	↓	2	2	1	5
PRE-ALPHA	206	2	44,072	↑	1	2	2	5
OFF-ROAD	160	6			3	1	1	5
CROSS-BUY	1030	1	83,385	↑	1	2	1	4
MULTITOUCH	779	1	64,352	↑	1	2	1	4
PSP2	773	2	171,731	↓	1	2	1	4
PADLET	715	2	99,497	↓	1	2	1	4
CELOWNICZEK	522	2	64,286	↓	1	2	1	4
CURRENT-GEN	504	2	68,833	↓	1	2	1	4
WTYCZKA	492	1	79,63	↑	1	2	1	4
SHMUP	446	2	78,149	↑	1	2	1	4
PSP GO	381	2	108,947	↓	1	2	1	4
OŚMIOBITOWY	291	2	70	↑	1	2	1	4
UNBOXING	264	1	85,039	↑	1	2	1	4
BUMPER	208	1	76,863	↓	1	2	1	4

KLATKOWAĆ	198	5		2	1	1	4
CROSS-CONTROLLER	65	3		2	1	1	4
PEESZCZÓRKA	57	4		2	1	1	4
CROSS-GEN	41	3		2	1	1	4
RANDOMIZACJA	27	3		2	1	1	4
PEESPLUSOWY	26	3		2	1	1	4
PEESDWÓJKA	24	3		2	1	1	4
SIDE-QUEST	185	1		1	1	1	3
MUST-HAVE	170	1		1	1	1	3
PAY-TO-WIN	170	1		1	1	1	3
MULTITASKING	168	1		1	1	1	3
QWERTY	142	1		1	1	1	3
NIEDZIELNIAK	141	1		1	1	1	3
RROD	137	1		1	1	1	3
PPM	133	1		1	1	1	3
BUDŻETÓWKA	132	2		1	1	1	3
CROSS-SAVE	124	1		1	1	1	3
METAGRA	110	1		1	1	1	3
TEKSTÓWKA	108	2		1	1	1	3
REAL-TIME	104	2		1	1	1	3
DEVKIT	102	1		1	1	1	3
YLOD	99	1		1	1	1	3
ONE-LINER	92	1		1	1	1	3
PAST-GEN	91	2		1	1	1	3
DOKANAŁÓWKI	74	1		1	1	1	3
SUB-BOSS	70	1		1	1	1	3
FLASHÓWKA	64	1		1	1	1	3
MULTITAP	54	1		1	1	1	3
SUB-QUEST	50	1		1	1	1	3
LPM	49	1		1	1	1	3
JEDNOSTRZAŁOWIEC	31	1		1	1	1	3
SZESNASTOBITOWY	28	2		1	1	1	3
ODBLOKOWAJKA	21	1		1	1	1	3
PRZETRWANIÓWKA	9	1		1	1	1	3
MULTIKART	9	1		1	1	1	3
CROSS-CHAT	6	1		1	1	1	3
ZBLUROWANY	6	1		1	1	1	3
EKSBOOKSŁAN	5	1		1	1	1	3
PRZEDŁUŻAJKA	4	1		1	1	1	3
STODWUDZIESTOOŚMIOBITOWIEC	4	1		1	1	1	3
THREE-SIXTY	4	1		1	1	1	3
TWARZOWANIE	4	1		1	1	1	3

W przypadku powyższego zestawienia zmodyfikowałem nieco jeden z warunków przyznawania punktów, mianowicie zwiększyłem zakresy w przypadku sumy wyrazów. Wynika to z tego, że wyrazów na tej płaszczyźnie jest dużo więcej niż w przypadku płaszczyzny zielonej.

Punkty przydzielałem następująco:

ponad 16000 wyrazów – 4 punkty

od 6001 do 16000 wyrazów – 3 punkty

od 201 do 6000 wyrazów – 2 punkty

i poniżej 200 – 1 punkt.

Podobnie, jak na płaszczyźnie zielonej, także tutaj, jeśli pole miało w sumie mniej niż 200 wyrazów, to w kategorii współczynnik zmienności dostawało 1 punkt. W przypadku tego typu pól nie podawałem danych dotyczących współczynnika oraz poziomu nachylenia.

Przyrównując dane z płaszczyzny czerwonej do tych, które pozyskane zostały na płaszczyźnie zielonej, widać, że w pierwszej grupie pól, czyli tych, które uzyskały maksymalną liczbę punktów w powyższym zestawieniu, zmianie uległa tendencja pola *RPG*. Wcześniej była ona malejąca, teraz jest natomiast rosnąca. Na tej podstawie tej zmiany należy wyciągnąć wniosek, że wahania między tendencją malejącą a rosnącą są na tyle małe, iż można przypuszczać, że tendencja tego pola jest bliska stałej. W pozostałych dwóch polach tej grupy tendencja wykazana podczas badania zielonego utrzymała się w badaniu czerwonym. Tym samym informacja ta wzmocniła hipotezę, że ich tendencja jest rosnąca.

Przyglądając się polom znajdującym się w zestawie tych, które uzyskały 11 z 12 możliwych punktów po zastosowaniu systemu 2.0, widać, że nastąpiła zmiana miejsc niektórych pól.

Do zbioru zakwalifikowane zostały pola, których nie było w ogóle na płaszczyźnie zielonej. Pojawiły się tu pola *strzelanina* ‘rodzaj gry, w której elementem wiodącym jest eliminowanie wrogów przy pomocy broni palnej’ oraz *hardkor* ‘coś niewyobrażalnie trudnego zrobienia, np. ukończenie bardzo trudnej gry’. Awansowały one w zestawieniu, ponieważ znacznej zmianie uległa łączna suma wyrazów graficznych reprezentujących te pola. Ze względu na zmianę sumy wyrazów, a mianowicie spadek w porównaniu z płaszczyzną zieloną pole *patch* znajduje się obecnie w trzeciej grupie powyższego zestawienia. Po tych rotacjach widać, że gracze bardzo lubią nie tylko gry *RPG*, ale także, niemalże w równym stopniu gry, w których strzela się do przeciwników. Preferują te, które są *hardkorowe*, czyli wymagają większego doświadczenia w graniu i mają wysoki próg wejścia, w przeciwieństwie od gier

casualowych, które są przeznaczone dla graczy grających rzadko, okazjonalnie. Pole *casual* zajęło trzecie miejsce na platformie zielonej oraz czwarte na czerwonej.

Analizując powyższe zestawienie pod kątem znalezienia pól charakterystycznych dla *konsolowców* i *komputerowców*, widać, że pole *pełniak* pogorszyło swoją pozycję. Na platformie zielonej było w grupie czwartej, natomiast teraz jest w piątej. Zrównało się ono tym samym z innym polem charakterystycznym dla *komputerowców*, mianowicie *gamingiem* ‘nazwa zjawiska, jakim jest granie w gry wideo’. W przypadku pól środowiska konsolowego nadal najważniejszymi są *Marian* oraz *Peesiks*, które w zestawieniu czerwonym uplasowały się wyżej niż w zielonym. W obu przypadkach było to spowodowane zmniejszeniem współczynnika zmienności.

W obu tabelach (zielonej i czerwonej), jak można zauważyć, umieściłem kolory przy nazwach danych pól. Zabieg ten miał na celu podkreślenie części mowy danego wyrazu, który posłużył jako nazwa całego pola semantycznego. Kolory pomarańczowe to rzeczowniki, zielone przymiotniki a czerwone czasowniki (zarówno bezokoliczniki, jak i imiesłowy). Widać, że zdecydowanie dominują rzeczowniki.

Na zakończenie pracy chciałbym jeszcze pokrótce przyjrzeć się systemowi 2.0 w odniesieniu różnych konfiguracji poszczególnych części korpusu.

18.2. SYSTEM 2.0 NA TLE INNYCH ZESTAWIEŃ I POŁĄCZEŃ POSZCZEGÓLNYCH CZĘŚCI KORPUSU – WSZYSTKIE PLATFORMY

System 2.0 jest na tyle uniwersalnym narzędziem, że można go stosować nie tylko do analizy danych w odniesieniu do sumy wyrazów zbadanych pól leksykalnych w całym korpusie, ale także do analizy różnych połączeń konkretnych części korpusu. Trzeba jednak pamiętać, żeby zmienić sposób przyznawania punktów za sumę wyrazów, ponieważ już nie jest analizowany cały korpus, a tylko jego wybrane części. Przykładowo, gdy badam przy pomocy systemu 2.0 materiał pozyskany z korpusów *konsolowców* i *komputerowców* platformy zielonej, to przedziały sumy wyrazów, za które przyznawałem dane punkty (patrz strona 239) zmniejszone zostały o połowę, ponieważ korpus *konsolowców* stanowi połowę całego głównego korpusu. Drugą połowę stanowi korpus *komputerowców*.

W przypadku analizy danych z uwzględnieniem kryterium użytkownika, każda jedna część korpusu zostanie oceniona (uwzględniając sumę wyrazów graficznych) w następujący sposób:

Platforma zielona	Platforma czerwona (p. zielona x ≈ 2)
mniej niż 50 wyrazów graficznych – 1 punkt	mniej niż 100 wyrazów graficznych – 1 punkt
od 51 do 1500 wyrazów graf. – 2 punkty	od 101 do 3000 wyrazów graf. – 2 punkty
od 1501 do 4000 wyrazów graf. – 3 punkty	od 3001 do 8000 wyrazów graf. – 3 punkty
ponad 4000 wyrazów graf. – 4 punkty	ponad 8000 wyrazów graf. – 4 punkty

Oceniając sumę pól dla części korpusu z uwzględnieniem kryterium podziału materiału ze względu na sposób publikacji, czyli na korpusy czasopism oraz newsów⁸⁸, zastosowałam następującą punktację:

Platforma zielona	P. żółta (p. zielona x $\approx 1,5$)	P. czerwona (p. zielona x ≈ 2)
mniej niż 33 w. graf. – 1 punkt	mniej niż 50 w. graf. – 1 p.	mniej niż 66 w. graf. – 1 p.
od 34 do 1500 w. graf. – 2 p.	od 51 do 1500 w. graf. – 2 p.	od 67 do 2000 w. graf. – 2 p.
od 1001 do 6667 w. graf. – 3 p.	od 1001 do 2667 w. graf. – 3 p.	od 2001 do 5334 w. graf. – 3 p.
ponad 2668 w. graf. – 4 p.	ponad 2668 w. graf. – 4 p.	ponad 5334 w. graf. – 4 p.

Jak widać na powyższej tabeli, także tutaj zastosowałam nieco inny sposób punktacji. Przedziały liczbowe w powyższej tabeli stanowią 1/3 przedziałów, które stosowałam przy ocenie całego korpusu (patrz strona 239). Stało się tak, ponieważ jedna część korpusu wyodrębniona ze względu na kryterium sposobu publikacji stanowi 33% całości. Na przykład korpus tekstów czasopism stanowi 1/3 całego korpusu obok takich części jak newsy i fora.

Na początku części właściwej tego rozdziału przeanalizuję i porównam wyniki systemu 2.0 w odniesieniu do korpusów *konsolowców* i *komputerowców*⁸⁹ z tymi, które dotyczyły całego korpusu i przedstawione zostały w poprzedniej części pracy na stronach od 235 do 238 i 243 do 246 (nie w aneksie, ale w głównej części pracy).

⁸⁸ Przy pomocy systemu 2.0 nie mogłem ocenić części korpusu, która powstała z tekstów opublikowanych na forach internetowych, ponieważ tam nie miałem możliwości sprawdzenia częstotliwości rozkładu wyrazów graficznych w poszczególnych kwartałach. Pomiąłem więc tę część w niniejszej analizie.

⁸⁹ W tej części będę odwoływał się do informacji przedstawionych w korpusie na stronach od 166 do 183.

Podczas analizy w odniesieniu do sumy wyrazów całego korpusu stwierdziłem, że najważniejszymi dla całego środowiska polskich graczy wideo są pola: *PC*, *PlayStation* oraz *RPG*, ponieważ to one uzyskały maksymalną liczbę punktów przy wykorzystaniu systemu 2.0. Po rozdeleniu zebranego materiału korpusowego na dwie części dzięki uwzględnieniu kryterium użytkownika widać, że identyczny zestaw pól jest najważniejszy dla grupy korpusów *komputerowców*, natomiast wśród zestawu najważniejszych pól w tekstach *konsolowców* zabrakło *PlayStation*. Warto zaznaczyć, że informacje podane w poprzednim zdaniu dotyczą obu badanych platform – zielonej i czerwonej, zatem nie można mówić tutaj o błędzie.

Pole *PlayStation* straciło jeden punkt za zbyt wysoki współczynnik zmienności. Pole to w odniesieniu do tekstów *konsolowców* jest więc nadal bardzo ważne (zajmuje drugie miejsce pod względem uzyskanych punktów) i jest często używane, ale jednak nie we wszystkich kwartałach.

Przyglądając się wynikom systemu 2.0 dla korpusu *konsolowców* widać, że najważniejszymi polami zaklasyfikowanymi przy pomocy systemu 1.0 do tych powiązanych z *konsolowcami* są *Peesiks* oraz *Marian*, które zajęły ex aequo to samo miejsce w zestawieniu, czyli uzyskały po 8 punktów na platformie zielonej i po 7 na platformie czerwonej. W tych samych zestawieniach widać, że najważniejszymi polami z grupy tych, które zaklasyfikowane zostały przez system 1.0 jako związane z *komputerowcami* są *pełniak*, *gaming* oraz *klatkowy*, które także uzyskały po 8 punktów na platformie zielonej. Natomiast w przypadku platformy czerwonej polem najważniejszym jest tylko *gaming*.

Analiza wyników systemu 2.0 w odniesieniu do korpusów *komputerowców* wskazuje, że najważniejszym polem przypisanym przez system 1.0 do tych, które związane są z *komputerowcami* jest pole *pełniak*, które uzyskało 9 punktów zarówno na platformie zielonej, jak i czerwonej. Natomiast najważniejszymi polami konsolowymi w tym zestawieniu są *cel-shading* oraz *QTE*, które uzyskały po 6 punktów.

Przyjrę się teraz wynikom systemu 2.0 w odniesieniu do korpusów czasopism oraz newsów platformy zielonej, żółtej i czerwonej.

Badania przy zastosowaniu systemu 2.0 w odniesieniu do korpusów czasopism pokazały, że najważniejsze są pola *RPG* oraz *PC*. Potwierdzeniem tego są identyczne wyniki w obu platformach – zielonej i czerwonej. Natomiast w przypadku newsów zaobserwować można pewne rozbieżności między wynikami trzech platform. Na platformie zielonej widać, że tylko jedno pole uzyskało maksymalną liczbę punktów – *PC*, natomiast na żółtej i czerwonej

poza wcześniej wspomnianym polem czołową pozycję zajęło także *PlayStation*. Zatem widać, że w dwóch z trzech badań wystąpiło nie jedno, ale dwa pola, więc trzeba przyjąć, że dla newsów najważniejsze są *PC* oraz *PlayStation*.

W obrębie pól zaklasyfikowanych jako te przynależące do graczy komputerowych wykazano, że najwyższą pozycję zajęło pole *gaming*, które na platformie zielonej i czerwonej uzyskało po 8 punktów. Najważniejszymi polami *konsolowców* w zestawieniach na w obu platformach są: *cel-shading*, *QTE*, *Peesiks* i *Marian*, które uzyskały po 6 punktów. Podczas analizy newsów pod kątem poszukiwania najważniejszych pól związanych z daną grupą graczy wykazano, że w przypadku pól *komputerowców* na wszystkich trzech platformach najważniejsze są *egranizacja* oraz *pełniak*, które uzyskały po 6 punktów.

Najważniejszymi polami *konsolowców* na platformie zielonej są *live-stream* i *Marian*, natomiast na żółtej i czerwonej – *Peesiks*, *live-stream* i *Marian*. Zatem to *live-stream* i *Marian* są tymi najważniejszymi polami.

Krótko podsumowując, najważniejszym polem leksykalnym moich badań jest *PC*, ponieważ niezależnie od tego, jakie przyjmujemy kryteria, to zawsze zajmuje ono czołowe miejsca.

ZAKOŃCZENIE

Badania nad odmianami socjolektalnymi polszczyzny nie mają długiej tradycji, a ich rozkwit datuje się na lata 70. poprzedniego wieku. Do tej pory zbadano między innymi socjolekty takich grup społecznych, jak np.: więźniowie, ochweśnicy, myśliwi, marynarze, parolotniarze czy kierowcy autokarów. Posługując się terminologią Walerego Pisarka można stwierdzić, że były to głównie badania subiektywne, czyli takie, które nie opierały się na wynikach ilościowych – obiektywnych. Niniejsza dysertacja jest próbą zmiany tej sytuacji, ponieważ stanowi propozycję wprowadzenia metodologii badań ilościowych do badań socjolektalnych na zdecydowanie większą skalę, niż obecnie ma to miejsce w różnego rodzaju badaniach tego typu.

Dysertacja miała na celu zbadanie socjolektu polskich graczy wideo pod kątem ilościowym, czyli, jak warto podkreślić jeszcze raz, w inny sposób niż obecnie bada się socjolekty.

W wyniku obserwacji środowiska polskich graczy wideo dostrzeżono jego wewnętrzne zróżnicowanie, czyli wydzielenie w nim dwóch grup graczy: komputerowych oraz konsolowych. Ich charakterystyce poświęcona została część 1.1 oraz 1.2. Poza tymi częściami rozdział pierwszy dopełniony został przeglądem prac związanych z jedną z młodszych dziedzin badań – groznawstwem (1.3). Skupiono się w nim głównie na ukazaniu publikacji, w których przedstawiono wyniki badań nad socjolektem polskich graczy wideo. Z przeglądu wynika, że ta odmiana socjolektalna nie została jeszcze dostatecznie zbadana. Wykazano także, że do tej pory najobszerniejszym ze znanym opracowań zagadnienia jest 30-stronicowy rozdział w książce Dominiki Urbańskiej-Galanciak – *Homo players. Strategie odbioru gier komputerowych*.

Drugi rozdział dysertacji jest przedstawieniem i omówieniem prac związanych z szeroko rozumianymi badaniami nad zróżnicowaniem językowym polszczyzny. Zaczyna się on od przywołania najstarszych prac związanych z dialektologią, czyli dziedziną, która dała początek badaniom nad szeroko rozumianymi odmianami językowymi polszczyzny. Najstarsza przywołana i omówiona pozycja w tym rozdziale to *Rzecz o dyalektach mowy polskiej* Wincentego Pola z 1869 r. autorstwa Drwęcy. To właśnie od daty wydania tej pozycji zaczyna się przegląd literatury dialektologicznej, który doprowadzony został do momentu powstania

socjolingwistyki – dziedziny, która przez dłuższy czas podlegała dialektologii, korzystała z jej dorobku oraz wypracowanych narzędzi naukowych. Po wielu latach istnienia w symbiozie dziedzina ta jednak w końcu uzyskała autonomię, a stało się to m.in. dzięki działaniom i pracom Władysława Lubasia, który uważany jest za ojca polskiej socjolingwistyki. Dokładniejsze informacje na temat dziejów polskiej socjolingwistyki zostały umieszczone w rozdziale 2.2.

Rozwój dialektologii i socjolingwistyki ściśle wiąże się z tworzeniem różnych klasyfikacji odmian językowych polszczyzny. Temat ten podjęty został w rozdziale 2.3, w którym przedstawiono, powstałe na przestrzeni ostatnich 65 lat, koncepcje różnych klasyfikacji języka polskiego. Ich autorami byli tacy badacze, jak: Z. Klemensiewicz, A. Furdal, T. Skubalanka, A. Wilkoń oraz A. Markowski.

Poza tym, co wymieniono powyżej, treść rozdziału drugiego wzbogacona została o:

- informacje na temat prac zapomnianych, ale ważnych dla badań językoznawczych (np. przywołanie artykułu K. Dejny, w którym twierdził on, że na terenie Polski jest tylko jeden język – język polski. Nie można więc używać tego terminu jako zamiennika dla *socjolektu*),
- polemikę z „faktami”, które przedstawione zostały w niektórych pozycjach naukowych (np. na temat pierwszego użycia w literaturze polskiej terminu *socjolekt*),
- interesujące anegdoty (np. o łódzkich warunkach życia H. Ułaszyna).

Rozdział trzeci poświęcony został ukazaniu rozwoju polskich badań ilościowych. W części tej przedstawiono, jakim problemom musiały sprostać różne grupy badawcze na przestrzeni wielu lat istnienia badań ilościowych w Polsce, np. pierwotnie problemem był brak odpowiedniej metodologii, następnie brak dobrych narzędzi badawczych, aż wreszcie brak dostępu do komputerów. Dużą przeszkodą w efektywnej pracy było także to, że wiele komputerów nie obsługiwało polskich znaków. Rozwiązaniem tego problemu było podwojenie liter w miejscu znaku diakrytycznego (np. wyraz *pokaż* zapisywano jako POKAZ, a *pokaż* jako POKAZZ). Ponadto w tej części dysertacji przytoczono interesujące rozwiązania metodologiczne, które okazały się przydatne podczas pracy z obszernym materiałem tekstowym. Jednym z takich rozwiązań było zastosowanie metodologii Kuraszkiewicza, czyli analizowanie wyłącznie pewnej części materiału (np. co piątą stronę) i na jej podstawie wyciąganie wniosków o całości materiału.

Wzorem wielu publikacji (np. W. Lubasia czy W. Miodunki) w dysertacji jako osobną część pracy wydzielono podrozdział – ustalenia terminologiczne (4), w których przedstawiono wybrane terminy niezbędne do precyzyjnego przeprowadzenia badań w pozostałych części

pracy. W rozdziale tym wyjaśniono takie terminy, jak: *język*, *socjolekt*, *wyraz*, *leksem* i *jednostka leksykalna* oraz podano, czym jest *pole leksykalne*. Rozdział 4 kończy pierwszą część pracy, czyli część teoretyczną.

Druga część pracy doktorskiej poświęcona została przedstawieniu zastosowanej w dysertacji metodologii. Jej powstanie wymuszone zostało tym, że współcześnie udostępniana publicznie jest tylko niewielka część narzędzi pomagających w badaniach ilościowych.. Profesjonalne programy są tworzone przez wąskie grona badaczy i nie są udostępniane publicznie. Dostęp do tych narzędzi mają wyłącznie wybrane zespoły badawcze/naukowe (np. pracownicy, którzy zajmują się korpusem PWN). Część narzędzi, co prawda jest dostępna publicznie, jednak nie obsługują one z kolei dużych zbiorów tekstów, więc nie mogły zostać wykorzystane do badań nad socjolektem polskich graczy wideo.

Narzędzie do pracy trzeba więc było przygotować i opracować samodzielnie, posiłkując się wyłącznie różnymi pracami teoretycznymi. Stworzenie takich narzędzi było celem nadrzędnym, który przyświecał powstaniu rozbudowanej części metodologicznej. Wprowadzaniem do niej jest rozdział 5, w którym pokrótce scharakteryzowano wybrany do badań materiał tekstowy. Dobrany został on tak, żeby uwzględniać dwa kryteria: kryterium użytkownika oraz kryterium pochodzenia tekstu. Pierwsze z nich wynikało bezpośrednio z podziału środowiska graczy wideo, którzy tworzyli grupy skupiające się wokół różnego rodzaju sprzętu do gier – wokół konsol (gracze konsolowi – *konsolowcy*) lub wokół komputerów (gracze komputerowi – *komputerowcy*). Każda z grup posiada także swoje czasopismo. Do graczy konsolowych skierowane jest czasopismo „PSX Extreme”, natomiast do *komputerowców* „CD-Action”. Do badań włączone zostały wybrane numery właśnie w tych dwóch czasopism.

Uwzględniając drugie kryterium – pochodzenia tekstu – zdecydowano, że aby badania były bogatsze i pełniejsze, warto jeszcze dodać teksty newsów ‘krótkich informacji na temat najnowszych wydarzeń ze świata’, które pochodziły ze stron www.ppe.pl (strona dla *konsolowców*) i www.cdaction.pl (strona dla *komputerowców*) oraz wpisy z forów internetowych. Tak więc kolejnymi tekstami włączonymi do badań były wpisy na forach internetowych: dla *konsolowców* (www.psxextreme.info) oraz dla *komputerowców* (www.forum.cdaction.pl).

W wyniku doboru materiału uzyskano 6 części korpusu: teksty z czasopisma „PSX Extreme”, teksty z czasopisma „CD-Action”, newsy ze strony www.ppe.pl, newsy ze strony www.cdaction.pl, teksty z forum internetowego www.psxextreme.info oraz teksty z forum www.forum.cdaction.pl.

W następnej części pracy (rozdziale 5.2) pokazano, z jakiego przedziału czasowego pozyskany został materiał. Do badań ostatecznie włączono 60 numerów czasopisma „PSX Extreme” oraz tyle samo numerów czasopisma „CD-Action”, po 6000 newsów ze stron ppe.pl i cdaction.pl oraz wpisy z forów internetowych. Wszystkie teksty zostały opublikowane w latach 2011–2015.

Uzupełnieniem rozdziału 5 jest część 5.3, gdzie wyjaśniono skrótowe nazwy korpusów, które wykorzystywane są we wszystkich częściach pracy.

W następnym rozdziale pracy zaprezentowano metodologię pozyskiwania i pracy z materiałem tekstowym. Rozdział ten jest swoistym dziennikiem pracy, czyli dominuje w nim opis pierwszoosobowy, a jego treść podzielona została na pięć etapów pracy, których ostatecznym celem było uzyskanie danych liczbowych. Posłużyły one do dalszych badań przedstawionych w części analitycznej niniejszej pracy doktorskiej.

Pierwszym etapem było pozyskanie edytowalnych materiałów tekstowych (takich, które można dowolnie modyfikować) z czasopism drukowanych. W tym fragmencie pracy skupiono się na pokazaniu sposobu pracy z programem ABBYY FineReader 12 Professional Edition, przy pomocy którego w łatwy sposób można osiągać zamierzony cel, czyli uzyskać wersje edytowalne czasopism. Dzięki skanowaniu poszczególnych stron z czasopism drukowanych oraz ich obróbce komputerowej uzyskano po 20 plików w formacie wordowskim – *docx*. Jeden plik zawierał w sobie jeden kwartał numerów danego czasopisma. Warto zaznaczyć, że w pracy przyjęto, iż kwartał jest najmniejszą jednostką podziału zebranego materiału tekstowego.

Etap drugi dotyczył pozyskania materiałów tekstowych z internetu. Podczas opisu metodologii pozyskania tekstów skupiono się na newsach oraz wpisach na forach internetowych. Większą część tego etapu stanowi opis pracy z programami HTTrack Website Copier oraz Total Commander. Pierwszy z nich umożliwił pobranie tekstów ze stron internetowych, natomiast drugi opracowanie pozyskanego materiału. Efektem prac z programami było uzyskanie po 20 plików *docx*, które zawierają po 1 kwartale newsów oraz dwóch zbiorów z wpisami na forach internetowych, których ze względu na specyfikę publikowanych tam treści nie udało podzielić się na kwartały.

Kolejnym etapem było przygotowanie materiału leksykalnego do pracy z korpusami tekstów. W tej części dysertacji przedstawiono sposób doboru jednostek socjolektalnych do badań ilościowych. Przy czym trzeba zaznaczyć, że wybrane zostały one z tekstów czasopism, newsów oraz wpisów na forach, po czym wynotowane i dodane do wspólnej bazy danych. W efekcie powstał duży zbiór jednostek leksykalnych (1732) który następnie podzielony został

na cztery kręgi tematyczne – *branża gier wideo, środowisko graczy, sprzęt gracza, gra*. Z każdego kręgu wybrano po 15% jednostek, czyli pozyskano ich łącznie 260.

W etapie czwartym przedstawiono sposób pracy ze zgromadzonymi jednostkami, czyli zademonstrowano, jak wyszukiwane były w korpusie ich realizacje tekstowe – wyrazy graficzne. W tej części pracy zaprezentowany został sposób wyszukiwania wyrazów graficznych przy pomocy wyszukiwarki wbudowanej w program Word z włączoną obsługą symboli wieloznacznych. Ponadto ten fragment pracy wzbogacony został sposobami rozwiązania pewnych problemów związanych z przedstawioną metodą wyszukiwania, czyli np. jak poradzić sobie z bardzo krótkimi wyrazami, które mogą stanowić problem dla wyszukiwarki oraz z wyszukiwaniem wyrazów graficznych, w których brakuje znaków diakrytycznych (ma to miejsce zwłaszcza w tekstach pozyskanych z internetu).

Zwieńczeniem prac przedstawionych w części metodologicznej niniejszej dysertacji było skumulowanie wcześniej wybranych jednostek leksykalnych w pola wyrazowe. Łącznie uzyskano ich 189. W skład jednego pola wchodziły różne formy podstawowe, których liczebność wahała się od 24 do 1. Być może nazwanie *polem* zbioru jednoelementowego (jest ich 69) może budzić kontrowersje, ale tak naprawdę badania ilościowe zawsze prowadzone są na mniejszych lub większych próbkach korpusów, a więc założono, że zbiór jednoelementowy mógłby ulec rozbudowaniu w kolejnych próbach badawczych. Z kolei tworzenie pól polegało na łączeniu ekwiwalentów znaczeniowych (nadrzędne było odniesienie do tego samego desygnatu), rodzin wyrazów (ta sama podstawa słowotwórcza mogła pojawiać się w rzeczowniku, przymiotniku, czasowniku, przysłówku), wariantów graficznych (różne sposoby zapisu). Każda z form podstawowych, o ile była odmienna, występowała w tekstach w różnych formach fleksyjnych. Dopiero cały wachlarz form podstawowych, i ewentualnych form fleksyjnych, został poddany sprawdzeniu za pomocą specjalnie stworzonych narzędzi pod kątem frekwencji w tekstach. Tak więc badaniu ilościowemu podlegała liczebność wyrazów graficznych reprezentujących określone pola wyrazowe.

Przykładowo pole *Marian* ‘główny bohater gry Super Mario Bros’ zawiera w sobie 2 formy podstawowe: *Marian* (od niej nazwę wzięło całe pole) oraz *Marianek*. Każda z tych form realizuje się w tekstach poprzez różne formy fleksyjne, np.: *Mariana, Marianem, Marianami* itd., natomiast *Marianek: Mariankiem, Marianku, Mariankom*. Ostatecznie badana jest frekwencja w tekstach konkretnych wystąpień wyrazów graficznych reprezentujących wszystkie możliwe formy z tego pola.

Część analityczną dysertacji otwiera rozdział *Pola leksykalne a ich złożoność*, w którym scharakteryzowane zostały pola wyrazowe pod kątem ich złożoności, czyli liczebności form podstawowych w nich zawartych. Uzupełnieniem tego rozdziału była analiza form podstawowych w celu zbadania występujących w nich części mowy. Badania wykazały, że wśród wszystkich form podstawowych dominują rzeczowniki, na drugim miejscu są przymiotniki, a na trzecim czasowniki.

Informacje przedstawione w kolejnym rozdziale (8) dotyczą sumy wyrazów graficznych w poszczególnych częściach korpusu. Ich analiza wykazała, że korpusy nie są równe pod względem liczebności wyrazów graficznych, dlatego w celu ich matematycznego wyrównania w dysertacji zastosowano różnego rodzaju mnożniki. Innymi słowy, zostały one wprowadzone w celu wyrównania liczby wszystkich wyrazów we wszystkich częściach korpusu, a więc także w celu wyrównania liczby wszystkich badanych wyrazów graficznych ze wszystkich pól leksykalnych. Efektem tego zabiegu było powstanie trzech platform badawczych: zielonej, żółtej oraz czerwonej. Platforma zielona nie wykorzystywała żadnych mnożników, więc jej dane są efektem wyłącznie wyszukiwania wyrazów graficznych przy pomocy wyszukiwarki programu Word. Platforma żółta wyrównywała dane liczbowe w korpusach newsów, natomiast czerwona we wszystkich częściach korpusu zrównując je z tymi, które zawarte zostały w korpusach pozyskanych z wpisów na forach internetowych (były to korpusy najobszerniejsze wyrazowo). Wyniki zrównania wyrazów w poszczególnych platformach przedstawione zostały w rozdziale 9 (ogólnie – w odniesieniu do sumy wyrazów graficznych) oraz 10 (szczegółowo – z podziałem na kwartały).

W rozdziale 11 zostały zestawione i porównane pod względem liczebności zbadane wyrazy graficzne graczy (te które połączone zostały w pola leksykalne) z resztą wyrazów pozostałych w korpusie (głównie wyrazów z polszczyzny ogólnej). W rozdziale 12 także zbadano wyrazy graficzne graczy, ale pod kątem występujących w nich części mowy. Identycznie, jak w przypadku form podstawowych, także i tutaj dominują rzeczowniki, na drugim miejscu pod względem liczebności są przymiotniki, a za nimi czasowniki.

Rozważania przedstawione w częściach omówionych powyżej doprowadziły do wniosku, że pewne pola występują wyłącznie w niektórych częściach korpusu, np. tych, które związane są z *konsolowcami*. Spostrzeżenie to doprowadziło do powstania systemu 1.0, któremu poświęcony został cały 13 rozdział dysertacji. Opracowany system pozwolił na podział i wskazanie pól, które związane są z daną grupą graczy wideo. Dodatkowo po uwzględnieniu zasad opisanych w części 13.2 wykazano, że istnieją pola, które wystąpiły wyłącznie w tekstach graczy konsolowych oraz takie, które są charakterystyczne wyłącznie dla

komputerowców. System 1.0 nie ogranicza się wyłącznie do podania informacji o tym, czy dane pole związane jest z konkretną grupą graczy, ale pozwala stwierdzić, jaki jest ich stopień takiego powiązania. Analiza wykazała, że grupa pól, które w jakimś stopniu związane są z daną grupą graczy (nie są polami ponad podziałami) stanowi ponad 25% wszystkich badanych pól leksykalnych.

Następny rozdział (14) jest rezultatem połączenia wyników systemu 1.0 oraz danych na temat złożoności pól leksykalnych, które wykazane zostały w jednym z poprzednich rozdziałów dysertacji.

Kolejnym fragmentem pracy jest rozdział 15, w którym przeanalizowano pola wyrazowe pod kątem sum wyrazów graficznych. Wskazano pola najbogatsze oraz najuboższe wyrazowo, a także pokazano, jaki procent one zajmują na tle pozostałych pól. Ponadto w tej części wskazano, które teksty (opublikowane w czasopismach, newsach lub na forach internetowych) są najbardziej nasycone najbogatszymi i najuboższymi wyrazowo polami. Uzupełnieniem tego rozdziału było przeanalizowanie danych korpusowych z uwzględnieniem kryteriów użytkownika oraz sposobu publikacji. Analiza ta dotyczyła wszystkich badanych platform (zielonej, żółtej i czerwonej).

Dane z rozdziału 15 wykorzystano także w następnym rozdziale – 16, w którym połączono pozyskane informacje na temat sumy wyrazów graficznych w całym korpusie z danymi o złożoności badanych pól leksykalnych.

W rozdziale 17, który poświęcony został częstotliwości wystąpień wyrazów graficznych pokazano, jak na przestrzeni pięciu lat zmieniała się liczba użycie wyrazów graficznych poszczególnych pól. Szczególną uwagę zwrócono na pola, które powstały, zaniknęły lub wykazywały pewne anomalie pod względem liczby użycia w badanym okresie. Ponadto omówiono, czym te anomalie są spowodowane oraz jaki jest wpływ czynników zewnętrznych na nie. Uzupełnienie stanowiły informacje na temat tego, czy pola wyrazowe wykazują tendencję spadkową pod względem użycia wyrazów graficznych w tekstach, czy też mają tendencje rosnącą. Analizy w tym rozdziale dokonano także z uwzględnieniem podziału na korpusy *konsolowców* oraz *komputerowców*, a także na te, które dotyczyły newsów lub czasopism.

Warto podkreślić, że badania przedstawione w tej części pracy potwierdzają wyraźną zbieżność między częstotliwością wystąpienia określonej leksyki a wydarzeniami z nią związanymi (np. z premierą nowego sprzętu do gier wzrost zaobserwowano w polach: *next-gen* oraz *rozdzielczość*).

Rozdziałem podsumowującym całą pracę jest *System 2.0*. Na podstawie zestawień wyników liczbowych, które dotyczyły sumy, złożoności oraz częstotliwości wystąpień wyrazów w badanych polach powstała klasyfikacja, którą nazwano systemem 2.0. Jest on swoistym zwieńczeniem tej pracy, ponieważ dzięki temu systemowi (zastosowaniu jego zasad) udało się wydzielić najważniejsze dla środowiska graczy wideo pola. Tym samym nieco obaląc istniejącą powszechnie opinię, że jeżeli danych wyrazów jest dużo w badanym korpusie, to są one ważne. Otóż nie zawsze tak jest, ponieważ w badaniach ilościowych tego typu jak moje trzeba jeszcze uwzględnić inne czynniki takie, jak np. częstotliwość czy złożoność, aby móc mówić o ważności danych pól leksykalnych. System 2.0 zastosowano w odniesieniu do wszystkich platform badawczych oraz podziału części korpusu, które uwzględniają różnego rodzaju kryteria. Z badań wynika, że polem, które wystąpiło we wszystkich wykonanych do tej pracy zestawieniach systemu 2.0 jest *PC*. Można więc stwierdzić, że jest ono najważniejszym polem całych badań nad socjolektem polskich graczy wideo.

Na zakończenie warto jeszcze sobie uświadomić, że badania przedstawione w tej dysertacji dotyczyły tylko niewielkiej części całego socjolektu polskich graczy wideo. Zbadano tylko jego wersję pisaną, zatem do zbadania została jeszcze odmiana mówiona. Badania obejmowały wyłącznie teksty opublikowane w latach 2011–2015, więc do zbadania pozostał okres poprzedzający te lata. Bardzo ciekawa byłaby także analiza socjolektu polskich graczy wideo poczynając od lat 90., kiedy to tak naprawdę on powstawał. Ciekawie byłoby przyjrzeć się także najnowszemu okresowi, czyli od roku 2016 do 2018 i porównać z wynikami opublikowanymi w niniejszej pracy.

W dysertacji skupiono się tylko na polach wyrazowych, które zawierały prawie wyłącznie jednowyrazowe formy graficzne, trzeba by więc jeszcze przyjrzeć się połączeniom wyrazowym charakterystycznym dla tego środowiska. A ponadto zbadać nowe środowisko graczy, czyli graczy mobilnych. Przyjrzeć się ich mowie oraz tekstom.

W pracy skupiłem się na przeanalizowaniu tekstów środowiska graczy wideo między innymi z podziałem na *konsolowców* i *komputerowców*. A trzeba pamiętać, że grupy dzielą się jeszcze na graczy hardkorowych i casualowych, których odmiany językowej w tej pracy nie zbadano. Zadanie to zostaje do zrealizowania w innych, kolejnych pracach.

Badanie socjolektu to zadanie rozległe pod względem wyboru okresów badawczych, aspektów analizy, zarówno o uwarunkowaniach społecznych, jak i rodzaju tekstów. Ilościowe rozeznanie w korpusach jest niezbędnym etapem poznania socjolektów, co otwiera wciąż nowe obszary badawcze.

BIBLIOGRAFIA

- Antas J., 1991, *Ogólna charakterystyka procentowego udziału poszczególnych części mowy w odmianach języka telewizji polskiej. Dane tekstowe*, w: Z. Kurzowa (red.), *Badania nad językiem telewizji polskiej. Ilościowy opis słownictwa*, Warszawa–Kraków, s. 13–22.
- Bańko M., 2003, *Wyraz a słowo*, www.sjp.pwn.pl/poradnia/haslo/wyraz-a-slowo;3506.html [dostęp: 18.12.2017].
- Bartłomiejczyk M., 2012, *O języku kierowców autokarów*, „Socjolingwistyka”, t. 26, s. 191–204.
- Błażejczyk-Majka R., Kala R., 2010, *Estymacja kombinowana granicznej funkcji produkcji*, w: B. Borkowski (red.), *Metody ilościowe w badaniach ekonomicznych*, Warszawa, s. 71–80.
- Budziszewska W., 1957, *Żargon ochweśnicki*, Łódź.
- Budzyk K., 1959, *Lingwistyka statystyczna a zadania wiedzy o stylu i wierszu*, „Przegląd Humanistyczny”, r. 3, Warszawa, s. 31–40.
- Buttler D., 1977, *Dobór wyrazów do słownika – minimum języka polskiego*, „Poradnik Językowy”, z. 6, Warszawa, s. 253–262.
- Buttler D., 1982, *Miejsce języka potocznego wśród odmian współczesnej polszczyzny*, w: S. Urbańczyk (red.), *Język literacki i jego warianty*, Wrocław–Warszawa–Kraków–Gdańsk–Łódź.
- Chmielewska-Luczak D., Smejliś T., 2005, *Komunikacja w świecie gier komputerowych*, w: A. Dytman-Stasienko, J. Stasienko (red.), *Język@multimedia*, Wrocław.
- Chruściej J., *Język komunikacji uczestników gier sieciowych*, niepublikowana praca magisterska.
- Czekanowski J., 1947, *Polska synteza slawistyczna w perspektywie ilościowej*, Kraków.
- Czeszewski M., 2001, *Słownik slangu młodzieżowego*, Piła.
- Dejna K., 1987, *Język polski i jego odmiany*, „Rozprawy Komisji Językowej”, t. 33, Wrocław–Warszawa–Kraków–Gdańsk–Łódź, s. 37–45.
- Doroszewski W., 1934, *Mowa mieszkańców wsi Staroźreby. Opis i tezy teoretyczne*, „Prace Filologiczne”, t. 16, Warszawa, s. 249–278.
- Dovey J., Kennedy H.W., 2011, *Kultura gier komputerowych*, Kraków.
- Drwęca, 1869, *Rzecz o dyalektach mowy polskiej Wincentego Pola*, Kraków.
- Dunaj B., 1986, *Dialektologia a socjolingwistyka*, „Acta Universitatis Lodzianensis. Folia Linguistica”, t. 12, Łódź, s. 15–23.
- Estreicher K., 1859, *Język złoczyńców*, „Rozmaitości”, nr. 12–14, Lwów.
- Estreicher K., 1903, *Szwargot więzienny*, Kraków.
- Filiciak M., 2006, *Wirtualny plac zabaw. Gry sieciowe i przemiany kultury współczesnej*, Warszawa.

- Furdal A., 1966, *Podział polskich dialektów miejskich*, „Rozprawy Komisji Językowej Wrocławskiego Towarzystwa Naukowego”, z. 6, Wrocław, s. 155–162.
- Furdal A., 1973, *Klasyfikacja odmian współczesnego języka polskiego*, Wrocław.
- Gajda S., 1982, *Podstawy badań stylistycznych nad językiem naukowym*, Wrocław.
- Gajda S., 2001, *Funkcjonowanie współczesnej polszczyzny*, w: S. Gajda (red.), *Najnowsze dzieje języków słowiańskich. Język polski*, Opole, s. 207–219.
- Garda M.B, Grabarczyk P., 2014, „Replay. The Polish Journal of Game Studies”, nr 1, Łódź.
- Gądek K., 2009a, *Metody opisu interakcji: gracz – postać – elementy świata gry w języku graczy komputerowych*, „Poradnik Językowy”, z. 4, Warszawa, s. 52–64.
- Gądek K., 2009b, *Metodologiczne podstawy opisu języka graczy komputerowych*, w: A. Czelakowska, M. Skarżyński (red.), *Biblioteka Lingwariów. Język z różnych stron widziany*, Kraków, s. 221–227.
- Gądek K., 2009c, *Cudzysłów jako wyraz niepewności i dystansu w tekstach dotyczących gier komputerowych*, „Poradnik Językowy”, z. 2, Warszawa, s. 58–68.
- Głowacki J., 1990, *O języku żołnierskim*, „Poradnik Językowy”, z. 2, Warszawa, s. 119–122.
- Górka S., 1901, *Skulscy „ochweśnicy”*, „Wisła. Miesięcznik geograficzny i etnograficzny”, t. 15, z. 1, Warszawa, s. 1–7.
- Grabias S., 1974, *Funkcyjna klasyfikacja socjalnych wariantów języka*, „Język Polski. Organ Towarzystwa Miłośników Języka Polskiego”, r. 54, z. 1, Kraków, s. 22–31.
- Grabias S., 1997, *Język w zachowaniach społecznych*, Lublin.
- Handke K., 1995, *Polski język familijny*, Warszawa.
- Hebal-Jeziarska M., Kaczmarska E. (red.), 2015, „Prace Filologiczne”, t. 67, Warszawa.
- Hrabec S., 1958, *Teoretyczne założenia „Słownika polszczyzny XVI wieku” i ich praktyczna realizacja*, w: P. Zwoliński (red.), *Z polskich studiów slawistycznych*, Warszawa, s. 233–239.
- Inglot K., 2013, *Nazwy własne a lokalizacja gier elektronicznych fantasy – o potrzebie przekładu oraz wybranych tendencjach w tłumaczeniu gier anglojęzycznych na rynek polski i niemiecki*, „Homo Ludens”, z. 1, Poznań, s. 73–79.
- Isaacson W., 2011, *Steve Jobs*, New York.
- Jaros V., 2012, *Język użytkowników CB-radia a dawna gwara szoferska w kontekście przemian językowych i społeczno-cywilizacyjnych w Polsce*, w: E. Laskowska, B. Morzyńska-Wrzosek, W. Czechowski (red.), *Język – natura – cywilizacja*, Bydgoszcz, s. 91–102.
- Jędrasiak K., *Gracz jako model uczestnika współczesnej kultury*, niepublikowana praca doktorska.
- Kamińska M., 1986, *Powiązania dialektologii z socjolingwistyką*, „Acta Universitatis Lodziensis. Folia Linguistica”, t. 12, Łódź, s. 47–53.
- Kaniuka W., 1972, *Tendencje rozwojowe współczesnej socjolingwistyki*, „Biuletyn Polskiego Towarzystwa Językoznawczego”, z. 30, Wrocław–Warszawa–Kraków–Gdańsk, s. 41–48.
- Karłowicz J., 1906, *Lud. Rys ludoznawstwa polskiego*, Lwów.

- Karwatowska M., Szypra-Kozłowska J., 2005, *Lingwistyka płci. Ona i on w języku polskim*, Lublin.
- Kasperczak M., 2004, *Język kobiet – język mężczyzn*, „Roczniki Humanistyczne. Językoznawstwo”, t. 52, z. 6, Lublin, s. 39–75.
- Klemensiewicz Z., 1953, *O różnych odmianach współczesnej polszczyzny*, Warszawa.
- Kniaginiowa M., 1962, *Próba zastosowania metod statystycznych w badaniach stylistyczno-składniowych*, „Język Polski”, r. 42, z. 2, Kraków, s. 92–116.
- Kochanowicz R., 2012, *Fabularyzowane gry komputerowe w przestrzeni humanistycznej*, Poznań.
- Kolberg O., 1857, *Pieśni ludu polskiego*, s. 1, Warszawa.
- Kołodziejek E., 1990, *O gwarze marynarzy i subkulturze marynarskiej*, „Socjolingwistyka”, t. 9, Wrocław–Warszawa–Kraków–Gdańsk–Łódź, s. 129–139.
- Kopernicki J., 1875, *Spostrzeżenia nad właściwościami językowymi w mowie górali bieskidowych, z dodatkiem słowniczka wyrazów góralskich*, „Rozprawy i sprawozdania z posiedzeń Wydziału Filologicznego Akademii Umiejętności”, s. 1, t. 3, Kraków, s. 343–379.
- Kosiński W., 1884, *Przyczynek do gwary zakopańskiej*, „Rozprawy i sprawozdania z posiedzeń Wydziału Filologicznego Akademii Umiejętności”, s. 1, t. 10, Kraków, s. 225–309.
- Kubiński P., 2016, *Gry wideo. Zarys poetyki*, Kraków.
- Kubiszyn-Mędrala Z., 1991, *100 najczęstszych wyrazów w języku telewizji polskiej*, w: Z. Kurzowa (red.), *Badania nad językiem telewizji polskiej. Ilościowy opis słownictwa*, Warszawa–Kraków, 41–60.
- Kujawa Ł., 2014a, *Knajpa dla gracza*, „PSX Extreme”, nr 207, Warszawa, s. 16–20.
- Kujawa Ł., 2014b, *Konsolowy e-sport*, „PSX Extreme”, nr 208, Warszawa, s. 18–21.
- Kupiszewski P., 1987, *Język żołnierski*, „Poradnik Językowy”, z. 4, Warszawa, s. 273–281.
- Kuraszkiewicz W., 1985, *Statystyczne badania słownictwa polskich tekstów XVI wieku*, w: P. Zwoliński (red.), *Z polskich studiów slawistycznych*, Warszawa, s. 240–257.
- Kuraszkiewicz W., Łukaszewicz J., 1951, *Ilość różnych wyrazów w zależności od długości tekstu*, Wrocław.
- Laskowski R., 1999a, *Wyraz*, w: *Encyklopedia języka polskiego*, red. S. Urbańczyk, M. Kucała, Wrocław–Warszawa–Kraków, s. 427.
- Laskowski R., 1999b, *Leksem*, w: *Encyklopedia języka polskiego*, red. S. Urbańczyk, M. Kucała, Wrocław–Warszawa–Kraków, s. 205.
- Laskowski R., 1999c, *Zagadnienia ogólne morfologii*, w: R. Grzegorzczkova, R. Laskowski, H. Wróbel (red.), *Gramatyka współczesnego języka polskiego. Morfologia*, Warszawa, s. 27–86.
- Lehr-Splawiński T., 1924, *Kilka uwag o praojczyźnie Słowian*, „Sprawozdania Towarzystwa Naukowego we Lwowie”, r. 4, z. 2, Lwów, s. 62–65.
- Lewandowska-Tomaszczyk B. (red.), 2005, *Podstawy językoznawstwa korpusowego*, Łódź
- Lewicki A., Kurcz I., Sambor J., Woronczak J., 1974, *Słownictwo współczesnego języka polskiego. Listy frekwencyjne. Teksty popularnonaukowe*, t. 1, Warszawa.

- Lewicki A., Masłowski W., Sambor J., Woronczak J., 1972, *Słownictwo współczesnej publicystyki polskiej*, Warszawa.
- Lubaś W., 1979, *Socjolingwistyka jako metoda badawcza*, „Socjolingwistyka”, t. 2, Warszawa–Kraków–Katowice, s. 11–27.
- Lubaś W., 1982, *Socjolingwistyka – metoda interdyscyplinarna*, „Socjolingwistyka”, t. 4, Warszawa–Kraków–Katowice, s. 11–18.
- Lutosławski W., 1896, *Sur une nouvelle méthode pour déterminer la chronologie des dialogues de Platon*, Paryż, s. 13–32.
- Malinowski L., 1871, *Listy z podróży etnograficznej po Szląsku*, „Na dziś”, z. 1, Kraków, s. 289–313.
- Mańkowski P., 2010, *Cyfrowe marzenia. Historia gier komputerowych i wideo*, Warszawa.
- Markowski A., 2012, *Wykłady z leksykologii*, Warszawa.
- Matusiak S., 1880, *Gwara Lasowska w okolicy Tarnobrzega. Studium dialektologiczne*, „Rozprawy i sprawozdania z posiedzeń Wydziału Filologicznego Akademii Umiejętności”, s. 1, t. 8, Kraków, s. 70–179.
- Mayenowa M.R., 1965, *Możliwości i niebezpieczeństwa metod matematycznych w poetyce*, w: M.R. Mayenowa (red.) *Poetyka i matematyka*, Warszawa, s. 5–20.
- Miodunka W., 1989, *Podstawy leksykologii i leksykografii*, Warszawa.
- Mizerski W., 2000, *Język polski. Encyklopedia w tabelach*, Warszawa.
- Morgenthaler R., 1958, *Statistiki des Neutestamentlichen Wortschatzes*, Zurych.
- Murrmann J., 2015, *Język sportu w ujęciu leksykografów. Analiza terminograficzna wybranych słowników z terminologią sportową*, „Socjolingwistyka”, t. 29, s. 245–267.
- Nagórko A., 2003, *Zarys gramatyki polskiej*, Warszawa.
- Nitsch K., 1910, *Próba ugrupowania gwar polskich*, „Rozprawy Akademii Umiejętności. Wydział Filologiczny”, s. 3, t. 1, Kraków, s. 336–365.
- Nitsch K., 1911, *Mowa ludu polskiego*, Kraków.
- Nowy słownik poprawnej polszczyzny*, 1999, red. A. Markowski, Warszawa.
- Pedzich B., 2010, *Jak powstaje środowiskowa odmiana języka (na przykładzie leksyki paralotniarskiej)*, w: M. Milewska-Stawina, E. Rogowska-Cybulska (red.), *Polskie języki. O językach zawodowych i środowiskowych*, Gdańsk, s. 117–122.
- Pędzich B., 2012, *Jak powstaje socjolekt? Studium słownictwa paralotniarzy*, Warszawa.
- Piekot T., 2008, *Język w grupie społecznej. Wprowadzenie do analizy socjolektu*, Wałbrzych.
- Pisarek W., 1972, *Frekwencja wyrazów w prasie. Wiadomości – komentarze – reportaże*, Kraków.
- Pisarek W., 1999, *Socjolekt*, w: *Encyklopedia języka polskiego*, red. S. Urbańczyk, M. Kucala, Wrocław–Warszawa–Kraków s. 364.
- Pol W., 1869, *Rzecz o dialektach mowy polskiej*, Kraków.
- Polański K., 1999, *Język*, w: *Encyklopedia języka polskiego*, red. S. Urbańczyk, M. Kucala, Wrocław–Warszawa–Kraków, s. 153–154.

- Przepiórkowski A., Bańko M., Górski R.L., Lewandowska-Tomaszczyk B., 2012, *Narodowy Korpus Języka Polskiego*, Warszawa.
- Roślowski B., 1978, *O zastosowaniu komputerów w badaniach języka mówionego*, w: S. Grabias, J. Mazur, K. Pisarkowa (red.), *Studia nad składnią polszczyzny mówionej. Księga referatów konferencji poświęconej składni i metodologii badań języka mówionego (Lublin 6–9 X 1975)*, Wrocław–Warszawa–Kraków–Gdańsk, s. 57–64.
- Rudolf M., 2004, *Metody automatycznej analizy korpusu tekstów polskich. Pozyskiwanie, wzbogacanie i przetwarzanie informacji lingwistycznych*, Warszawa.
- Sambor J., 1969, *Badania statystyczne nad słownictwem (na materiale „Pana Tadeusza”)*, Wrocław–Warszawa–Kraków.
- Sambor J., 1972, *Słowa i liczby. Zagadnienia językoznawstwa statystycznego*, Wrocław–Warszawa–Kraków–Gdańsk.
- Sambor J., 1975, *O słownictwie statystycznie rzadkim*, Warszawa.
- Seretny A., 2011, *Kompetencja leksykalna uczących się języka polskiego jako obcego w świetle badań ilościowych*, Kraków.
- Skarżyński M., 1991, *Charakterystyka ilościowa form homonicznych w języku telewizji polskiej*, w: Z. Kurzowa (red.), *Badania nad językiem telewizji polskiej. Ilościowy opis słownictwa*, Warszawa–Kraków, s. 115–127.
- Skubalanka T., 1976, *Założenia analizy stylistycznej*, w: H. Markiewicz, J. Sławiński (red.), *Problemy metodologiczne współczesnego literaturoznawstwa*, Kraków.
- Słownik frekwencyjny współczesnej polszczyzny*, 1990, red. I. Kurcz, A. Lewicki, J. Sambor, K. Szafran, J. Woronczak, Kraków.
- Słownik gramatyki języka polskiego*, 2002, red. W. Gruszczyński, J. Bralczyk, Warszawa.
- Słownik gwar polskich, 1910–1911*, red. J. Karłowicz, Kraków.
- Słownik języka polskiego, 1958–1969*, red. W. Doroszewski, Warszawa.
- Słownik wiedzy o języku*, 2004, red. D. Podlaska, I. Płóciennik, Bielsko-Biała.
- Szczerbowicz L.W., 1890, *Gwara warszawska w społecznym języku literackim*, „Biblioteka warszawska”, t. 199, Warszawa, s.202–220.
- Świdziński M., Derwojoweda M., Rudolf M., 2002, *Dehomonizacja i desynkretyzacja w procesie automatycznego przetwarzania wielkich korpusów tekstów polskich*, „Biuletyn Polskiego Towarzystwa Językoznawczego”, z. 58, Kraków.
- Taszycki W., Milewski T., 1956, *Polski język literacki powstał w Małopolsce*, w: K. Budzyk (red.), *Pochodzenie polskiego języka literackiego*, Wrocław, s. 364–436.
- Tokarski J., 1993, *Schematyczny indeks a tergo polskich form wyrazowych*, Warszawa.
- Tomiałoje L., 1970, *Badania ilościowe nad synantropijną awifauną Legnicy i okolic*, „Acta Ornithologica”, t. 12, nr 9, Warszawa, s. 292–325.
- Ułaszyn H., 1951, *Język złodziejski. La langue des voleurs*, Łódź.

- Urbańczyk S., 1956, *Rozwój języka narodowego. Pojęcia i terminologia*, Z. Stieber (red), *Z dziejów powstania języków narodowych i literackich*, Warszawa, s. 9–36.
- Urbańczyk S., 1981, *Zarys dialektologii polskiej*, Warszawa.
- Urbańska-Galanciak D., 2005, *Elementy socjolektu graczy komputerowych*, w: A. Dytman-Stasienko, J. Stasienko (red.), *Język@multimedia*, Wrocław.
- Urbańska-Galanciak D., 2009, *Homo players. Strategie odbioru gier komputerowych*, Warszawa.
- Wielki słownik ortograficzny*, 2017, red. E. Polański, Warszawa.
- Wielki słownik poprawnej polszczyzny*, 2006, red. A. Markowski, Warszawa.
- Wilkoń A., 1987, *Typologia odmian językowych współczesnej polszczyzny*, Katowice.
- Wilkoń A., 1988, *Typologia socjolektów*, „Socjolingwistyka”, t. 8, Wrocław–Warszawa–Kraków–Gdańsk–Łódź, s. 83–93.
- Wilkoń A., 2000, *Typologia odmian językowych współczesnej polszczyzny*, Katowice.
- www.arstechnica.com/gaming/2014/11/ubisoft-working-to-fix-game-breaking-assassins-creed-unity-bugs, dostęp: 23.07.2018.
- www.gamerankings.com, dostęp: 9.04.2017.
- www.gieldamedia.pl/archiwum/652-psx-extreme, dostęp: 8.04.2017.
- www.gry-online.pl/S008.asp, dostęp: 9.04.2017.
- www.gry-online.pl/slownik-gracza.asp, dostęp: 26.05.2017.
- www.korpusomat.nlp.ipipan.waw.pl/manual, dostęp: 22.12.2017.
- www.metacritic.com, dostęp: 9.04.2017.
- www.sjp.pwn.pl/poradnia/haslo/konsola;3437.html, dostęp: 5.06.2017.
- www.sjp.pwn.pl/slowniki/konsoli.html, dostęp: 6.04.2017.
- www.statystyka.cba.pl/wspolczynnik.zmiennosci.html, dostęp: 22.05.2018.
- www.zkdp.pl/images/komunikat_2013.pdf, dostęp: 7.04.2017.
- www.zwoleja.blog.onet.pl/2014/07/22/w-poznaniu-wrocilem-na-swoj-level-kamil-majchrzak, dostęp: 8.04.2017.
- Zabrocki L., 1972, *Z teorii socjolingwistyki*, „Biuletyn Polskiego Towarzystwa Językoznawczego”, z. 30, Wrocław–Warszawa–Kraków–Gdańsk, s. 17–25.
- Zarębina M., 1971, *Najczęstsze wyrazy polszczyzny mówionej*, „Język Polski”, r. 51, z. 5, Kraków, s. 336–347.
- Zarębina M., 1978, *Ustalenie korpusu tekstowego dla badań statystycznych nad słownictwem polszczyzny mówionej*, [w:] S. Grabias, J. Mazur, K. Pisarkowa (red.), *Studia nad składnią polszczyzny mówionej. Księga referatów konferencji poświęconej składni i metodologii badań języka mówionego (Lublin 6–9 X 1975)*, Wrocław–Warszawa–Kraków–Gdańsk, s. 103–110.
- Zarębina M., 1982, *Badania statystyczne nad słownictwem polszczyzny mówionej*, „Poradnik Językowy”, z. 9, Warszawa, s. 587–599.

- Zarębina M., 1985, *Próba statystycznej analizy słownictwa polszczyzny mówionej (synteza danych liczbowych)*, Wrocław–Warszawa–Kraków–Gdańsk–Łódź.
- Zgółka T., 2010, *Parametry odróżniające odmiany językowe*, w: M. Milewska-Stawina, E. Rogowska-Cybulska (red.), *Polskie języki. O językach zawodowych i środowiskowych*, Gdańsk, s. 34–40.
- Zgółkowa H., 1983, *Słownictwo współczesnej polszczyzny mówionej. Lista frekwencyjna i rangowa*, Poznań.
- Zgółkowa H., 1999, *Język subkultur młodzieżowych*, w: W. Pisarek (red.), *Polszczyzna 2000*, Kraków, s. 252–261.
- Żmigrodzki P., 2005, *Wprowadzenie do leksykografii polskiej*, Katowice.