



**UNIwersytet  
MIKOŁAJA KOPERNIKA  
W TORUNIU**

Wydział Nauk o Zdrowiu  
Collegium Medicum w Bydgoszczy

Monika Betyna

**Zaburzenia fluencji słownej u osób z dysfunkcjami poznawczymi  
w ujęciu neuropsychologicznym i lingwistycznym.**

**Rozprawa na stopień doktora nauk medycznych i nauk o zdrowiu  
w dyscyplinie nauki o zdrowiu**

**Promotor:**

**prof. dr hab. n. med. Alina Borkowska**

**BYDGOSZCZ 2022r**

Nr rozdziału	Tytuł rozdziału	strony
I	<b>WSTĘP</b>	5
II	<b>Fluencja słowna</b>	7
	2.1. Fluencja słowna- wprowadzenie	7
	2.2. Techniki badania fluencji werbalnej i niewerbalnej	9
	2.3. Adaptacja oraz implementacja testu fluencji słownej	13
	2.4. Funkcje poznawcze zaangażowane w wykonanie testu fluencji słownej	17
	2.5. Przegląd badań oraz zastosowanie analiz lingwistycznych u chorych z otępieniem	19
	2.6. Psychometryczne właściwości i ograniczenia narzędzia	22
	2.7. Wykonanie testu a zmienne demograficzne	24
III	<b>Lingwistyka korpusowa</b>	30
	3.1. Badania korpusowe	30
	3.2. Podstawowe zagadnienia w badaniach wspomaganym korpusem	33
	3.3. Podejścia korpusowe do analizy językowej	38
	3.3.1 Odmiany korpusów językowych	43
	3.3.1.a. Podejście wielowymiarowe	43
	3.3.1.b. Podejście socjolingwistyczne	44
	3.4. Listy słów, listy frekwencyjne, słowa klucze i konkordancje	48
IV	<b>Frazy kolokacje i wiązki leksykalne</b>	54
	4.1. Zakres frazeologii	54
	4.2. Frazeologia- definicje i założenia	55
	4.2.1. Jednostki wielowyrazowe	56
	4.2.2. Kolokacje	57
	4.3. Podejście do frazeologii- perspektywa polska	61
	4.4. Podejście do frazeologii - perspektywa rosyjska	64
	4.4.1 Klasyfikacja jednostek frazeologicznych Winogradowa	65
	4.5. Wiazki leksykalne w ramach kategorii frazeologicznych	67
	4.6. Ekserpowanie wiązek leksykalnych	73
V	<b>Cel badan i hipotezy badawcze</b>	84
	5.1. Cel badań	84
	5.2. Hipotezy badawcze	85
	5.3. Szczegółowe cele badawcze	85
VI	<b>Metodologia badań</b>	87
	6.1. Osoby badane: opis badanej grupy i zgoda komisji	87
	6.2. Metody badań	88

	6.2.1. Opis testu fluencji słownej	88
	6.2.2. Metoda analizy korpusowej	91
	6.3. Metody analizy statystycznej	94
VII	<b>Wyniki badań</b>	95
	7.1. Analiza spójności wewnętrznej stosowanych narzędzi	95
	7.2. Wyniki testów MMSE i fluencji słownej	97
	7.3. Analiza korelacji	101
	7.4. Analiza predykcji – wyniki analizy czynnikowej	103
	7.5. Analiza korpusowa	115
VIII	<b>Dyskusja</b>	121
	8.1. Wyniki testów poznawczych	121
	8.2. Wyniki analizy korpusowej	126
IX	<b>Wnioski</b>	132
X	<b>Streszczenie w języku polskim</b>	133
	10.1. Streszczenie	133
XI	<b>Streszczenie w języku angielskim</b>	135
	11.1. Summary	135
XII	<b>Bibliografia</b>	137
XIII	<b>Załączniki</b>	160
	13.1. Załącznik A- Spis rycin	160
	13.2. Załącznik B- Spis tabel	161

Dziękuję Pani Profesor Alinie Borkowskiej za ofiarowany czas,  
ogromną pomoc merytoryczną i nieocenione wsparcie  
oraz pełne zaangażowanie w proces powstawania pracy.

Niniejszą rozprawę doktorską dedykuję Wszystkim,  
którzy wspierają zarówno mój rozwój naukowy  
jak i zawodowy.

## ROZDZIAŁ I

### Wstęp

Język jest najważniejszym źródłem komunikacji międzyludzkiej. W związku z tym badania funkcji językowych są jednymi z najważniejszych w ocenie różnych grup klinicznych w szczególności u chorych z zaburzeniami funkcji poznawczych.

U pacjentów we wczesnych stadiach choroby otępiennej dochodzi do licznych zaburzeń funkcji werbalnych m.in. zaburzeń w zakresie fluencji słownej. Analiza fluencji słownej u chorych z różnymi dysfunkcjami, np. z uszkodzeniem prawej, lewej lub obu półkul mózgu [270] znalazła zastosowanie w diagnozie różnicowej zaburzeń neurologicznych u dzieci i młodzieży (Biechowska, Kaczmarek, Witkowska, Steinborn 2012), jak również w diagnostyce neuropsychologicznej [216], która od kilku lat jest systematycznie eksplorowana.

Niemniej jednak, niewiele badań nad zaburzeniami fluencji słownej chorych z zaburzeniami poznawczymi dotyczy kompleksowej interdyscyplinarnej analizy neuropsychologiczno lingwistycznej. W związku z tym postanowiono dokonać analizy specyfiki zaburzeń fluencji słownej u chorych we wczesnej fazie otępienia na etapie łagodnych zaburzeń poznawczych oraz u pacjentów z kliniczną manifestacją chorób otępiennych.

Celem niniejszego projektu badawczego jest zbadanie i przeanalizowanie fluencji słownej u osób z dysfunkcją poznawczą na poziomie łagodnych zaburzeń poznawczych i otępienia w tym chorych na chorobę Alzheimera. Analizie zostaną poddane wyniki badań testu fluencji słownej (MMSE) pacjentów wykazujących podobne zaburzenia. Za pomocą narzędzi do analizy danych korpusowych zbadane zostaną najczęściej występujące słowa wyekscerpowane z badań pacjentów za pomocą skali oceny stanu psychicznego Mini-Mental State Examination (MMSE).

Niniejsza dysertacja składa się z części teoretycznej i empirycznej. Część teoretyczna zawiera analizę lingwistyczną, w tym różne podejścia i klasyfikacje, komunikacja werbalna, test fluencji słownej. Opisane są także typologie testów do mierzenia fluencji słownej oraz zaprezentowane badania własne uwzględniające narzędzia językoznawstwa korpusowego.

Zaprezentowano analizę z zakresu słowotwórstwa, która może być pomocna np. przy rozpoznaniu użytych form fleksyjnych lub neologizmów przez osoby z zaburzeniami fluencji słownej oraz elementy analizy semantycznej, która pomaga ocenić zasadność użytego słownictwa.

Część empiryczna koncentruje się na zgromadzeniu i dogłębnym przeanalizowaniu materiału badawczego pozwalającego na stwierdzenie w jakim stopniu nasilenie zaburzeń poznawczych ma wpływ na stopień dysfunkcji fluencji słownej przy szczególnym uwzględnieniu wieku i płci pacjentów, wyekscerpowanie słów najczęściej występujących w języku polskim (z wykorzystaniem narzędzi korpusowych Narodowego Korpusu Języka Polskiego).

## ROZDZIAŁ II

### FLUENCJA SŁOWNNA

Rozdział drugi zawiera definicję fluencji, metodologię badań zagadnienia dotyczącą płynności posługiwania się językiem oraz opis stosowanych narzędzi badawczych. Testy fluencji słownej znajdują zastosowanie w badaniu dysfunkcji funkcji poznawczych u osób z wielu grup klinicznych, włączając w to pacjentów z zespołami otępiennymi. W literaturze przedmiotu podział testów fluencji obejmuje płynność niewerbalną (opartą na zdolnościach psychomotorycznych i wzrokowo-przestrzennych) oraz werbalną (dotyczącą ekspresyjności oraz zaburzeń mowy). W rozdziale drugim szczegółowo opisano zastosowanie testu fluencji słownej jako najprostszego i najbardziej użytecznego narzędzia badania procesów wydobywania pojęć. Test ma zastosowanie nie tylko w badaniach naukowych, ale również w praktyce klinicznej. Test fluencji słownej umożliwia diagnozę licznych zaburzeń funkcji poznawczych, obejmując zarówno płynność werbalną jak również szereg procesów związanych z uważnością, pamięcią oraz sprawnością funkcji wykonawczych.

#### **2.1. Fluencja słowna- wprowadzenie**

Fluencja jest określana jako sprawność służąca do planowania oraz wykonywania sekwencji niezautomatyzowanych czynności w określonym czasie (Łojek, Stańczak 2010: 91; [269], [228]. Sprawność tę odnosi się do szeregu działań psychicznych oraz neuropsychologicznych, takich jak płynność werbalna, płynność słowna, płynność niewerbalna, płynność konstrukcyjna, jak i płynność ruchowa [269], [271].

Fluencja werbalna jest to „konkretny termin medyczny, który określa umiejętność doboru słów zakodowanych i skonsolidowanych w wyniku nabywania doświadczenia językowego” [174], [269]. Do wykonania zadań z zakresu płynności czy też fluencji

werbalnej, zwanej również semantyczną, nieodzowna jest wiedza kategorialna. Kategorią semantyczną opisuje się minimum znaczeniowe, które znajduje się w ludzkiej świadomości, co buduje ich wiedzę referencjalną [298]. Typologia wywołuje skojarzenia semantyczne niezbędne do przywołania odpowiedniego nazewnictwa z leksykonu umysłowego. Sprawność wymieniania słów relewantnych do wyznaczonego kryterium możliwa jest dzięki sieci semantycznych łączących pojęcia w określone zbiory, mechanizmom aktywacji leksykalnej oraz sprawnościom podstawowym takim jak umiejętność rozumienia treści komunikatów werbalnych [144].

W przypadku osób starszych występujące zaburzenia rozumienia słów mogą prowadzić do zmniejszenia frekwencji użycia i występowania słów, a tym samym do znacznej prymitywizacji fluencji słownej [276]. Na jakość procesów językowych, jak również na stan pamięci semantycznej, wpływa sposób realizacji zadań fluencji słownej. Są to między innymi ilość wygenerowanych pojęć zgodnych z danym kryterium, jak również zgodność pojęcia z przyjętym kryterium, liczba klasterów / wiązek leksykalnych oraz ich zasób, charakter błędów, wliczając neologizmy oraz persewercje, jak również słowa nie przynależące do wskazanej kategorii [269], [97].

Wykaz trudności z aktualizacją konkretnych pojęć uznawane są za istotny symptom rozwijających się procesów otępiennych, zachodzących najczęściej we wczesnych stadiach choroby Alzheimera, kiedy to głównym kryterium rozpoznania dysfunkcji stanowią deficyty pamięci. Następstwem pogłębiającego się zaburzenia w procesie postępowania choroby są zakłócenia sfery językowej, przejawiające się w większości przypadków obecnością anomii, parafazji semantycznych oraz żargonu [205], [217], [80].

Wczesne jak i współczesne badania oraz ich dogłębne analizy zarówno z zakresu neuropsychologii jak również psychiatrii opisują i podkreślają, że istnieje wiele czynników które mają bezpośredni i pośredni wpływ na poziom wykonania zadań fluencji słownej oraz wskazują na wielopłaszczyznowość nie tylko mechanizmów psychologicznych ale również neuronalnych, które są ściśle powiązane z badanym procesem. Wykazano niewiele prac naukowych o badanej problematyce, które odnosiłyby się do dostosowań technik fluencji słownej do badań prowadzonych w polskich realiach. Wymieniony został szereg czynników znacząco wpływających na



proces realizacji płynności, do których zalicza się wiek, płeć jak również wykształcenie. Ponadto nacisk kładziony jest na możliwości poznawcze badanych osób, ich zaburzenia afektywne i osobowościowe, o czym świadczyć może podwyższony poziom lęku, nasilenia depresyjno kompulsywne oraz pojawiające się choroby współistniejące [269]. Badania naukowe przeprowadzone nad interpretacją psychologicznych mechanizmów fluencji, wskazują istotny związek zaburzeń fluencji z udziałem „pamięci semantycznej, epizodycznej, operacyjnej funkcji wykonawczych, uwagi, wyobraźni, szybkości psychomotorycznej oraz procesów metapoznawczych w realizacji fluencji słownej” [229].

## **2.2. Techniki badania fluencji werbalnej i niewerbalnej**

Testy fluencji słownej należą do typowych psychologicznych narzędzi stosowanych w diagnozie funkcji poznawczych. Osoba poddana badaniu ma za zadanie wymienianie w określonym czasie słów związanych z konkretnym hasłem. Czas ten najczęściej ograniczany jest do 60 sekund, jako optymalny wyznacznik sprawdzający zakres fluencji danego pacjenta. Do najczęściej wykorzystywanych technik badawczych należą testy na fluencję fonemiczną, zamiennie w literaturze nazywanej literową (litery: k, f lub inne) oraz fluencję semantyczną/kategorialną, jak również fluencję czasownikową, w której osoba badana jest proszona o wymienienie czasowników. Fluencja rzeczownikowa polega analogicznie na wymienieniu rzeczowników, natomiast fluencja asocjacyjna polega na wymienianiu wszystkich randomowych słów, przychodzących pacjentowi jako pierwsza myśl [269].

Najczęściej przeprowadzane badania dotyczą ilościowych parametrów realizacji zadań fluencji, analizy jakościowo-ilościowe są znacznie rzadziej przeprowadzane. Schwatz (2003) opisuje organizację pojęć przeprowadzoną na podstawie słów wygenerowanych w zadaniach fluencji literowej z zastosowaniem analizy skupień i analizy korespondencji [241]. Jak twierdzi Schwarz (2003) wygenerowane słowa tworzą układy spójne semantycznie, na przykład zwierzęta domowe, meble itd.

Literatura przedmiotu wskazuje że ekwiwalencja testów psychologicznych jest uwarunkowana w dużym stopniu od różnych czynników kulturowych. Implementacja testów fluencji polega na przystosowaniu czynności podjętych przy wdrażaniu narzędzia badawczego do zróżnicowanych okoliczności kulturowych. Jak powszechnie wiadomo, specjalistyczna wiedza językoznawcza jest nieodzownym elementem w celu przystosowania testów do cech systemowych konkretnego języka. Podczas analizy lingwistycznej uwzględnia się wiele aspektów, m.in. strukturę wyrazów, ich znaczenie oraz długość [141], a także ilość wyrazów na określoną literę, co ma ogromne znaczenie w przypadku testu fluencji słownej [155].

Chociaż wartość diagnostyczna zadań fluencji została uznana i testy uważa się za naukowe narzędzie badawcze, trudno jest jednoznacznie określić, jakie wskaźniki wykonania zadań fluencji słownej powinny być stosowane w klinicznej diagnozie różnicowej. Przy badaniu uwzględnia się nie tylko ilość poprawnie podanych słów ale również charakter popełnionych błędów (biorąc pod uwagę występowanie neologizmów, perseweracji, słów niezgodnych z przyjętym kryterium), czy też frekwencję występowania klasterów/wiązek leksykalnych (ang. clusters/ lexical bundles- definiowanych jako indywidualne pojęcie albo grupę pojęć, które są ściśle powiązane ze sobą nie tylko brzmieniowo, ale również na gruncie semantycznym) [135]. Wszystkie te czynniki świadczą o zasobach pamięci semantycznej oraz ścisłym związku funkcji poznawczych z wydolnością pamięci operacyjnej i przełączeń (ang. switching- definiowanych jako przejścia pomiędzy klasterami/ wiązkami leksykalnymi) [234].

Jakościowa analiza obejmuje badanie pojęć, grup tematycznych oraz frekwencji części mowy, czyli w zadaniach „otwartych”, ściśle związanych z płynnością asocjacyjną, sytuacją, w przypadku gdy osoba badana wymienia słowa nasuwające się na myśl. Obecnie proponowane są nowe rodzaje zadań na fluencję. Oprócz metod tradycyjnych, testów o charakterze nieafektywnym, polegających na wymienieniu słów rozpoczynających się określoną głoską bądź też należących do określonych kategorii semantycznych (imiona, części ciała, zwierzęta, przedmioty na konkretną literę lub głoskę itd.), do badań implementowane zostały kategorie afektywne wymagające od badanych wymienienia słów z kategorii słów przyjemnych i nieprzyjemnych [269]. Wyróżnić można również fluencję figuralną, inaczej nazywaną niewerbalną, bazującą na

przedstawieniu przez badanego jak największej ilości nowych wzorów. Ten rodzaj fluencji jest ściśle związany z ogólną płynnością motoryczną, jak również z umiejętnością wykonywania i planowania ruchów. Fluencja figuralna to kreatywny proces tworzenia nowych układów ruchowych (Suchy i in., 2010).

Uwzględnienie podanych wskaźników ilustruje, iż testy fluencji słownej zostają wykonywane na zróżnicowanym poziomie sprawności werbalnej przez osoby z dysfunkcjami mózgu [264], [241] w tym z zespołem otępiennym o różnym mechanizmie aktywności [146], Capitani i in., 2009). Dowiedziono znacznej różnorodności w realizacji zadań fluencji przez chorych z dysfunkcjami prawej półkuli mózgu w porównaniu do badanych z dysfunkcjami lewej półkuli mózgu. U osób z patologią lewej półkuli mózgu zaobserwowano znaczne obniżenie umiejętności produkowania słów rozpoczynających się na konkretną głoskę czy literę. Omówione kompetencje pozostają w kontraście względem wyników testów fluencji semantycznej. Uszkodzenie prawej półkuli mózgu wpływa na odwrotność powyższej sytuacji [99].

Badania wykazują, że sposób przetwarzania informacji oraz realizacja zadań fluencji są odmienne dla każdej z półkul mózgowych. Zadania lewej półkuli polegają między innymi na płynności literowej oraz klasyfikacji pojęć w klastery lub też związki leksykalne według wcześniej przyjętego kryterium. Fluencja semantyczna określa wiązki leksykalne semantyczne które nazywamy zgodnymi, lub brzmieniowe czyli niezgodnymi z typem zadania. Zadania prawej półkuli natomiast polegają między innymi na realizacji płynności semantycznej oraz warunkowaniu wyboru strategii semantycznych dotyczących fluencji literowej, co świadczy o wykorzystywaniu strategii heurystycznych. Udowodniono, że zaburzenia prawej półkuli wpływają istotnie na produkcję znacznie uboższej ilości słów z kategorii semantycznych niż u pacjentów z dysfunkcjami lewej półkuli [264], [99].

Wyżej wymienione badania pozwoliły na założenie hipotezy w której silne skojarzenia mogą służyć uruchomieniu funkcji lewopółkulowych co gwarantuje efektywność wykonania zadań fluencji przy każdym jej typie w fazie wstępnej. Cardebata (1996) potwierdza tę hipotezę w swoich badaniach, które wskazują istotne różnice we frekwencji wymienionych egzemplarzy przez badanych z dysfunkcjami lewej lub też prawej półkuli. Znaczące różnice pokazują się w ciągu pół minuty od rozpoczęcia

testu, a wyniki u badanych z zaburzeniami prawej półkuli mózgowej są zdecydowanie gorsze niż te u pacjentów z dysfunkcjami lewej półkuli mózgowej [55].

Schwartz i Baldo (2001) wykazały szereg różnic jakościowych w realizacji zadań na fluencję. Jak podaje Schwartz i Baldo (2001) prawa półkula inicjuje strategie ekscerpowania pojęć niezależnie od pochodzenia czy przynależności danego słowa do konkretnych kategorii. Pacjenci po uszkodzeniu prawej półkuli mózgu wykazują tendencje do podawania pojęć lub zwrotów powszechnie uznawanych za typowe, zaś ci z zaburzeniami lewej półkuli do podawania nietypowych pojęć lub zwrotów. Zróżnicowane zaangażowanie półkul mózgowych w proces wykonania testów fluencyjnych mają potwierdzenie w badaniach neuroobrazowych. W zależności od kategorii aktywizowane są różne obszary mózgu. Przykładem może być aktywacja tylnego obszaru mózgu, głównie prawej półkuli, który jest odpowiedzialny za produkcję słów w kategorii „zwierzęta”. Ponadto, aktywność obszaru ciemieniowego prawej półkuli mózgu jest największa podczas generowania słów w kategorii „narzędzia” ze względu na ich aspekty funkcjonalne [176], [292].

Najnowsze metody diagnostyczne zakładają stworzenie zadań z grupy fluencji afektywnej, które wsparłyby proces wytwarzania słów z kategorii np. „przyjemne” i „nieprzyjemne”. Osoby funkcjonujące w normie, w kontekście funkcjonowania OUN (ośrodkowego układu nerwowego), wykazują tendencję do generowania mniejszej ilości egzemplarzy w obrębie konkretnych kategorii. Warto nadmienić, iż stopień realizacji zadań na fluencję słowną jest współzależny pośrednio lub bezpośrednio od czynnika wieku oraz płci [272] jak również stanu emocjonalnego respondenta (Rossell, 2006; [40], [231]. Zdaniem badaczy oraz autorów publikacji naukowych zależności występujące pomiędzy prawą i lewą półkulą a realizacją zadań na fluencję potwierdzają, że lewa półkula bierze udział w realizacji zadań związanych z emocjami dodatnimi jak również funkcjami językowymi. Prawa półkula natomiast uczestniczy w realizacji emocji negatywnych oraz procesów niejęzykowych [13], [118].

W oparciu o inspirację metodologiczną wielowymiarowego skalowania słów pochodzących z zadań fluencji literowej, zaproponowano ilościowo-jakościową analizę słów wygenerowanych w zadaniach fluencji emocjonalnej w celu nie tylko charakterystyki organizacji pojęć, jak to uczynił zespół Schwartz i wsp. (2003), co opisu wewnętrznej struktury pojęć emocjonalnych [242]. Założono bowiem, iż analiza zestawów i znaczeń wygenerowanych słów pozwala nie tylko na rekonstrukcję relacji między słowami, ale także na rekonstrukcję wewnętrznej struktury pojęć. Struktura wewnętrzna pojęcia jest rozumiana jako treść pojęcia, indywidualny sposób jego rozumienia wyrażony wewnętrznym układem elementów danego pojęcia.

### **2.3. Adaptacja oraz implementacja testu fluencji słownej**

Jak pokazuje literatura przedmiotu, czynniki kulturowe warunkują przydatność testów psychologicznych. Implementacja testów fluencji wymaga szeregu czynności związanych z ich przystosowaniem do różnych kontekstów kulturowych. Jak wykazano, przystosowanie testów fluencji do specyficznych cech systemowych badanego języka stanowi wiele trudności. W takich przypadkach wymagana jest specjalistyczna wiedza językoznawcza, a nie tylko psychologiczna, co niejednokrotnie stanowi duży problem.

W procesie implementacji językowej narzędzia badawczego mnogość czynników brana jest pod rozwagę. Do najważniejszych elementów należą między innymi: struktura wyrazów, znaczenie wyrazów, długość klastera lub też wiązki leksykalnej jedno lub wieloelementowej [141], a także, co ma znaczący wpływ na fluencję, ilość egzemplarzy rozpoczynających się na określonej literę [155]. Można wyróżnić wiele testów angażujących analizę funkcji językowych. Do tych najbardziej znanych i użytecznych należy test fluencji słownej [259].

Istnieją empiryczne dowody na różnice w poziomie trudności testów fluencji słownej. Borkowska, Tomaszewska i Wilkość (2006), bazując na analizie słownika, sklasyfikowały litery C, F, P, A i S jako najłatwiejsze, natomiast litery L i R określone zostały jako najtrudniejsze w zastosowaniu testów fluencji. Należy nadmienić, iż wynik danego badania zawiera wszystkie wypowiedziane słowa, wliczając w to błędy takie jak intruzje (wyrazy wtrącone, które nie należą do danej kategorii) oraz persewercje.

Różnice w wersjach testu występują między obszarami językowymi jak również w obrębie obszarów językowych. Na przykład w Polsce w testach fluencji wykorzystywany jest wariant oryginalny, czyli słowa na określoną literę: F, A, S, jak również wariant podrzędny uwzględniający litery takie jak K, M, W, S, P. Jak wynika z przeprowadzonych badań [218], test fluencji słownej jest częścią szerokiego spektrum baterii neuropsychologicznej i z powodzeniem wykorzystuje się go podczas oceny kompetencji psychofizycznych przedstawicieli wielu kultur posługujących się różnymi językami.

W większości prac badawczych [129] uwzględniono oba warianty testu fluencji słownej: semantyczny i fonetyczny. Analiza szczegółowych badań potwierdziła zasadniczo wysoki stopień ich uniwersalności w sytuacji wykorzystywania kategorii ogólnych tj. "zwierzęta". Oznacza to stosunkowo niski poziom wpływu czynników kulturowych na tego typu kategorie słowne. Warto podkreślić, że określanie jakości poziomu fluencji słownej jest możliwe dzięki kategoryzacji semantycznej. Według badań Ponichtera-Kasprzykowskiej i Sobów (2014) za przykład może posłużyć język chiński, który w charakterystyce lingwistycznej nie uwzględnia wyodrębniania liter. Można założyć, że niektóre ze znaków mają przełożenie w konkretnych sylabach. Niemniej jednak, większość ze znaków ma wyodrębnione znaczenia lingwistyczne [56], [57], [159], [190].

Zagadnienie wariantu literowego zastosowanego w teście fluencji stanowi zdecydowanie większy problem. Jest to narzędzie, które występuje zarówno w wersji oryginalnej, jak i z zastosowaniem indywidualnego kryterium doboru liter wybranych na podstawie badań empirycznych, lub też te, których frekwencja zastosowania jest

przyjęta umownie za największą. Ponadto, nawet w obrębie jednego obszaru językowego, wykorzystywane są różne zestawy wariantu fonetycznego testu fluencji.

Jak wynika z analizowanych prac przez Ponichtera-Kasprzykowską oraz Sobów (2014), procedura adaptacji tak specjalistycznego narzędzia wymaga uwzględnienia wielu aspektów. Pod uwagę bierze się między innymi indywidualne cechy danego języka. Osoby hiszpańskojęzyczne wytwarzają mniej słów na litery F, A, S w porównaniu z osobami anglojęzycznymi. Sytuacja ma związek z różnorodnością frekwencji występowania danych wyrazów w językach hiszpańskim i angielskim [169]. Kempler i wsp. (1998) argumentują, iż Hiszpanie w odróżnieniu od Wietnamczyków oraz osób posługujących się językiem angielskim generują najmniejszą ilość nazw zwierząt. Różnica ta ma związek z długością słów, gdyż w języku hiszpańskim nominatywy zwierząt są znacznie dłuższe niż w języku wietnamskim [141].

Badania korelacyjne i confirmacyjne stanowią podstawę do analizy zagadnienia dotyczącego określenia stopnia wpływu czynników takich jak wiek, płeć, wykształcenie na poziom płynności słownej [5], [155], [179]. Część badań zawiera analizę porównawczą wyniku testu fluencji słownej przeprowadzonej na grupie osób zdrowych oraz grupie pacjentów cierpiących na stwardnienie rozsiane [83], łagodne zaburzenia poznawcze [193] czy też zaburzenia obsesyjno-kompulsywne [145]. Zakres czynników wpływających na zróżnicowanie wyników przeprowadzanych badań jest szeroki. Określenie zależności psychospołecznych w kontekście przeprowadzanej obserwacji dotyczącej poziomów fluencji jest konieczne dla stworzenia kryteriów stanowiących fundamentalną bazę w zestawach testów fluencji fonetycznej i słownej. Szereg przyczyn różnicuje poziom wykonanych zadań testujących poziom posługiwania się narzędziami komunikacyjnymi. Natomiast warto podkreślić, że istnieje niewiele przykładów, które potwierdzałyby, iż płeć w sposób znaczący różnicuje poziom wykonania testu fluencji słownej. Niewątpliwie należałoby opracować odrębne normy badawcze kategoryzujące pytania testowe w kontekście płci. Słuszne byłoby stworzyć klaster/ wiązki leksykalne odrębne dla kobiet i mężczyzn oraz dokonać ich analizy porównawczej [9]. Przewagę kompetencji fluencyjnych kobiet potwierdzają niektóre badania [168]. W opracowaniach innych badaczy nie odnajdziemy analizy porównawczej zagadnienia w kontekście

kompetencji z racji cech dominujących płci, bądź spotkamy się z konkluzją, iż nie wykryto tego rodzaju zależności [278].

Wielokrotnie badano wpływ wieku oraz wykształcenia na poziom wykonania testu fluencji, ze szczególnym uwzględnieniem różnic w zakresie kategorii fonetycznej i semantycznej (Koiovisto i in., 1992; [278], [110]. Jak wykazano, ilość generowanych słów jest zależna od wieku i maleje proporcjonalnie wraz z procesem starzenia [278], [168]. Niemniej jednak, istotny spadek obserwuje się w zakresie kategorii semantycznej [5], [179], [155]. Pozytywną korelację odnotowano również podczas badań nad płynnością słowną a poziomem wykształcenia badanych osób, ze szczególnym uwzględnieniem ilości lat edukacji, co nie miało wcześniej miejsca [12], [283], [225].

Polscy badacze podkreślają, że w polskiej literaturze przedmiotu brakuje danych empirycznych na temat wykorzystania wariantów testu fluencji słownej u osób z zespołami otępiennymi. Przeprowadzone badania w grupie pacjentów z chorobą Alzheimera wykazały znaczące deficyty językowe, które, obok zaburzeń pamięci należących do najbardziej rozpoznawalnych i dominujących czynników, mogą stanowić przesłanki demencji starczej [270], [41]. U pacjentów z chorobą Alzheimera, mających trudności z generowaniem nazw i zaburzoną fluencją słowną, postępuje zaburzenie mowy. Do podstawowych objawów zaburzeń mowy można zaliczyć trudności z nazywaniem rzeczy i przedmiotów oraz obniżona fluencja słowna, zwłaszcza semantyczna [121], [135]. Uwzględniając aspekty ilościowe, pacjenci cierpiący na otępienie typu alzheimerowskiego dysponują znacznie uboższym zasobem słów niż ich zdrowi rówieśnicy [191], [280], [286].

W polskiej publicystyce naukowej nie wskażemy wielu publikacji na temat adaptacji lingwistycznej testu fluencji słownej, mimo że to narzędzie łatwo dostępne oraz dające podstawy do pracy nad stworzeniem modelu działań mających na celu przeciwdziałanie regresowi języka w grupie osób z ryzykiem zaburzenia. Natomiast w licznych badaniach wykorzystuje się różne zestawy liter, lub też tylko pojedyncze litery, jak na przykład. M, P, K, Z, L [147] S, K, W [259] K, F [73], [74] lub K [267], [270], [274], [287], [296].



Uzasadnieniem na zastosowanie wyżej wymienionych zestawów okazuje się być frekwencja wyrazów na daną literę (np. K – wysoka frekwencja, F – niska frekwencja). Istotną część badaczy w testach fluencji stosuje oryginalny zestaw liter, w skład którego wchodzi: F, A, S [41], [134], [178], [220]. Nie bez znaczenia pozostaje fakt, że opracowano normy fluencyjne dla osób w wieku od 18 do 55 roku życia należących do poszczególnych grup: osoby zdrowe, cierpiące na schizofrenię, zaburzenia układu nerwowego oraz cukrzycę typu 1. i 2. (kategorie: „zwierzęta”, „ostre przedmioty”, litera „K”) [296]. Płynności semantycznej i literowej w różnych fazach dorosłości przyjrzeni się i opisali Daniluk i Szepietowska (2009a, 2009b) [73], [74].

Test fluencji słownej jest często stosowany jako narzędzie służące do pomiaru zaburzeń wśród osób cierpiących na depresję lub zaburzenia depresyjne [274], w przypadku zaburzeń afektywnych dwubiegunowych [73], [74] oraz w przypadku cukrzycy [273]. Opublikowano ponadto szczegółowe dane dotyczące wykonania testu fluencji przez dzieci w wieku 6-17 lat z problemami neurologicznymi takimi jak padaczka [259] oraz przez osoby cierpiące na niewydolność nerek [115].

#### **2.4. Funkcje poznawcze zaangażowane w wykonanie testu fluencji słownej**

Thurstone (1938) opisuje fluencję słowną jako wskaźnik wydajności umysłowej oraz zapis procesów lingwistycznych [275]. Fluencja słowna określana jest jako część składowa czynników mających istotny wpływ na podwaliny ludzkiej inteligencji. Chertkow oraz Bub (1990) oraz Ruff i wsp. (1997) w swoich badaniach wskazują, iż fluencja słowna wymaga dostępu do zasobów pamięci długotrwałej zarówno tej semantycznej jak i leksykalno- gramatycznej. „Fluencja słowna wymaga wysokich umiejętności ekscerpowania informacji oraz wysokiej efektywności funkcji wykonawczych, pamięci operacyjnej, koordynujących i monitorujących przebieg wykonania zadania” [58], [234].

Keil i wsp. (2002) stwierdzili stanowczo, że funkcje wykonawcze, występujące w sytuacjach nowych lub nietypowych, odpowiadają za efektywność i procedury badania pamięci, kontrolę jakości przebiegu zadania, w szczególności relewantności odpowiedzi z przyjętym kryterium, wyeliminowanie błędów, takich jak powtórzenia, neologizmy czy słowa w nieodpowiedniej kategorii [140]. Biorąc pod uwagę kryterium literowe zadania mają charakter bardziej nietypowy względem powszechnie przyjmowanego semantycznego sposobu magazynowania oraz wykorzystywania wiedzy w pamięci deklaratywnej [198].

Jako czynniki wpływające na poziom rozwiązania zadań należy uznać kryterium czasu, rodzaj wykonania zadania, długość przerw, gdyż te mogą stanowić wyznacznik sprawności przebiegu procesu poznawczego. Badania Ruffa i wsp. (1997), przeprowadzone na grupie 360 zdrowych pacjentów, wykazały istotę zestawienia danych badawczych testu FAS (COWAT) z wynikami uzyskanymi w testach Słownik i Powtarzanie Cyfr Skali Inteligencji Wechslera (WAIS-R), z próbą pamięci listy słów oraz testem uwagi. Weiss i wsp. (2006) zauważyli znaczący związek pomiędzy poziomem wykonania zadaniach w teście na fluencję literową i fluencję semantyczną wykonaną przez mężczyzn a wynikiem w teście Symbole Cyfr. W przypadku kobiet wyższe noty związane były ze zdecydowanie lepszym wynikiem uzyskanym w testach pamięci. Proporcjonalnie odwrotna sytuacja zachodzi w kontekście badań fluencyjnych przeprowadzanych wśród mężczyzn [288].

Jak zauważają Troyer i wsp. (1998) wskaźnikiem poziomu wykonania zadania w przypadku testu fluencji semantycznej są klastery, wiązki leksykalne oraz przełączenia [280]. W testach fluencji fonemicznej to przełączenia są głównym wskaźnikiem poziomu wykonania zadania. Zakłada się ścisły związek procesów pamięciowych ze zjawiskiem przełączania jako fundamentalnego indykatora sprawności procesów wykonawczych, wpływających istotnie na wielkość związków leksykalnych generowanych w testach fluencji [289]. Mayr (2002) zwrócił szczególną uwagę na rolę procesów wykonawczych w produkcji każdego egzemplarza, wskazując zjawisko przełączenia jako nie do końca relewantnego wskaźnika sprawności funkcji poznawczych.

## **2.5. Przegląd badań oraz zastosowanie analiz lingwistycznych u chorych z otępieniem**

Badania socjologiczne wykazują, że społeczeństwo Europy podlega nieustannemu procesowi starzenia się. Z każdym kolejnym rokiem wzrasta odsetek osób w starszych w społeczeństwie. Prognozy demograficzne wskazują, że w 2030 roku populacja osób powyżej 65. roku życia, licząca obecnie około 14%, będzie wynosiła 23% w krajach rozwiniętych, a odpowiednio 5% i 10% w krajach rozwijających się. Jak podaje Latkowski i Lukas (2004) w 2000 roku w Polsce odsetek osób w wieku powyżej 60. roku życia wynosił 16,7%, w porównaniu do innych krajów Europejskich w których ten odsetek wynosił 22,8% [158]. Ponadto z roku na rok liczba osób powyżej 65. roku życia nieustannie się zwiększa. W Polsce w 2005 roku wskaźnik ten wynosił 13,3% całego społeczeństwa, a prognozy przewidują, iż liczba ta wzrośnie do 31,3% do roku 2050. W 2015 roku odnotowano w Polsce ponad 6 000 pacjentów w wieku ponad 65 lat, co stanowi 15,8% całej populacji. Obserwując dynamikę starzenia się najstarszych grup wiekowych w krajach europejskich, prognozy zapowiadają, że przeciętna długość życia w Polsce stale rośnie a obecnie wynosi około 3,6 miesiąca dla mężczyzn oraz 2,1 miesiąca dla kobiet rocznie. Szacuje się, że w roku 2030 średnia długość życia dla kobiet wynosić będzie 83,6 lat, a dla mężczyzn 77,5 lat.

Zwiększająca się populacja osób w podeszłym wieku powoduje, iż liczba osób z demencją czy depresją wzrasta. Podobnie proporcjonalnie można przewidzieć rosnący wskaźnik ludności zmagającej się z chorobą Alzheimeta czy innymi zespołami zaburzeń, w tym stanami otępiennymi. Wzrost liczebności populacji osób starszych ściśle wiąże się z narastającymi problemami psychofizycznymi [200].

Badanie BERCAL zdaje się potwierdzać tezę dotyczącą wpływu wieku na postępowanie chorób wśród seniorów - analiza czynników ryzyka różnych form demencji zdecydowanie wskazuje na wiek jako czynnik w tym wypadku istotny, a nawet dominujący. Należy przewidywać, że wraz ze starzeniem się społeczeństw rozrosną się problemy związane z postępowaniem zaburzeń, chorób wśród seniorów [111]. Demencja starcza jest znacznie trudniejszym problemem do zdiagnozowania niż

przyczyny czy jej przebieg. Jak pokazuje Kiejna (2011) choroba Alzheimera należy do najczęstszych przyczyn zespołów otępiennych. Jak stwierdzono na podstawie ponad 60% zbadanych przypadków, choroba ta dotyka jedną na dziesięć osób powyżej 65. roku życia i niemal połowę osób w wieku powyżej 85 lat [143]. Na chorobę Alzheimera choruje na świecie od 12 do 15 mln osób, natomiast badania na gruncie polskim pokazują, że liczba ta przekracza 200 tys. osób [10], [41]. Zdiagnozowanie choroby Alzheimera pozwala na szybką implementację leczenia farmakologicznego, które okazuje się być efektywne tylko w spowalnianiu deterioracji poznawczej pacjenta, przede wszystkim w początkowej fazie choroby.

Międzynarodowa Statystyczna Klasyfikacja Chorób i Problemów Zdrowotnych ICD–10 klasyfikuje otępienie na podstawie występowania zespołu objawów spowodowanych chorobą mózgu, które musi trwać co najmniej 6 miesięcy. U pacjentów cierpiących na chorobę otępienną najczęściej zaburzone zostają funkcje poznawcze (funkcje korowe), do których zalicza się: myślenie, pamięć, rozumienie, orientację, liczenie, zdolności językowe oraz ogólne zdolności uczenia się i oceny. Takie zaburzenia mają najczęściej charakter przewlekły lub postępujący [40], [44].

Ćwiczenie funkcji poznawczych przez pacjentów cierpiących na chorobę mózgu jest naturalnym, poza farmakologicznym, sposobem zatrzymania postępowania choroby otępiennej. Spektrum zachowań osoby cierpiącej na demencję starczą uwarunkowane jest zmianami w obrębie mózgu, które mają charakter postępujący w sposób nieustanny i powodują zaburzenia pamięci oraz procesów poznawczych, co ma bezpośredni wpływ na zaburzenie zdolności postrzegania, rozumienia i opisywania otaczającej rzeczywistości [45].

Na Światowym Kongresie Choroby Alzheimera i Chorób Pokrewnych w Honolulu w 2010 r., określono nowe kryteria diagnozy zaburzenia zwanego otępiennym. Jako priorytetowy objaw jego występowania przyjęto zakłócenia w funkcjonowaniu zawodowym jak również zaburzenia w wykonywaniu czynności życia codziennego [188], [11]. Niestety niejednokrotnie zdarza się opóźnienie lub zakłócenie w diagnozowaniu zespołów otępiennych z powodu nieodróżniania zespołów otępiennych od zaburzeń depresyjnych lub zaburzeń lękowych [222].

Obok otępienia coraz częstszym zaburzeniem cywilizacyjnym XXI wieku, występującym wśród osób w podeszłym wieku, jest depresja. Według WHO, zaburzenie depresyjne jest najczęstszą chorobą psychiczną. Wskazuje się, iż roczna zachorowalność u osób w podeszłym wieku dotyka nawet 15% pacjentów, a 30% badanych z tej grupy umiera z powodu popełnienia samobójstwa w stanie skrajnej depresji. Daniluk i Borkowska (2008) podkreślają, iż depresja znacznie zwiększa ryzyko nasilenia się zaburzeń funkcji poznawczych, w szczególności u pacjentów z chorobą Alzheimera.

Spośród podstawowych objawów depresji w chorobie Alzheimera najczęściej występującymi są: obniżenie nastroju, wystąpienie epizodu depresji, utrata zainteresowań, zdecydowany spadek energii. Depresja bardzo często określana jest jako pierwszy objaw choroby Alzheimera. Dane epidemiologiczne wskazują, iż zaburzenia nastroju najczęściej dotyczą łagodnej i umiarkowanej postaci choroby [77]. Występowanie epizodu depresyjnego u pacjentów w ciągu ich całego życia wystąpiło u ponad 50% badanych z chorobą Alzheimera, jak wykazuje Zubenko i wsp. (2003). Ponadto podaje się, że około 33% pacjentów doświadcza epizodu depresyjnego w trakcie następowania zaburzeń otępiennych [299]. Według kryteriów DSM-IV, w celu rozpoznania depresji, niezbędne jest nasilenie się u pacjenta co najmniej dwóch objawów przez większą część dnia, w okresie nie krótszym niż dwa tygodnie [43].

Interesujący jest związek pomiędzy depresją a zaburzeniami funkcji poznawczych występującymi w chorobie Alzheimera. W przypadku postępującej depresji dysfunkcje te mogą osiągnąć nawet poziom pseudo otępienia depresyjnego. Ponadto, zaburzenia te dotyczą większości obszarów kognitywnych, w których często obserwowane jest spowolnienie psychomotoryczne, zdecydowane spowolnienie procesów zaburzenia uwagi, przetwarzania informacji, zaburzenia funkcji przestrzennych oraz wykonawczych [42].

Większość badań prowadzonych w tym obszarze zajmuje się m. in. zwiększeniem skuteczności rozpoznania zaburzeń otępiennych [299], [42], [222] wyspecjalizowanym postępowaniem terapeutycznym, w szczególności przy porównaniu zaburzeń takich jak otępienie czołowo-skroniowe a otępienie w chorobie Alzheimera [42] jak również poprawą jakości leczenia pacjentów [238] oraz przy szczególnym uwzględnieniu

frekwencji występowania zaburzeń otępienia i depresji, zidentyfikowaniem czynników, takich jak: wiek, płeć, wykształcenie oraz warunki mieszkaniowe, które mogą mieć istotny wpływ na zachorowalność pacjentów na otępienie i depresję [53].

Niestety, literatura przedmiotu nie opisuje odrębnie wyników badań dotyczących zastosowań analiz lingwistycznych u chorych z dysfunkcjami otępiennymi. Do tej pory nie zgromadzono wystarczających danych na temat analiz językowych przy pomocy narzędzi korpusowych u osób ze zmianami otępiennymi w kontekście wykonywania testów werbalnych. W związku z powyższym niniejsza dysertacja skupia się na analizie lingwistyczno-korpusowej u chorych z deficytami otępiennymi, w tym cierpiącymi na chorobę Alzheimera.

## **2.6. Psychometryczne właściwości i ograniczenia narzędzia**

Należy podkreślić, iż brakuje ściśle określonych norm i przeprowadzonych badań naukowych na temat psychometrycznego charakteru testu fluencji w języku polskim. Klinicyści i badacze bazują głównie na własnym doświadczeniu lub też zagranicznych badaniach naukowych. Ruff i wsp. (1996) wskazują na wysoką zgodność wewnętrzną testu fluencji słownej, normalizację testu literowego wersji F, A, S jak również innych alternatywnych wersji testu literowego [233].

Nieprawidłowości w zakresie wdrażania narzędzia badawczego mogą wynikać z różnych stanów patologicznych, o czym nadmieniono w sekcji 2.1., należą do nich choroba Alzheimera, choroba Parkinsona, wszelkiego rodzaju stany otępienne wywołane wirusem HIV, schizofrenia, jak również depresja. Najbardziej popularne są dwie wersje testu fluencji słownej: z zastosowaniem kryterium formalnego (literowego, ang. phonemic fluency) oraz kryterium semantycznego (kategorialnego, ang. semantic fluency).

Należy nadmienić, iż frekwencja występowania słów na daną literę w różnych językach może być zróżnicowana, co może mieć istotny wpływ na rozbieżności w przebiegu i ocenie przeprowadzonego testu. W polskiej literaturze przedmiotu można wyróżnić testy literowe w różnych wersjach, na przykład F, A, S, lub też K, W, M, P i S [260], [269]. Praktyka kliniczna wskazuje na nieprawidłowy wynik testu, gdy pacjent wygeneruje poniżej piętnastu poprawnych słów na określoną literę w trakcie 60 sekund. Aby wynik badania był relewantny wspomniane wcześniej zmienne demograficzne powinny być również uwzględnione. Przyjmuje się wydłużony czas testu fluencji słownej jeżeli u pacjentów występuje spowolnienie psychoruchowe w obrazie klinicznym choroby [216], [44].

Rzetelność, lub też wskaźnik rzetelności (*ricc*), jako zgodność czasowa (ang. *test-retest*) narzędzia badawczego, jest odpowiednia i relewantna gdy współczynnik korelacji wynosi ponad 0,70. Taka sytuacja powinna mieć miejsce w testach o krótkich, kilkumiesięcznych przedziałach czasowych, jak i kilkuletnich badaniach. Pojawić się może efekt uczenia się przejawiający się niewielkim wzrostem produkcji słów w retestach, co może wystąpić nawet po kilkunastu miesiącach [233], [14], [165], [260].

Jak twierdzi Dikmen i wsp. (1999) wykorzystanie innego zestawu liter podczas retestu zmniejsza prawdopodobny efekt nauki, jednakże nie wyklucza go całkowicie [78]. Basso i wsp. (1999) wskazują, że prostota wykonania testu fluencji słownej a nie efekt uczenia się czy błąd pomiaru statystycznego wpływa istotnie na różnice w poziomie wykonania testu.

Ross i wsp. (2005) w swoich kilku tygodniowych badaniach wykazali niewielki współczynnik korelacji wewnątrzklasowej dla ilości występowania klasterów, związków leksykalnych oraz przełączeń, przy założeniu że średnia wielkość egzemplarzy nie uległa zmianie [230]. Przy każdorazowym zastosowaniu testu zauważyć można różnorodność stosowanych strategii podczas ekscerpowania słów. Troyer i wsp. (1997) podają, iż standaryzacja instrukcji testów pozwoliłaby na obiektywną ocenę efektywności danego narzędzia, co adekwatnie pozwoliłoby ocenić zdolność efektywnego wykorzystania informacji. Różnorodność wspomnianych klasterów oraz związków leksykalnych stanowi cenne źródło informacji na temat stosowanych strategii [279].

## 2.7. Wykonanie testu vs zmienne demograficzne

Wyniki badań wskazują, że istnieje istotna zależność między wykształceniem oraz wiekiem badanego a poziomem wykonania zadania. Do czynników wpływających na wynik testu zaliczają się między innymi iloraz inteligencji, przynależność etniczna, dwujęzyczność oraz stan afektywny [141], [95], [278], [50].

Wyniki badań przeprowadzonych przez Bäckman i wsp. (2004) oraz Kave i wsp. (2005) wykazują, że poziom zdolności wykonania zadań fluencji fonemicznej wzrasta, by osiągnąć apogeum w okresie do czwartej dekady życia, a następnie, wraz z wiekiem, obniża się [139]. Natomiast wyniki polskich badań przeprowadzonych przez Borkowską i wsp. (2006), Szepietowską i wsp. (2011) oraz Piskunowicz i wsp. (2013) ujawniają znaczący wpływ wieku pacjenta na spadek liczby poprawnych odpowiedzi w teście określającym poziom fluencji werbalnej, jak również jakość i ilość przełączeń.

W jednym z przeprowadzonych przez Sumerall i wsp. (1997) badań nad fluencją literową wykazano, iż błędy, do których zaliczyć można powtórzenia, nazwy własne i słowa posiadające ten sam rdzeń, okazały się być zjawiskiem powszechnie spotykanym u badanych w wieku powyżej 70. roku życia [265]. Co więcej, badania potwierdziły, iż w danej grupie badawczej błędy typu słowa na nieprawidłową literę pojawiają się zdecydowanie rzadziej. Troyer i wsp. (1997) wykazali, że w testach fluencji semantycznej zaobserwować można spadek liczby przełączeń, na co wpływa wiek pacjentów. Natomiast wielkość klasterów, czy też wiązek leksykalnych, pozostaje czynnikiem niezależnym od wieku.

Wniosek ostatni jest równoznaczny z konkluzją badawczą Szepietowskiej i Gawdy (2011) oraz Mayra i Kliegla (2000), którzy twierdzą, że wiek ma istotny wpływ na proces wykonawczy, a nie wartość semantyczną testu fluencji słownej [181]. Mayr (2000) podkreśla jednocześnie, że na wynik testu fluencji semantycznej u osób w podeszłym wieku może mieć wpływ fakt, iż badani potrzebują zdecydowanie więcej czasu na generowanie słów niż osoby w średnim wieku. Liczba przełączeń związana jest również



z wielkością klasterów, im większa wiązka leksykalna, tym mniejsza liczba przełączeń [1], [180].

Stwierdzono, że na wynik testu fluencji słownej wpływa poziom wykształcenia. Szepietowska i Gawda (2011) wykazały proporcjonalną zależność pomiędzy liczbą poprawnych odpowiedzi oraz przełączeń u zdrowych pacjentów a wykształceniem respondentów. Badania osób w podeszłym wieku wykazały, że na poziom fluencji literowej ma wpływ poziom wykształcenia. Sytuacja dotyczy również pacjentów cierpiących na otępienie w stopniu lekkim. Fluencja semantyczna u badanych nie zależy natomiast od poziomu wyedukowania osoby badanej, ani jego wieku oraz płci. Na podstawie analizy badań uznano test za istotny wskaźnik postępowania procesu otępiennego [71].

Płeć nie jest wskazywana jako istotny indyktor powodujący różnice w poziomie wykonania testu fluencji słownej. Jak wykazali Weiss i wsp. (2006) kobiety charakteryzują się większą tendencją do przełączeń w teście fluencji literowej, w przeciwieństwie do mężczyzn, którzy generują zdecydowanie większą ilość wiązków leksykalnych [288]. Podobnie badania na gruncie polskim dowodzą wpływu czynników takich jak wiek, płeć i wykształcenie na poziom i jakość wykonywania zadań fluencji. Przy badaniu diagnostycznym mężczyzn z wykształceniem wyższym ilość poprawnych odpowiedzi wzrasta wraz z wiekiem, u kobiet natomiast maleje. Podobna tendencja nie została zaobserwowana w symultanicznych badaniach osób z wykształceniem średnim [268].

## KONKLUZJE

Niniejszy rozdział traktuje o fluencji słownej, jej etymologii, podziale, zastosowaniu oraz rodzajach testów weryfikujących fluencję pacjenta. Fluencja słowna określana jest jako umiejętność wymieniania słów zgodnie z określonym kryterium (są to między innymi kategorie takie jak zwierzęta, przedmioty ostre lub też słowa rozpoczynające się na określoną literę) i ograniczona jest limitem czasowym (60 sekund). Metoda uznawana jest za uniwersalny wskaźnik sprawności nie tylko językowej ale i artykulacyjnej. Psycholingwistyka zakłada, że zadania fluencyjne stwarzają niepowtarzalną możliwość analizy struktury badanego języka, w szczególności jego poziomu semantycznego tworzącego klaster umysłowy. Metoda kreśli drogę infiltracji człowieka do treści i znaczenia słowa. Frekwencja występowania Frekwencja występowania zarówno semantycznych jak i fonemicznych subkategorii, takich jak klastery czy też wiązki leksykalne należy do najczęściej uwzględnianych znaczników zasobów językowych i pamięci. Dane subkategorie semantyczne i fonemiczne konstruowane są przez wymieniane słowa, istotnym indykatorem natomiast sprawności funkcji wykonawczych jest frekwencja przełączeń [279].

Sposób realizacji fluencji słownej stanowi źródło danych o psychologicznej sieci połączeń wiedzy semantycznej [44] oraz o jej zaburzeniach w przypadku patologii mózgu i licznych dysfunkcji neuropsychologicznych. Ralph (2014) wskazuje, że odrębne kategorie są wykorzystywane do współtworzenia wiedzy semantycznej [224]. Inne badania sugerują segregację pojęć według przyjętych cech obiektów, percepcyjnych i funkcjonalnych. Marques i wsp. (2008) za przykład podają porównanie pojęć naturalnych i sztucznych [177]. Hillis i Caramazza (1995) do segregacji pojęć według przyjętych cech obiektów zaliczają uwzględnienie ich cech gramatycznych, porównanie rzeczowników i czasowników [120]. Współcześni badacze powołują się na powyższe ustalenia i traktują elementarne techniki realizacji fluencji słownej z najwyższą uwagą dążąc do rozpowszechnienia i poszerzenia obszaru jej wpływu na rozwój nauki z różnych dziedzin.

Badacze opisują zależności sięgające „od założeń sięgających do niezależności systemów neuronalnych stanowiących podstawę organizacji klas semantycznych lub cech obiektów zaliczanych do tych klas” [224] po współczesne podejście prezentowane przez Pulvermüller (2013), gdzie podkreślany jest „dynamiczny i rozproszony charakter aktywacji mózgowej z zaznaczeniem roli obszarów wyspecjalizowanych w przetwarzaniu podstawowych cech obiektów oraz struktur organizujących te dane” [221].

Fluencja słowna bazuje głównie na pamięci semantycznej. Generowanie podczas testów pojęć ze względu na ich liczne powiązania semantyczne może mieć charakter automatyczny. Przeszukiwanie wielu egzemplarzy wymusza udziału dodatkowych procesów, takich jak selekcji pożądaných pozycji, produkowania subkategorii, czy występowania przełączeń [138].

W rozdziale opisano techniki fluencji werbalnej i niewerbalnej, które należą do typowych psychologicznych narzędzi stosowanych w diagnozie funkcji poznawczych. Podaje się również ogólny opis narzędzia - testu fluencji słownej. Test ten występuje w wielu wersjach. Thurstone stworzył pierwowzór testu fluencji słownej w pierwszej połowie XX wieku, który nazwany został Thurstone's Word Fluency Test. Charakterystyka testu pisemnego określa metodę pozyskiwania materiału badawczego. Polega ona na generowaniu jak największej ilości słów na konkretną literę w ściśle określonym czasie. Test ten używany był do oceny stopnia produktywności umysłowej badanych z dysfunkcjami ośrodkowego układu nerwowego.

Zakłada się, że do wykonania testu włączone są podstawowe funkcje takie jak mowa, pamięć oraz wszelkie procesy wykonawcze. Test stanowi ważny wskaźnik funkcjonowania między innymi płatów czołowych i skroniowych. Literatura anglojęzyczna wykazuje szereg badań na ten temat. Niestety istnieje deficyt jasno określonych norm w literaturze polskiej dotyczących sposobu ewaluacji wyników tekstów fluencyjnych. Sytuacja ograniczenia dostępu do opisu badań na polskim gruncie skutkuje zmniejszeniem wartości narzędzia badawczego i wymaga zdecydowanej ostrożności oraz skrupulatności w ocenie wyników przeprowadzonych badań.

Wielokrotnie podkreśla się, iż funkcje poznawcze odgrywają znaczącą rolę podczas wykonania testu fluencji słownej. Jak uważają Troyer i wsp. (1998), w przypadku fluencji semantycznej wskaźnikiem poziomu wykonania zadania są zarówno klaster czy też wiązki leksykalne, jak i przetęczenia, które są głównie spotykane w przypadku fluencji fonemicznej. Niewątpliwie przetęczenie jest wskaźnikiem służącym do weryfikacji sprawności procesów wykonawczych, gdzie wielkość utworzonych związków wyrazowych związana jest przede wszystkim z procesami pamięci [280]. Jednak Mayr (2002) podkreślił rolę procesów wykonawczych w generowaniu każdego słowa, a nie tylko tych rozpoczynających nowy klaster lub też wiązka leksykalna i uważał, że frekwencja występowania przetęczeń może nie być adekwatnym wskaźnikiem sprawności badanych funkcji.

Przeprowadzone, przeanalizowane i opisane badania wskazują na bardzo silną korelację pomiędzy wiekiem i poziomem wykształcenia, a stopniem poprawności wykonania testu. Do innych czynników wpływających na poziom wykonania testu fluencji należą pochodzenie etniczne, iloraz inteligencji, wielojęzyczność czy też stan afektywny [141], [95], [278], [50]. Wyniki badań przeprowadzonych przez Bäckman i wsp. (2004) oraz Kave i wsp. (2005) udowadniają ciągły wzrost poprawności wykonania zadań fluencji fonemicznej w okresie do czwartej dekady życia, co następnie z wiekiem ulega zdecydowanemu obniżeniu [139]. Natomiast wyniki badań na gruncie polskim przeprowadzonych przez Borkowską i wsp. (2006), Szepietowską i wsp. (2011) oraz Piskunowicz i wsp. (2013) zauważają wpływ podeszłego wieku pacjenta na znaczący spadek ilości poprawnych odpowiedzi, jak również ich przetęczeń.

W polskiej literaturze przedmiotu brakuje ściśle określonych norm i publikacji naukowych na temat psychometrycznych właściwości testu fluencji słownej, dlatego klinicyści i badacze odwołują się zwykle do zagranicznych badań normalizacyjnych lub doświadczenia klinicznego. Nieprawidłowości, które mogą wystąpić podczas wykonania testu, mogą być wynikiem wielu zaburzeń chorobowych, takich jak choroba Alzheimera, choroba Parkinsona, różnego rodzaju procesy otępienne wywołane wirusem HIV, schizofrenia, jak również stany głęboko depresyjne jak i epizody depresyjne. Do najbardziej powszechnych wersji testu fluencji słownej należą te posługujące się kryterium formalnym oraz kryterium semantycznym.

Czynniki kulturowe mają bezpośredni wpływ na diagnostyczną relewancję testów psychologicznych. Należy zgodzić się z postawioną tezą mając na uwadze liczne publikacje omawiające narzędzia, sposób ich wykorzystania oraz wyniki badań. Wdrażanie testów fluencji wymaga czynności związanych z dostosowaniem do różnych kulturowych zawiłości badanego języka. Ponadto, przystosowanie testów do cech systemowych danego języka wymaga specjalistycznej wiedzy językoznawczej, co niejednokrotnie stanowi, trudną do pokonania, zaporę lingwistyczną.

Starzenie się społeczeństwa jest powszechnym procesem dotyczącym zarówno kraje rozwijające się jak i te rozwinięte. Dzięki wzrostowi efektywności opieki zdrowotnej, średnia długość życia ciągle się wydłuża. Większa liczba coraz starszych seniorów oznacza wzrost występowania chorób związanych z wiekiem. W związku z tym rośnie ryzyko zachorowań wśród osób w podeszłym wieku. Różne kompilacje chorób otępiennych, w tym choroba Alzheimlera, czy choroba Parkinsona, które charakteryzują badaną grupę są powszechnie występującymi zaburzeniami. Dlatego wzrasta liczba badań wykonywanych w celu zwiększenia skuteczności rozpoznania zaburzeń otępiennych, poprawy jakości leczenia pacjentów, wykrycia bardziej efektywnych postępowań terapeutycznych oraz identyfikacji czynników wpływających na zachorowalność na depresję, chorobę Alzheimlera oraz inne postępujące deficyty otępienne.

Dotąd w literaturze przedmiotu nie znaleziono danych dotyczących analiz językowych przy użyciu narzędzi korpusowych dla osób z dysfunkcjami otępiennymi w kontekście wykonywania testów werbalnych. Zagadnienie w świetle powyższego sformułowania stanowi niszę badawczą dla niniejszej dysertacji. Niniejsza suma badawcza skupia się na wykładni lingwistyczno-korpusowej u chorych z deficytami otępiennymi, w tym chorymi na chorobę Alzheimlera.

## ROZDZIAŁ III

### LINGWISTYKA KORPUSOWA

W niniejszym rozdziale uwaga zostaje skierowana na badania korpusowe, w kontekście których, w procesie diagnozy zagadnienia problemowego, odbiorca poznaje definicję korpusu. W rozdziale trzecim zostaną przedstawione podstawowe zagadnienia dotyczące badań wspomaganych korpusem oraz dwa główne podejścia do językoznawstwa korpusowego. Pierwsze podejście, wielowymiarowe, opiera się na przetwarzaniu danych korpusowych mających na celu identyfikację wiązek leksykalnych i klasterów. Podejście socjolingwistyczne koncentruje się na analizie danych, klasyfikowanych w zależności od kontekstu i typu interakcji. Ostatnim aspektem omawianym w rozdziale, który będzie przedmiotem badań empirycznych niniejszej rozprawy, jest krótkie przedstawienie podstawowych narzędzi badawczych, które można wykorzystać do analizy korpusów: list słów, list frekwencyjnych, słów kluczowych i konkordancji.

#### **3.1. Badania korpusowe**

Badania językowe oparte na korpusach zakładają, że istnieje ogromna baza danych obliczeniowych zwana korpusami. Wybitni lingwiści korpusowi podają jego różne definicje, w zależności od ich naukowego punktu widzenia, celu badania, czy podejścia metodologicznego. Wczesną definicję terminu - korpus podaje John Sinclair (1991), o którym mówi się, że jest ojcem nowoczesnej lingwistyki korpusowej, natomiast Sinclair i Mauranen (2005) podają zaktualizowaną definicję korpusu:

[...] a collection of pieces of language text in electronic form, selected according to external criteria to represent, as far as possible, a language or language variety as a source of data for linguistic research [251].

Leech (1992) wskazuje, że

a corpus is a *helluva* lot of text, stored on a computer...computer corpora are rarely haphazard collections of textual material: they are generally assembled to be (informally speaking) *representative* of some language or text type [161].

Francis (1964) opisuje korpus jako:

[...] a collection of texts assumed to be representative of a given language, dialect, or other subset of a language to be used for linguistic analysis [91].

Z bardziej współczesnej perspektywy McEnery (2006) definiuje:

A corpus is a collection of (1) machine-readable (2) authentic texts (including transcripts of spoken data) which is (3) sampled to be (4) representative of a particular language or language variety [186].

Gries (2016) podkreśla, że

[...] corpus is a collection of files that contain text or/and transcribed speech that is supposed to be representative and balanced for a certain language, variety, register, or dialect [108].

Definicje te podkreślają, że korpus to baza danych, która może istnieć zarówno w formie pisemnej jak i ustnej, stworzona według wcześniej ustalonych kryteriów. McEnery i wsp. (2006) wskazują, że korpus zawiera teksty autentyczne lub, jak określa to Sinclair (1991), teksty występujące naturalnie, co oznacza, że są one gromadzone bez ingerencji językoznawców [245]. Teksty powstają w rzeczywistych sytuacjach, a następnie są gromadzone w formie kwestionariuszy lub wywiadów. Gries (2016) podkreśla, że korpus składa się nie tylko z dyskursu pisanego lub mówionego, ale zawiera również dodatkowe adnotacje zawierające części mowy oraz cechy semantyczne i morfologiczne [108].

Inny aspekt dotyczy idei reprezentatywności, o której wspomina definicja Sinclaira (1991). Język lub odmiana językowa (zgodnie z definicją Sinclaira 1991), pobrana w formie korpusu, działa jako reprezentacja określonej grupy ludzi. Oznacza to, że cechy wspólne, wyprowadzone z analizy korpusowej, są reprezentatywne dla wszystkich respondentów. Według Leech (2011) ustalenie, czy korpus jest reprezentatywny, jest niezwykle trudne. Sinclair i Mauranen (2005) wspierają tę kwestię, dodając, że bez względu na to, jak starannie kompilowane są dane, cechy języka w próbie mogą nieznacznie różnić się od języka źródłowego. W tym przypadku korpus może nie być w pełni reprezentatywny, jak stwierdzili Reppen i Simpson (2002) oraz Sinclair i Mauranen (2005) [227]. Aby osiągnąć cel maksymalnej reprezentatywności, teksty muszą być odpowiednio dobrane zgodnie z rodzajami tekstów i dziedziną. Wielkość korpusu również ma znaczenie, Sinclair i Mauranen (2005) stwierdzają, że im większy korpus, tym bardziej reprezentatywne jego cechy można znaleźć.

Warto zauważyć, że korpus jest odczytywalny maszynowo [184], co umożliwia zaawansowana technologia. Internet jest praktycznym narzędziem, jak zauważył Baker (2009), umożliwiającym gromadzenie różnych tekstów w formie elektronicznej. Dzięki temu analiza wskazanych tekstów nie nastręcza żadnych trudności. Dodatkowo znaczenie ma format pliku. Narzędzia do analizy korpusu nie będą przetwarzać tekstów w innym formacie niż tekstowy. W rezultacie, jeśli artykuły są pobierane z Internetu w formacie PDF i WORD, muszą zostać przekonwertowane w celu dalszej analizy.

Wreszcie Francis i wsp. (1996) definiując korpus podkreśla jego cel, którym jest analiza dużego zbioru danych z językowego punktu widzenia [92]. Najbardziej przydatnymi programami do analizy korpusów, które są również wykorzystywane w części empirycznej niniejszej rozprawy, są WordSmith [244], Sketch Engine [149] oraz AntConc [4]. Programy te umożliwiają szybkie wyliczanie i zliczanie wystąpień elementów leksykalnych, klasyfikowanie kontekstu w konkordancjach oraz badanie kontekstu tych elementów leksykalnych. Badanie korpusu pokazuje, jakie wzorce wiążą się z cechami leksykalnymi i gramatycznymi charakterystycznymi dla danego dyskursu.



### 3.2. Podstawowe zagadnienia w badaniach wspomaganych korpusem

Dziedzina nauki, jaką jest analiza lingwistyczna, w ostatnich latach dynamicznie się rozwija, dokonując przełomowych badań dzięki ewolucji globalnej, ze szczególnym uwzględnieniem coraz bardziej zaawansowanych technologicznie urządzeń i oprogramowania, które pozwalają naukowcom, przy ich kierunkowym wykorzystywaniu, analizować szeroki zakres tekstów, dostępnych za pośrednictwem uporządkowanych zbiorów danych, znanych jako korpusy. Do najbardziej podstawowych narzędzi stosowanych w analizie korpusów należą narzędzia do obsługi tekstu. Wybrane zostaną omówione w dalszej części pracy - znane jako lista słów; liczba słów; analiza słów kluczowych lub konkordancje; analiza klasterów i wiązek leksykalnych. Częstotliwość występowania słów w tekście poddanym badaniu jest istotnym zagadnieniem, niemniej jednak jeśli chodzi o artykuły naukowe, konieczna jest również interpretacja jakościowa. Zasadniczo celem badania jest nie tylko przedstawienie surowych danych, ale także odnalezienie wzorców rządzących językiem, które będą konsekwentnie stosowane w celu tworzenia narzędzi dydaktycznych i materiałów dla na przykład przyszłych lekarzy [26].

W ciągu ostatnich czterdziestu lat korpusy elektroniczne okazały się ważnym narzędziem dla językoznawców. W niniejszym rozdziale zostaną przybliżone zagadnienia natury lingwistyki korpusowej, ogólny wkład korpusów w naukę oraz trzy obszary analizy językowej wyróżnione przez McEnery i Gabrielatos (2006):

- ogólny opis języka,
- analiza leksykalno-gramatyczna,
- nauczanie języka obcego.

W ciągu ostatnich kilkudziesięciu lat liczni autorzy, tacy jak Biber i wsp. (1998), Kennedy (1998), McEnery i Wilson (2001) wyrazili przekonanie, że lingwistyka korpusowa nie jest tylko jedną teorią językową, ale raczej metodologią [26], [142], [185]. I odwrotnie, inni badacze twierdzą, że „lingwistyka korpusowa to coś więcej niż tylko metodologia, to raczej nowe przedsięwzięcie i nowe filozoficzne podejście do tematu” [161].

Zasadniczo korpusy stanowią obszerne tablice tekstu przechowywanego cyfrowo w bazach danych, które można wdrażać za pomocą oprogramowania znanego jako narzędzia do analizy korpusów. Jednak sposób, w jaki zainteresowani wchodzą w interakcję z korpusami, różni się od sposobu, w jaki wchodzić oni w interakcję z drukowanymi wersjami tekstów. Z reguły drukowane teksty są sprawdzane sekwencyjnie od początku do końca. Z kolei w trakcie badania korpusu bada się zwykle drobne fragmenty treści, np. poszczególne jednostki leksyki lub poszczególne linijki tekstu, umożliwiając w ten sposób jednoczesną analizę wielu odcinków [161].

Komputery i oprogramowanie komputerowe ułatwiły zbieranie, przechowywanie i analizę dużych ilości danych, przyczyniając się do skuteczności i szybkości metody [161]. Dodatkowo, ponieważ proces analizy został zautomatyzowany, łatwiej jest powtórzyć i zweryfikować wyniki. Leech (1992) stwierdza, że korzystanie z komputerów:

daje nam możliwość zrozumienia i wyjaśnienia zawartości [...] korpusów w sposób, o jakim nie śniło się w erze przedobliczeniowej językoznawstwa korpusowego [161].

Jednym z najbardziej znaczących ulepszeń, dokonanych przy wykorzystaniu metody badania języka opartego na korpusie, jest możliwość skoncentrowania uwagi wielu użytkowników języka na półautomatycznych fragmentach wyrazów [249]. Dlatego wielu badaczy, zamiast skupiać się wyłącznie na pojedynczych słowach, dostrzegło znaczenie kolokacji i wielowyrazowych jednostek znaczeniowych [104], [123], [293]. Wybrane fragmenty i jednostki wielowyrazowe zostaną omówione w rozdziale czwartym.

Badania oparte na korpusach wykazały, że podczas gdy dla native speakerów takie wielowyrzowe jednostki nie stanowią problemu, inaczej problem wygląda w kontekście omawianej sprawności językowej u obcokrajowców [76], [123], [196]. O ile zrozumienie takich jednostek nie powinno sprawiać trudności, o tyle używanie ich do tworzenia naturalnie brzmiących tekstów może pozostać problematyczne. Takie fragmenty leksykalne są niezbędne zarówno w komunikacji ustnej, jak i pisemnej. Aby tworzyć wypowiedzi, które będą rodzime, użytkownik języka powinien osiągnąć określony poziom znajomości często używanych kombinacji słów i ich właściwego kontekstu, zarówno w języku ogólnym jak i ESP, co znacznie poprawia styl wypowiedzi. Powyższe założenie należy uwzględnić przy projektowaniu metod i programów nauczania [70], [124], [201], [293].

Metody korpusowe ukierunkowane zostały bardziej na posługiwanie się językiem niż na jego strukturze, co wpłynęło na poszerzenie i wzbogacenie wielu dziedzin lingwistycznych, takie jak analiza dyskursu, leksykologia, dialektologia, semantyka, badania wariacji językowych, pragmatyka, socjolingwistyka, przekład, językoznawstwo historyczne, stylistyka i psycholingwistyka. Jak donoszą Bower i Pearson (2002) „[...] małe wyspecjalizowane korpusy, zawierające teksty w określonym stylu, mogą być przydatnym źródłem podczas nauki pisania w określonym stylu lub tworzenia określonego rodzaju tekstu” [47].

Na uwagę zasługuje fakt, że podejście korpusowe jest skorelowane z pracami z zakresu językoznawstwa korpusowego prowadzonymi przez Leech (1991) oraz Halliday (1994), które traktują priorytetowo „informacje uzyskiwane przez wzorce syntaktyczne, a nie leksykalne” (Tognini- Bonelli, 2001: 81). Z kolei podejście korpusowe koresponduje z badaniem Sinclaira (1991) dotyczącym kontekstualnej teorii znaczenia, które podkreśla ogromny wpływ leksykologii na badania korpusowe. Rozróżnienie między stanowiskiem Leecha (1991) czy Hallidaya (1994) w kwestii podejścia korpusowego a poglądem Sinclaira (1991) obejmuje nie tylko kwestie metodologiczne, ale także teoretyczne [113].

Korpusy to bazy danych, które umożliwiają analizę słów w różnych kontekstach, takich jak struktury składni, funkcje i konteksty, w których mogą być nadreprezentowane lub niedostatecznie reprezentowane. Ponieważ analiza korpusu opiera się na częstotliwości słowa lub frazy (ang. the usage-based approach), umożliwia badanie słów od bardziej szczegółowych do bardziej ogólnych (ang. the bottom-up approach) [277].

W dziedzinie dydaktyki Bączkowska (2005a, 2005b) doszła do podobnych ustaleń i przedstawiła techniki probabilistyczne i restrykcyjne. Technika probabilistyczna wymaga, aby wiele próbek z większego zbioru było analizowanych w celu określenia ogólnych zasad zalecających użycie badanej struktury [15], [16]. Podejście restrykcyjne natomiast jest podejściem odwrotnym. Najpierw formułowana jest hipoteza dotycząca języka, a następnie przeprowadzana jest dogłębna analiza licznych przykładów w celu ich potwierdzenia. Dlatego podejście restrykcyjne wykazuje analizę odgórną ang. a top-down analysis (od ogółu do szczegółu), podczas gdy podejście probabilistyczne wykazuje podejście oddolne ang. a bottom-up approach (od szczegółu do ogółu) [15], [16].

Sinclair (1991) stwierdza jednak, że należy ufać tekstowi i pozwolić, aby dane z korpusu same sugerowały wzorce, co z kolei zapewnia nowy wgląd w język, którego wcześniej nie zauważano. Jest to rdzeń podejścia opartego na korpusie ang. the core of the corpus-driven approach [277].

Stubbs (1996) wyraźnie stwierdza, że najważniejszą techniką, umożliwiającą osiągnięcie zadowalających wyników analizy językowej, jest połączenie dwóch przywołanych podejść - the corpus-based approach oraz the corpus-driven approach [261]. „Koncepcja lingwistyki opartej na danych musi stawić czoła klasycznemu problemowi, że nie ma czegoś takiego jak czysta indukcja, którą deklaruje [...] – Językoznawca zawsze podchodzi do danych z hipotezami i przecuciami, jakkolwiek niejasnymi” (1996: 25), co ukazuje istotę podejścia korpusowego (the corpus-based approach) w pracy językoznawcy.

Wprowadzenie korpusów elektronicznych silnie wpłynęło na metody tworzenia słowników. Językoznawcy korpusowi zdali sobie sprawę z nowych źródeł, które można wykorzystać w poszukiwaniu dowodów językowych, a także oryginalnych metod analizy, prezentacji i zarządzania danymi leksykograficznymi.

Pierwszym znaczącym projektem korzystającym z korpusów elektronicznych jest słownik języka angielskiego Collins COBUILD [248]. Został opracowany i oparty na Międzynarodowej Bazie Języków Uniwersytetu Collins-Birmingham (COBUILD), która rozwinęła się w obszerny zasób o nazwie Bank of English. Oryginalny korpus COBUILD powstał w latach 80. w Birmingham z Johnem Sinclairem na czele. Pomógł w stworzeniu wielu słowników i tekstów skupiających się na regułach gramatycznych, w tym tak znanych publikacji jak Collins COBUILD Dictionary i Collins COBUILD Grammar Patterns [92].

W konsekwencji starsza metoda, mniej efektywna i mniej obiektywna, została zastąpiona metodą korpusową jako główną metodą gromadzenia danych leksykograficznych. W dzisiejszych czasach wszyscy liczący się wydawcy, przygotowując słowniki czy materiały źródłowe, polegają na korpusach słownych zawierających miliony haseł.

Leksykografowie mają dostęp do licznych źródeł autentycznych i obiektywnych danych językowych, występujących w postaci korpusów językowych, w celu zliczania częstotliwości słów: MI - analizy statystycznej; podobieństw leksyko-gramatycznych w obrębie danego dyskursu, jak również odniesień do zależności między innymi czynnikami:

„Lingwistyka korpusowa polegała na analizie słów w kontekście, w celu zademonstrowania jego użycia w określonym kontekście, co wiązało się ze zmianą formatu słownika, aby umożliwić transfer tej kontekstowej wiedzy z powrotem do użytkownika. Doprowadziło to do rewolucji w słownikach jedno- i dwujęzycznych. Podejście kontekstowe doprowadza obecnie do zmiany terminologii, ponieważ w takim rzeczywistym użyciu sztywność pojęciowa już nie obowiązuje” [290].

W leksykografii opartej na korpusie analitycy nie polegają już na swojej intuicji, ale na wzorcach znalezionych w korpusie. Dzięki temu słowniki są bardziej aktualne kontekstowo, a definicje prezentowanych słów leksykalnych i gramatycznych oparte są na języku autentycznym. Ponadto wykorzystanie korpusów do celów leksykograficznych zaowocowało stworzeniem bardziej wydajnych sposobów dostępu, przechowywania i odwoływania się do materiałów źródłowych, a także opracowaniem nowych narzędzi, które mają pomóc w zarządzaniu ogromnymi ilościami danych [291]. W konsekwencji wprowadzenie korpusów oraz narzędzi korpusowych umożliwiło leksykografom bardziej szczegółowe badanie wzorców językowych, co z kolei stworzyło możliwość tworzenia rozbudowanych haseł słownikowych.

### **3.3. Podejścia korpusowe do analizy językowej**

Twórcy nowej szkoły myślenia w badaniach językoznawczych, znanej jako językoznawstwo korpusowe, rozpoczęli swoją pracę w zasadzie niezauważeni. Rudolph Quirk ujawnił swój plan stworzenia korpusu brytyjskiego angielskiego w mowie i piśmie w 1959 roku, który stał się znany jako Survey of English Usage (SEU) [223]. Niedługo potem dwóch lingwistów, Nelson Francis i Hendry Kučera, zebrali zespół specjalistów z Brown University, biegłych w zawiłościach korpusów. W rezultacie Brown Corpus został uznany za „standardową próbkę” drukowanego amerykańskiego angielskiego „do użytku z komputerami cyfrowymi” [91].

W ten sposób rozpoczęła się druga era językoznawstwa korpusowego, z korpusami SEU i Brown na czele. Choć korpusy SEU nie były korpusami komputerowymi, epokę zdecydowanie zdominowała obecność komputera, pomimo jego głównej wady, tj. braku koncentracji na danych mówionych (ze względu na problemy z transkrypcją i wprowadzaniem danych). SEU nie został jednak pominięty w tej kwestii ze względu na równomierny podział na dane mówione i pisane, a także charakterystycznych dla jego „nieobliczalnych” danych mowy.

Jan Svartvik wykorzystał mocne strony korpusów Brown i SEU z pomocą swojego zespołu w Lund. W 1975 roku jego grupa badawcza Survey of Spoken English zaczęła renderować nieskryptowane teksty mówione z korpusu SEU do odczytu maszynowego. Zwieńczeniem ich wysiłków był London-Lund Corpus (LLC), który do dziś nie ma sobie równych pod względem bogactwa zasobów wykorzystywanych w procesie nauki mówionego języka angielskiego (głównie brytyjskiego) (czas opracowania: 1959–1990). Tworzenie i analizowanie korpusów niezmiennie cieszy się dużym zainteresowaniem. Oprócz MICASE (czas kompilacji: 1997–2002), CALLHOME (czas kompilacji: 1997), CALLFRIEND (czas kompilacji: 1996-1997), British National Corpus (czas kompilacji: 1991-1994) i Corpus of Contemporary American English (czas opracowania: 1990-obecnie), istnieją różne korpusy, takie jak The Cambridge English Corpus (pobrane: 2019), corpus of Academic Written and Spoken English (CAWSE) (pobrane: 2020), English as a Lingua Franca w Academic Settings (ELFA) (dostęp: 2020), akademicki korpus ELF (dostęp: 2020), International Corpus of Learner English (ICLE) (dostęp: 2020), Louvain International Database of Spoken English Interlanguage (LINDSEI) (dostęp: 2020), Trinity Lancaster Corpus, jeden z największych korpusów języka angielskiego w mowie L2, Vienna-Oxford International Corpus of English (VOICE) (czas opracowania: 2001-2009).

Chociaż niniejsza rozprawa koncentruje się na korpusach brytyjskich, innym naukowcem, o którym należy wspomnieć, jest Mark Davies, profesor lingwistyki na Uniwersytecie Brigham Younga w USA. Jego dorobek naukowy jest znaczący. Profesor opracował liczne bazy danych służące do analizy ogromnych ilości materiałów lingwistycznych pochodzących z różnych źródeł. Współpracował z naukowcami, którzy zaznaczyli swoją obecność w świecie nauki jako badacze, teoretycy oraz promotorzy nowych rozwiązań z zakresu lingwistyki stosowanej, m.in. z Douglasem Biberem i Dee Gardner. Do jego największych osiągnięć należą:

- Corpus of LDS General Conference talks,
- Corpus del Español (2002),
- Register variation in Spanish (2004),
- Frequency dictionary of Spanish (2005),
- Corpus do Português (2006),

- Frequency dictionary of Portuguese (2007),
- Corpus of Contemporary American English (COCA) (2008),
- Frequency dictionary of American English (2009),
- English word frequency, collocates, and n-grams (2010),
- Corpus of Historical American English (COHA) (2010),
- Google Books Corpus (2011),
- GloWbE: Corpus of Global Web-Based English (2013),
- Downloadable full-text corpus data (2014),
- Wikipedia Corpus (2015),
- Hansard Corpus (British Parliament) (2015),
- Billion word extensions to the Spanish and Portuguese corpora (2015-2007),
- CORE corpus (Corpus of Online Registers of English) (2016),
- Corpus of US Supreme Court Opinions (2017),
- iWeb Corpus (2018),
- TV and Movie corpora (2019),
- The Corpus of Contemporary American English (COCA) (2020)<sup>1</sup>.

Nowe spojrzenie i poszukiwanie dróg rozwoju dla językoznawstwa korpusowego (oparte na korpusach SEU i Browna) wpłynęły na dynamikę ich ewolucji od 1961 roku do tego stopnia, że te stały się główną metodą przetwarzania języka naturalnego. Nie spełniło się marzenie amerykańskich lingwistów strukturalnych o w pełni wystarczającym korpusie, nie mniej jednak wartość korpusu nie jest już lekceważona – jest on wykorzystywany jako źródło danych i poligon doświadczalny dla językowych eksperymentów. Było jednak oczywiste, że korpus komputerowy oferuje nową metodologię budowania systemów przetwarzania języka naturalnego.

Na podstawie odniesienia do opisanych w literaturze przedmiotu badań, McEnery i Wilson (2001) twierdzą, że rzeczowniki, przymiotniki i przysłówki wymagają większych korpusów ze względu na ich mniejszą częstotliwość występowania w porównaniu ze słowami gramatycznymi [185]. Biber (1990) przeciwnie, wskazał, że mniejsze korpusy

---

<sup>1</sup> <http://davies-linguistics.byu.edu/personal/> [dostępność: 20.03.2020]



są więcej niż wystarczające do badania takich standardowych cech języka jak elementy gramatyczne.

Korpus wykorzystywany jest nie tylko w leksykografii i analizie języka, ale także jako pomoc w nauczaniu języka obcego. Jak stwierdzili Bower i Pearson (2002), „można zastosować podejście oparte na korpusie, aby nauczyć się, jak tworzyć określony rodzaj tekstu lub pisać w określonym stylu” [47]. Istnieje kilka sposobów wykorzystania korpusu do celów edukacyjnych. Po pierwsze, lingwistyka korpusowa oferuje autentyczną bazę danych, która pozwala sprawdzić rozumienie języka i dowiedzieć się więcej o zawiłościach językowych, których wcześniej nie dostrzegano (O Keeffe i in., 2007).

Co więcej, korpusy były również wykorzystywane do tworzenia podręczników, które umożliwiały uczniom języka obcego kontakt z bardziej naturalnym językiem niż ten stosowany w tradycyjnych podręcznikach. Wiele dominujących wydawnictw ma dostęp do pokaźnych korpusów, które wykorzystywane są do tworzenia materiałów korpusowych, takich jak podręczniki do nauki języka obcego, podręczniki do gramatyki i słownictwa, przygotowanie do testów i wielu innych materiałów dla uczniów i nauczycieli.

Badania w dziedzinie ESP (English for Specific Purposes- Angielski do Celów Specjalistycznych) często analizują różnice między cechami językowymi ogromnie zdefiniowanej kategorii, takiej jak proza akademicka lub spotkania biznesowe, a tymi, które pojawiają się w innych ogólnych i szerzej stosowanych rejestrach. Uzyskana konfiguracja pokazuje, że pomimo wewnętrznej zmienności można wyróżnić poszczególne wzorce zmienności. Często badanie poszczególnych słów koncentruje się na potrzebach uczniów EAP (English for Academic Purposes- Angielski do Celów Akademickich).

Ocena Byrda i Coxheada (2000), która wygenerowała Academic Word List (AWL), w znacznym stopniu ukształtowała zasoby leksykalne dla EAP. Różne cechy gramatyczne prozy akademickiej stanowią jeden z segmentów opisanych w Longman Grammar of Spoken and Written English (1999) Bibera i wsp. (1999) [70]. Zawiera ponad 300 opracowań o różnych cechach gramatycznych, w dużej mierze zestawiających prozę

akademicką z komunikacją werbalną, literaturą, a także prasą. Składają się one z prostych obliczeń częstotliwości, takich jak leksykalne klaster, ale można też znaleźć bardziej złożone cechy gramatyczne i leksyko-gramatyczne, takie jak zmienne warunkujące wybór między stroną czynną i bierną. Rozważono również kompilację cech gramatycznych. Biber i wsp. (2011) oraz Biber i Gray (2010) przedstawili kilka analiz, które wydają się być największym wyzwaniem dla praktyków EAP. Badania dotyczyły dyskursu akademickiego i komunikacji werbalnej w zakresie wykorzystania opracowań z mowy zależnej (np. zdania dopełniające lub zdania względne) wobec skompresowanych modyfikacji za pomocą fraz (np. fraz rzeczownikowych przyimkowych lub fraz przyimkowych). Badacze odkryli, że proza akademicka stosuje znacznie więcej modyfikatorów frazowych, podczas gdy komunikacja werbalna wykazuje znacznie więcej opracowań klauzulowych [33].

Proponując stosunkowo odmienne stanowisko w sprawie licznych rejestrów uniwersyteckich, Biber (2006) porównuje dziesięć rejestrów pisanych i mówionych na uniwersytetach (na przykład nauczanie w salach lekcyjnych lub podręcznikach) pod kątem różnych cech leksyko-gramatycznych. Należy przeanalizować różne rozległe różnice dotyczące uczestników, celu i sposobu. Liczne badania leksykalne badają leksykę dla poszczególnych przedmiotów, na przykład inżynierii [192], biznesu [195], czy nawet wykładów z teologii [164]. Segmenty wielowyrazowe zostały również przeanalizowane w różnych dyscyplinach. W jednym z nich Cortes (2004) zestawiał wiązki leksykalne w czasopismach biologicznych i historycznych, odkrywając kilka istotnych różnic w ich strukturze i funkcjach. Na przykład biologia pozwala na większą różnorodność swoich kategorii strukturalnych i wykorzystuje więcej wiązek postaw, aby wyrazić brak pewności.

### **3.3.1. Odmiany korpusów językowych**

Istnieją dwie podstawowe brytyjskie tradycje językoznawstwa korpusowego [90], na których koncentruje się niniejsza rozprawa. Pierwsza pochodzi z wczesnych lat 60. XX wieku i jest zakorzeniona w podejściu probabilistycznym stosowanym w projekcie Survey of English Usage (SEU) prowadzonym przez Jana Svartvika z University College London z pomocą Quirk, Leech i Greenbaum. Leech później przeniósł się do Lancaster i stanął na czele kompilacji korpusu Lancaster-Oslo- Bergen (LOB) na tych samych zasadach co SEU. Drugie podejście, znane jako podejście neo-Firthian, było popierane przez Halliday (2004), Sinclair (2004), Stubbs (2004), Hunston (2002) i Hoey (2005).

Pierwsze podejście wykorzystuje dane korpusowe podczas badania kategorii gramatycznych, przede wszystkim tych, które stanowią podstawę kompleksowej gramatyki języka angielskiego [262], [223]. Należy jednak podkreślić, że to drugie podejście zbudowało ramy teoretyczne na kontekstualnej teorii znaczenia i kontekstu jednostek frazeologicznych Firtha (1957). Uważa się, że znaczenie zależy od słów, z którymi współwystępują (więcej szczegółów w rozdziale 4). Dane korpusowe pozwalają naukowcom weryfikować idee Firtha poprzez analizę dużych zbiorów danych (dlatego też to podejście nosi nazwę „neofirthian”) i dostarczają dowodu na znaczenie informacji kontekstowych (współwystępowanie słów zazwyczaj łączących się z badanym słowem, np. bazy kolokacyjnej) w opisie znaczenia słowa [88].

#### **3.3.1.a. Podejście wielowymiarowe**

Badania korpusowe nie stałyby się tym, czym są, gdyby nie znaczący wpływ prac Bibera. Zarówno teoretyczne jak i metodologiczne podstawy podejścia badacza do medycyny można znaleźć w jego przełomowej pracy „Wariacje w mowie i piśmie” [23], a ostatnio to właśnie podejście obliczeniowe do analizy korpusu zostało również wykorzystane w identyfikacji wiązek leksykalnych.

Opisując obliczeniową, wielowymiarową analizę statystyczną stosowaną w badaniu zmienności językowej w celu określenia typologii tekstu, Biber (1988) wyodrębnił 67 cech charakterystycznych dla języka. Obejmują one zarówno cechy składniowe (znaczniki czasu), jak i leksykalne (różne typy czasowników), a także kategorie obejmujące przyimki i czasowniki modalne. Dalej [29] rozszerzył początkową liczbę 67 do 129 cech językowych, które obejmują znaczną ilość nowych dodatków leksyko-gramatycznych (np. zdania typu „that-clauses”).

Wyniki badań Bibera pozwoliły ustalić osiem typologii tekstów opartych na wspólnych cechach składniowych i leksykalnych. Następnie zinterpretował te wzorce współwystępowania, aby znaleźć podstawowe funkcje – sytuacyjne, społeczne i poznawcze. Wyprowadził więc swoją klasyfikację z wewnętrznych kryteriów językowych, co oznacza, że teksty z tego samego rejestru (np. konwersacja, listy, podręczniki) nie należą do tego samego typu tekstów, chyba że mają wspólne cechy językowe.

Charakter wiązek leksykalnych powoduje, że identyfikuje się je nie intuicyjnie, ale statystycznie i jako takie częściej są niekompletne strukturalnie. Nawiązując do przykładu z języka angielskiego przedimek „the” jest wiązką, która składa się z niepełnej frazy rzeczownikowej i zawiera początek osadzonej frazy „of”.

### **3.3.1.b. Podejście socjolingwistyczne**

Wartą odnotowania cechą wspólną wyżej opisanych podejść naukowych do procesu badawczego struktur językowych w celu ustalenia i weryfikacji korpusu jest to, że łączy je powiązanie socjolingwistyczne, które ujawnia się podczas wykorzystania danych korpusowych. We wczesnej lingwistyce korpusowej pojawiły się zauważalne uprzedzenia dotyczące analizy tekstów pisanych i dyskursu mówionego. Trafną ilustracją takiego wczesnego korpusu, zawierającego 5 milionów terminów, jest Korpus the Cambridge and Nottingham Corpus of Discourse in English, zwanym również CANCODE.

Korpus CANCODE został założony w 1996 roku przez Cartera i McCarthy i zajmuje się nieformalnym mówionym językiem angielskim. Co więcej, jego konstrukcja i późniejsze analizy były punktem zwrotnym w erze skomputeryzowanych korpusów mówionych – przed 1990r. korpusy mówione koncentrowały się jedynie na gromadzeniu danych demograficznych, definiowaniu kategorii mowy i kierowaniu do określonych grup osób. Chociaż CANCODE wykorzystuje również dane demograficzne i jest do tego przeznaczony, jego główna zasada organizacji obejmuje cztery szerokie konteksty gatunkowe języka mówionego: profesjonalny, intymny, towarzyski i transakcyjny.

Obecnie coraz częściej podkreśla się tę predyspozycję do rejestrowanego i wzorcowego języka angielskiego w językoznawstwie korpusowym. Ze względu na ogromne nakłady pracy, jakie trzeba wykonać przy tworzeniu kompilacji, korpusy języka mówionego, wraz z ich prozodią, nie zostały jeszcze dostatecznie zobrazowane. Niemniej jednak, w przeciwieństwie do korpusów języka mówionego skomponowanych w przeszłości, te, które powstały niedawno, takie jak British National Corpus czy Longman Corpus of Spoken American English, zapewniają dostęp do mowy o szerokim zakresie społecznym.

Ponadto już funkcjonujące korporacje utworzyły międzynarodowe agencje dystrybucji korpusów. Na przykład Lee (2010) w swojej analizie odwołuje się do: Linguistic Data Consortium (LDA), International Corpus of Modern and Medieval English (IAME), Open Language Archives Community oraz Oxford Text Archive (OTA) [160]. Lokalizacje te mogą być niezmiernie cenne, ponieważ nadają się do realizacji indywidualnych projektów badawczych.

Ewentualnie, jeśli żaden dostępny korpus nie zawiera odpowiednich danych wymaganych do analizy, wówczas nowy korpus może być budowany przez badacza lub zespół współpracujący. McEnery i wsp. (2006: 71) nazywają to korpusem DIY. Budowa korpusu DIY polega na zebraniu odpowiednich, autentycznych fragmentów pisemnych, a następnie zarządzaniu nimi jako materiałem elektronicznym z wykorzystaniem procedur projektowych. Badacz decyduje się operować na czystych lub surowych fragmentach korpusu lub opatrzyć teksty adnotacjami.

Anotacja korpusu stanowi jednoznaczne uzupełnienie odpowiednich danych językowych w korpusie elektronicznym za pomocą znaczników. Jak stwierdza Leech (1997), jest to sposób na wzmocnienie lub uzupełnienie użyteczności korpusu i może być korzystny w zależności od celu badawczego. Zgromadzone dane są następnie uporządkowane pod kątem rodzaju kontekstu i interakcji, przy czym te pierwsze odzwierciedlają relacje międzyludzkie między rozmówcami i ich chęć przekazywania informacji na odpowiednim poziomie. To, wraz z otoczeniem, w którym powstało nagranie, kwalifikuje dane do określonej kategorii.

Najbardziej fragmentaryczny typ anotacji w korpusie to taki, który przyporządkowuje poszczególnym słowom określone części mowy. Jednak prowadzenie takiej anotacji ręcznie, w stosunkowo dużych korpusach, może stać się niezwykle męczące i żmudne. Do tego wszystkiego pomyślnie można skorzystać z predefiniowanych tagerów, które potrafią wykonać to zadanie, osiągając niemal stuprocentową dokładność, w krótszym czasie.

English CLAWS tagset wersja 7, jak określono w Sketch Engine, jest dostępny w angielskich korpusach. Jest to automatyczny system anotacji słów (ang. The Constituent Likelihood Automatic Word-tagging System). Został opracowany przez University Center for Computer Corpus Research on Language na Uniwersytecie Lancaster. Mówi się, że CLAWS niezmiennie osiąga ponad dziewięćdziesiąt siedem procent dokładności, przy czym konkretna dokładność zależy od rodzaju tekstu. Stąd, chociaż mogą wystąpić pewne błędy w tagowaniu, z reguły pozostają one marginalne i łatwe do wyeliminowania. Tagowanie jest możliwe na poziomie pól semantycznych, a także morfologii, składni, fonologii oraz pragmatyki.

Instytut Podstaw Informatyki PAN, będący kordynatorem Instytutu Języka Polskiego PAN, Wydawnictwa Naukowego PWN oraz Zakładu Językoznawstwa Komputerowego i Korpusowego Uniwersytetu Łódzkiego stworzył Narodowy Korpus Języka Polskiego (NKJP). Projekt Narodowego Korpusu Języka Polskiego został zrealizowany jako program badawczy Ministerstwa Nauki i Szkolnictwa Wyższego.

Wszystkie cztery wyżej wymienione instytucje wspólnie zbudowały tak zwany korpus referencyjny języka polskiego zawierający ponad półtora miliarda słów. Wyszukiwarki korpusowe, które są ogólnodostępne pozwalają przeszukiwać zasoby korpusowe z użyciem zaawansowanych narzędzi, których zadaniem jest nie tylko uwzględnienie odmian i rodzajów wyrazów polskich, ale również analiz budowy zdań w języku polskim.

Narodowy Korpus Języka Polskiego zawiera klasykę literatury polskiej, prasę codzienną oraz teksty specjalistyczne w tym nagrania rozmów, czyli teksty autentyczne. NKJP charakteryzuje się zróżnicowaną tematyką i gatunkowością tekstów, które zawiera, a jednocześnie reprezentacje rozmówców obu płci, w różnym wieku i z różnych regionów, oraz grup etnicznych. Wszystkie te aspekty są szczególnie istotne w ocenie wiarygodności korpusu jako narzędzia badawczego.

Narodowy Korpus Języka Polskiego jest relewantną bazą materiałową dla nowego Wielkiego Słownika Języka Polskiego stworzonego w ramach projektu badawczo-rozwojowego w Instytucie Języka Polskiego PAN. Większość tekstów zebranych w korpusie jest wykorzystywana w projekcie Korpus Polsko-Rosyjski afiliowanym na Wydziale Polonistyki Uniwersytetu Warszawskiego we współpracy z Uniwersytetem Pedagogicznym w Ufie oraz Narodowym Korpusem Języka Rosyjskiego.<sup>2</sup>

Dla każdego kompilatora etapem końcowym, następującym po udostępnieniu odpowiedniego korpusu, jest wykonanie analiz lingwistycznych z wykorzystaniem wybranego zestawu narzędzi korpusowych. Poniżej wymieniono kilka dostępnych. Analityk może dokonywać wszelkiego rodzaju ocen ilościowych i jakościowych tekstów występujących w korpusie.

---

<sup>2</sup> Narodowy Korpus Języka Polskiego  
<http://www.nkjp.pl> (dostępność: 26.09.2021r)

### 3.4. Listy słów, listy frekwencyjne, słowa kluczowe i konkordancje

W tej sekcji zostaną zaprezentowane podstawowe narzędzia zarządzania stosowane w językoznawstwie korpusowym. Aby rozpocząć analizę opartą na korpusie, należy stworzyć listę wszystkich terminów w odpowiednim korpusie. Będzie to wspomagane przez oprogramowanie do przetwarzania danych. Oprogramowanie użyte do analizy w tej rozprawie nazywa się WordLister i należy do zestawu narzędzi o nazwie WordSmith Tools.1. Program ten pozwala na szybką budowę listy słownictwa, a mianowicie jednego katalogu wszystkich słów we wszystkich wybranych plikach lub wielu list słownictwa, które rozróżniają listy słownictwa dla każdego z plików. Kompilacja występuje w „trzech różnych formatach: analiza statystyczna, lista słów uszeregowana według frekwencji i alfabetycznie uporządkowana lista słów” [47].

Według Scotta i Tribble'a (2006) lista słów to spis wszystkich słów w tekście lub korpusie, zwykle wraz z danymi dotyczącymi liczby wystąpień. Dostępne są również listy lematów, które nie rejestrują odmienionych form wyrazów jako przykładów występowania hasła lub lematu. Do innych skwantyfikowanych i wymienionych zjawisk należą tagi [243].

Przeglądanie list słów dostarczonych przez pokaźne korpusy referencyjne może pozwolić na analizę frekwencji występowania słów w badanym języku. Spisy słownictwa dostarczone przez tekst lub korpus specjalistyczny mogą być wykorzystane podczas wstępnego badania leksyki tekstu lub typu tekstu i podobnie pouczające może okazać się porównanie takiej listy słownictwa z listą z ogólnego korpusu bibliograficznego. Może to ujawnić na przykład, że najczęściej występujące słowa różnią się od tego, co uważa się za normę [243].

Kolejny etap zakłada utworzenie list okurencyjnych; WordLister konwertuje wszystkie pliki w korpusie i tworzy rejestr wszystkich form (różnych wariacji danego słowa) sklasyfikowanych według częstotliwości występowania [210]. Na przykład niska częstotliwość przyimka „of” może wskazywać na ograniczoną liczbę postmodyfikowanych fraz rzeczownikowych, co z kolei sugeruje bardziej bezpośredni



i mniej formalny styl. Użycie funkcji słów kluczowych może być również wykorzystane do wyszukiwania frekwencji słów, jeśli ktoś pragnie bardziej systematycznej metody [243].

Listy frekwencyjne są bardzo wygodne w dostarczaniu naukowcom informacji o przedmiotowym charakterze tekstu lub korpusu, przy użyciu narzędzi do przetwarzania korpusu (na przykład WordSmith Tools lub Sketch Engine Tools, które są wykorzystywane do analizy korpusu w niniejszej rozprawie). Kolejną funkcją tego narzędzia jest identyfikacja słów kluczowych (ang. keywords), czyli identyfikacja słów pojawiających się w tekście lub korpusie z wyjątkowo wysoką częstością występowania. W rezultacie funkcja słowa kluczowego zapewnia lepszą wskazówkę dotyczącą tematycznego charakteru danego tekstu [34].

Słowa kluczowe posiadają wiele funkcji. Mogą być narzędziem wyszukiwania służącym do eksploracji i klasyfikacji tekstu, a także narzędziem analitycznym do interpretacji tekstu i analizy dyskursu. *Keyness* to termin zdefiniowany przez Scotta i Tribble'a (2006) jako taki, który pozwala wydobyć główne idee dominujące w tekście. *Keyness* identyfikuje się na podstawie analizy statystycznej: „słowa kluczowe to słowa, które w próbie tekstu występują znacznie częściej niż można by się spodziewać, biorąc pod uwagę ich częstość w dużym ogólnym korpusie referencyjnym” [263].

Słowa kluczowe nie zawsze są pojedynczymi słowami: istnieją grupy kluczy (powtarzające się ciągi słów) lub frazy kluczowe. John Sinclair (2004) stwierdza, że perspektywa korpusowa umożliwia spojrzenie na kombinację słów i wykorzystanie frazeologii jako idealnego punktu wyjścia w dyskusji o systematycznej relacji między tekstem a formą.

Innym sposobem na właściwą analizę tekstu jest użycie funkcji konkordancji. Konkordancja to lista słów (lub wzorców) w korpusie (lub tekście) wraz ze słowami otaczającymi. W istocie konkordancje to metoda prezentacji danych, w której wyszukiwany termin i jego bezpośredni tekst są zorganizowane w taki sposób, aby można było łatwo oszacować wzorzec tekstowy i zidentyfikować wzorce otaczające wyszukiwane hasło. Bowker i Barlow (2008) definiują konkordancje jako przekształcenia tekstu, dające analitykowi różne perspektywy [48].

Bower i Pearson (2002) wymieniają dwa typy konkordancji: konkordancje jednojęzyczne i konkordancje dwujęzyczne. Te pierwsze są stosowane w tekstach pisanych w jednym języku i pozwalają na szybkie odnalezienie wszystkich wzorców wraz z ich bezpośrednimi kontekstami. Konkordancje dwujęzyczne z kolei są używane na równoległych korpusach; te konkordancje ułatwiają wzorzec w jednym języku i zestawiają go z wzorcami w odpowiednich korpusach w języku docelowym [47].

Zastosowanie konkordancji umożliwia automatyczne wydobycie informacji o występowaniu i współwystępowaniu słów w odpowiednich tekstach lub korpusie. Uważa się, że dane zgodności we współczesnej lingwistyce korpusowej oferują stymulujący wgląd w strukturę i użycie języka poprzez ocenę słów w prawdziwych tekstach oraz identyfikowanie wzorców i znaczenia leksyki i gramatyki [243]. Taką samą opinię podzielają McEnery i Hardie (2012), którzy dodają, że materiał korpusu musi być analizowany za pomocą narzędzia konkordancyjnego, aby uzyskać właściwą i wiarygodną analizę wzorców leksykalno-gramatycznych.

Automatyczne ekscerpowanie list słów pozwala badaczom zajmować się tylko słowami zebranymi z korpusu, w którym się pojawiły. Funkcja zgodności daje analitykowi możliwość rozważenia słów w kontekście ich występowania w tekstach i pozwala dostrzec znaczenie słowa w tekście oraz sposób, w jaki to znaczenie jest kreowane w poszczególnych przypadkach [243].

## KONKLUZJE:

Trzeci rozdział niniejszej rozprawy dotyczy problematyki językoznawstwa korpusowego. Począwszy od definicji korpusu, z perspektywy historycznej jaką prezentowali Francis (1964) czy Leech (1992), aż po bardziej współczesne definicje zaproponowane przez Sinclaira i Mauranena (2005), czy McEnery i wsp. (2006). Najbardziej przekonująca jest jednak interpretacja Griesa (2016), która koncentruje się nie tylko na tekstach pisanych, ale także na dyskursie mówionym. Studia z zakresu językoznawstwa korpusowego pozostają w procesie rozwoju, wprowadzane są nowe trendy językowe, o których mowa w pierwszej części rozprawy [108].

W niniejszym rozdziale omówiono istotne kwestie w zakresie badań wspomaganych korpusem, a także dokonano analizy porównawczej elementarnych podejść korpusowych (w tym podejścia wielowymiarowego i socjolingwistycznego). Pierwsze podejście omawia przetwarzanie danych korpusowych, które koncentruje się na rozpoznawaniu jednostek dyskursu opartych na słownictwie. Nauczanie oraz proces uczenia się słownictwa okazuje się być najistotniejszą kwestią dla socjolingwistów, o czym świadczą przywołane dyskursy. Drugie podejście obejmuje analizę danych, które są kategoryzowane pod kątem różnych kontekstów i typów interakcji. Badania nad korpusami elektronicznymi okazały się niezbędną techniką dla językoznawców. W rezultacie omawiana gałąź studiów bardzo się rozwinęła w ostatnich latach.

Rozdział trzeci przedstawia istotę językoznawstwa korpusowego i koncentruje się na trzech obszarach analizy językowej, wyróżnionych przez wielu naukowców, m.in. McEnery'ego i Gabrielatos'a (2006). Te trzy obszary obejmują ogólny opis języka, analizę leksykalno-gramatyczną i nauczanie języka. W zakresie opisu języka wspomina się o odwołaniu do różnych tekstów i dyskursów [161], [26], [142], [185], co zostanie podkreślone w części empirycznej niniejszej rozprawy - w analizie dyskursu medycznego. Analiza leksykalno-gramatyczna ściśle podkreślała znaczenie tzw. wiązek słów (ang. word-chunks: [249]) lub (ang. multi-word units of meaning: [105], [124], [295]). Szerokie stosowanie tych wiązek znacznie poprawia styl wypowiedzi, co stwierdzili Bower i Pearson (2002). Fragmenty i jednostki wiązek będą przedmiotem dyskusji w dalszej

części rozprawy. Sumę rozważań stanowi opis wyniku analizy aspektu metodologicznego wraz z narzędziami korpusowymi, które mogą uprościć jej proces i przyspieszyć sformułowanie wyników językowych.

Ponadto Tognini-Bonelli (2001) określa korpusy jako bazę danych, która umożliwia analizę w różnych kontekstach, przy różnych strukturach składniowych lub funkcjach. Wystąpienie danego słowa lub frazy w tych kontekstach nazywa się podejściem opartym na użyciu, ang. the usage-based approach, Tognini-Bonelli (2001). Podejście oddolne umożliwia analizę od szczegółowej do ogólnej, co stanowi punkt odniesienia dla studiów korpusowych. Bączkowska (2005a, 2005b) doszła również do podobnych wniosków w dziedzinie dydaktyki, pokazując i analizując podejścia probabilistyczne i restrykcyjne. Pierwsza technika wymaga analizy dużej bazy danych w celu wygenerowania uniwersalnych reguł. Drugie założenie przedstawia szczegółową analizę licznych przykładów, które mają wykazać poszczególne ograniczenia.

W poprzednich dekadach zaobserwowano duże zainteresowanie badaniami korpusowymi. Sinclair (1987) podkreśla, że Collins COBUILD English Language Dictionary jest pierwszym istotnym podejściem do efektywnego skorzystania z korpusów elektronicznych. Obecnie obserwuje się niezmiennie rosnące zainteresowanie tworzeniem korpusów. W rozdziale trzecim omówiono różne współczesne korpusy, które powstały w ramach badanego dyskursu, z których niektóre zasługują na szczególną uwagę. Niniejsza rozprawa koncentruje się głównie na korpusie brytyjskim, ponieważ pierwszym korpusem był korpus londyńsko-lundskiego języka mówionego brytyjskiego angielskiego jako prekursor dla różnych korpusów [248], [161].

Obecnie powstało wiele współczesnych korpusów, w tym MICASE, CALLHOME, CALLFRIEND, British National Corpus, Corpus of Contemporary American English, The Cambridge English Corpus, CAWSE, ELFA, akademicki korpus ELF, ICLE, LINDSEI, VOICE i Trinity Lancaster Korpus, który jest uważany za jeden z największych korpusów mówionego języka angielskiego osób posługujących się językiem angielskim jako drugim. Jednym z naukowców, który wniósł ogromny wkład pracy badawczej w językoznawstwo korpusowe, jest Mark Davies, profesor lingwistyki na Uniwersytecie Brigham Younga w USA. Jest autorem Corpus of Contemporary American English (COCA), który jest chyba

jedynym na świecie tak odmiennym gatunkowo korpusem, obejmującym nie tylko strony internetowe, ale także napisy do filmów (np. *The Little Prince* Antoine de Saint-Exupéry lub *Harry Potter* JK Rowling). COCA zawiera ponad miliard słów.

Do najbardziej podstawowych metod stosowanych w analizie korpusowej należą narzędzia do obsługi tekstu, wśród których znajdują się listy słów, listy frekwencyjne, słowa kluczowe i konkordancje. Metody zostaną bliżej omówione w końcowej części niniejszej rozprawy, gdyż posłużą jako narzędzie dowodzące przedstawionej teorii. Jak podkreślają Bower i Pearson (2002), odpowiednią analizę danych korpusowych należy rozpocząć od przygotowania list słów za pomocą specjalnych narzędzi konkordancyjnych. Narzędzia te umożliwiają zestawienie odpowiednich list frekwencyjnych, które pokazują, jakie wielowyzrazowe jednostki lub fragmenty występują najczęściej w danym dyskursie. Dzięki temu łatwo jest wskazać słowa kluczowe jako najczęściej jednostki w korpusie lub konkordancje, wszystkie z nich są określane nie tylko pojedynczymi słowami, ale także tzw. ko-tekstem (ang. co-text). To narzędzie korpusowe umożliwia ilościową i jakościową analizę bazy danych poprzez określenie odpowiednich wzorców leksykalno-gramatycznych.

## ROZDZIAŁ IV

### FRAZY, KOLOKACJE I WIĄZKI LEKSYKALNE

Rozdział czwarty wprowadza i definiuje szeroki zakres terminologii frazeologicznej, w którego zbiorze szczególne miejsce zajmują takie pojęcia jak wiązki leksykalne i kolokacje. Istotną część rozprawy stanowi omówienie wyników pracy zespołów lingwistów, prezentujących odmienne stanowiska co do zagadnień z zakresu frazeologii. W gronie naukowców, którzy przyczynili się do rozwoju badań, znajdują się polscy badacze; warto w tym miejscu przywołać m.in. Skorupkę, Bogusławskiego, Chlebdę czy Lewickiego. Pośród badaczy rosyjskich szczególną uwagę należy skierować w stronę teorii badawczych Vinogradova, Smirnitsky'ego i Arnolda.

Wiązki leksykalne należy poddawać analizie w kontekście frazeologicznym, ze szczególnym uwzględnieniem kwestii ich ekstrakcji oraz wykorzystania w dyskursie akademickim pisanym przez rodzimych i nierodzimych użytkowników języka angielskiego, a także ich aspektu dydaktycznego.

#### **4.1. Zakres frazeologii**

Cowie (1994, 2004) definiuje frazeologię jako „badanie struktury, znaczenia i użycia kombinacji słów” [66], [69]. Zainteresowanie uczonych badaniami dotyczącymi sposobu łączenia się słów narodziło się dość szybko jako gałąź dynamicznie rozwijającej się myśli lingwistycznej. Firth (1968), Palmer (1933) i Jespersen (1924) publikowali prace dotyczące wyrażen stałych i kolokacji [131]. Kolejna fala zainteresowania zagadnieniem miała miejsce w latach 70. i trwała aż do początku lat 90., kiedy to nowe pokolenie badaczy przeprowadziło badania nad zdaniami schematycznymi w języku pisanym

i mówionym, zwłaszcza za pomocą ogromnej bazy danych znanej jako korpusy językowe [89], [112], [194], [25], [183], [166], [203], [250].

Analiza obszernego zbioru danych (korpusów) ujawniła informacje dotyczące powtarzających się wzorców w języku ojczystym oraz kluczowej roli kontekstu w odkrywaniu znaczenia słów i struktur gramatycznych. Sinclair (1991) stwierdza, że „użytkownik języka ma do dyspozycji dużą liczbę częściowo prekonstruowanych fraz, które stanowią pojedyncze wybory, nawet jeśli mogą wydawać się analizowane w segmentach”, sugerując tym samym, co powinno być poddane analizie językowej. Odtąd podkreśla się charakter konstrukcji frazeologicznych oraz elementy idiomatyczne obecne w większości tekstów niespecjalistycznych. Podobny wniosek przedstawił Kjellmer (1987) pisząc o jednostkach frazeologicznych zwanych kolokacjami – „we wszelkiego rodzaju tekstach kolokacje są nieodzownym elementem, z którego w dużej mierze składają się nasze wypowiedzi” [150], [152], [163].

#### **4.2. Frazeologia- definicje i założenia**

W zależności od podejścia badacza do jednostek frazeologicznych stosuje się szeroki zakres ich nazewnictwa: wyrażenia skonwencjonalizowane (ang. conventionalised expressions: [127], kolokacje (O'Keeffe i in., 2007: 59), kawałki (ang. chunks: [252]), jednostki wielowyrazowe, frazy leksykalne i wiązki leksykalne (ang. multi-word units, lexical phrases and lexical bundles [30]), a także bloki tekstowe (ang. text building blocks: [28]). Co więcej, koncepcje dotyczące ich analizy są tak liczne, jak nazwy użyte do ich oznakowania. Biber i wsp. (2005) stwierdzają, że różnice między badaniami nad kombinacjami słów są następujące:

- a) „cele badawcze (opis pełnego zakresu vs. mały zestaw jednostek wielowyrazowych);
- b) kryteria identyfikacji jednostek wielowyrazowych (istotność percepcyjna, kryteria częstotliwości itp.);

- c) charakterystyka formalna jednostek wielowyrazowych (ciągi ciągłe, ramki nieciągłe lub wzorce leksykalno-gramatyczne; kolokacje dwuwyrazowe a ciągi dłuższe);
- d) liczba użytych próbek tekstów (małe vs duże korpusy);  
oraz
- e) obecność lub brak porównań rejestrów (tylko teksty pisane, tylko teksty mówione, oba)” [31].

Howarth (1998) uważa, że istnieje zasadnicza różnica w podejściu do badań frazeologicznych, wynikająca z tego, że koncentrują się one na jednym elemencie spektrum frazeologicznego, takim jak idiomy, kolokacje czy formuły mowy. Co więcej, zdaniem Howartha frazeologia rozwijała się niemal niezależnie w wielu dziedzinach, takich jak językoznawstwo opisowe, analiza dyskursu i leksykografia, przyswajanie drugiego języka, a w niektórych przypadkach nawet badaniu właściwości językowych sztucznej inteligencji. Wynalazku, który w obecnej chwili stanowi odrębne zagadnienie naukowe, z elementarną kategorią określającą jego funkcjonalność - przetwarzaniem i wykorzystaniem mowy naturalnej przez boty. Według Howartha, w świetle powyższego, ujednoczenie istniejącej terminologii i pojęć okazuje się być trudniejsze [124].

#### **4.2.1. Jednostki wielowyrazowe**

Jednostki wielowyrazowe zostały zdefiniowane na wiele sposobów, chociaż wszystkie mogą być postrzegane jako synonimy. Etykiety, nadane jednostkom wielowyrazowym, zawierają liczne odpowiedniki, takie jak „wzorce prefabrykowane” (ang. “prefabricated patterns” [112]), „formuły rutynowe” (ang. “routine formulae” [64]), „frazy leksykalne” (ang. “lexical phrases” [194]), „kawałki” (ang. “chunks” [75]), „zleksykalizowane korzenie” (ang. “lexicalized stems” [209]) lub „sekwencje formuł” (ang. “formulaic sequences” [293], [294], [239]). Jednocześnie nadano im bardziej



popularne nazwy, takie jak „wyrażenia stałe”, „(zastrzeżone) kolokacje” czy „jednostki/wyrażenia wielowyrazowe” (ang. “fixed expressions”, “(restricted) collocations” or “multi-word units/expressions” [201]. Pomimo różnych pozycji terminologicznych w tej i kolejnych sekcjach, termin jednostki wielowyrazowe jest używany jako termin bardziej ogólny [173].

Niezależnie od nazw, powszechnie uważa się, że jednostki wielowyrazowe stanowią istotną cechę użycia języka. Takie „gotowe” (tj. prefabrykowane) frazy ułatwiają przepływ produkcji wyrażen w czasie fizycznym, a ich znaczenie jest równe przepływowi słownictwa jednowyrazowego przy badaniu pragmatyki lub semantyki języka [194].

Jak stwierdził Bolinger już w 1976 r., obecność wyrażen językowych, które nie byłyby, przynajmniej do pewnego stopnia, „prefabrykowane”, wydaje się mało możliwa i jest uniwersalną cechą każdego języka. Kwestią sporną jest pytanie, jak płynnie użytkownik lub osoba ucząca się języka obcego może przyjąć wielowyrazową płynność użytkownika natywnego [38].

Frazeologiczne jednostki wielowyrazowe stanowią wyzwanie nawet dla zaawansowanych uczniów drugiego języka, jeśli chcą naśladować znajomość języka native speakerów, czyli jeśli chcą brzmieć naturalnie [209], [103]. Często zdarza się, że frazy tworzone i używane przez obcokrajowców wydają się sztuczne. Jedną z najczęściej analizowanych jednostek wielowyrazowych, które przyczyniają się do osiągnięcia naturalności, są kolokacje, o których mowa w kolejnym podrozdziale.

#### **4.2.2. Kolokacje**

Mówi się, że kolokacje są częściowo kompozycyjnymi kombinacjami wyrazów, ponieważ „jeden element (tu czasownik) ma specjalistyczne znaczenie określone przez drugi element (tu rzeczownik)” [67]. Co więcej, kolokacje są oznaczone jako motywowane, co należy rozumieć jako możliwe do analizy [123]. Hausmann (1985) stwierdził, że podstawa jest zazwyczaj rzeczownikiem, ponieważ „rzeczowniki wyrażają

rzeczy i zjawiska w świecie, o których ludzie znajdują coś do powiedzenia”, podczas gdy z drugiej strony „czasowniki i przymiotniki byłyby możliwe jako słowa podstawowe tylko w kombinacjach z przysłówkami” [116].

Czasowniki specjalistyczne wykazują słabszą zależność od rzeczownika. Dlatego niektórzy badacze postanowili stworzyć odrębną grupę „wyspecjalizowanych kombinacji leksykalnych”, popierając swoją decyzję argumentem, że jeśli można je rozpatrywać w kategoriach ograniczeń, nie należy ich traktować jako kolokacji, które są bardziej ogólne (L. Homme i Bertrand, 2000).

Cowie i wsp. (1983) przedstawiają następujące trzy kryteria, które można zastosować do kategoryzacji i opisu kolokacji (leksykalnych):

- (a) „czasownik używany w znaczeniu przenośnym („oferować” lub „rozszerzać”),
- (b) określenie kontekstowe przez (arbitralnie) ograniczony zbiór rzeczowników,
- (c) określenie czasownika w tym samym znaczeniu przez ten sam ograniczony kontekst” [65].

Co więcej, Howarth (1996) analizuje środki, które angażują wyrozumiałość wobec tych kryteriów. Na wstępie stwierdza, że zasada (a) może być rozumiana szeroko, tak, że obejmuje czasowniki używane w „piśmie technicznym” lub „wyspecjalizowanym znaczeniu”, przy czym te pierwsze mogą być postrzegane jako „obszary życia codziennego”. Kryterium zawieszające (b) pozwoliłoby „włączyć do pasma frazeologicznego spektrum te kombinacje czasownika w sensie specjalistycznym i potencjalnie dużą liczbę kolokujących rzeczowników”. Stąd kryterium (c) czyni przepisy mniej restrykcyjnymi i umożliwia stosowanie pewnych czasowników równoznacznych, jak proponują [67].

Kwestia, którą należy poruszyć w tym miejscu, dotyczy zakresu „ograniczonego wyboru” słowa związanego z podstawą kolokacji, a także zakresu odmiany, jaka powinna być dozwolona, aby można było założyć, że kolokacja jest „niezależna” i nie jest zdeterminowana jej podstawą [209]. Warto wiedzieć, że zwroty w języku angielskim, takie jak prowadzenie sklepu, czy prowadzenie firmy, mogą być używane jako właściwe

kolokacje. Praktyczne byłoby poznanie możliwych alternatyw w każdym przypadku, aby można było wybrać kombinację, która byłaby zarówno właściwa, jak i pasująca do sytuacji.

W języku angielskim kolokacje zawierające czasownik „zdelekskalizowany” lub „lekki”, taki jak *make*, *break*, *have* i *take* (pl. *robić*, *łamać*, *mieć* i *brać*), są zwykle postrzegane jako unikalna forma kolokacji, mimo że stanowią grupę, w której bardzo często występuje niedosłowne znaczenie czasowników. Tym, co sprawia, że są one postrzegane jako różne od kombinacji typu *follow a regulation* i *launch an enterprise*, jest to, że łączy się je z szeroką gamą rzeczowników, których nie da się umieścić w określonych granicach (*make an improvement, sense, a decision, a mistake, a contribution, a choice, etc.*). W rezultacie mogą być używane jako formy skrajnie polisemiczne, łatwo zmieniające swoje znaczenie w celu dopasowania do rzeczowników lub posiadania słabszych, ale stabilnych „gramatycznych” znaczeń.

Howarth (1966) zaproponował podział fraz, takich jak kombinacje czasownik + dopełnienie nominalne, w zależności od stopnia idiomatyczności:

- (1) „w swobodnych kolokacjach, poza zakresem frazeologicznym, oba elementy są używane w sensie dosłownym,
- (2) kolokacje ograniczone mają jeden element, który jest używany w sensie dosłownym, podczas gdy drugi jest wyspecjalizowany (czasownik zdelekskalizowany, techniczny lub figuratywny),
- (3) idiomy figuratywne mają całokształt znaczenia figuratywnego, ale zachowują obecną interpretację dosłowną,
- (4) czyste idiomy mają w całości znaczenie figuratywne, ale nie zachowują aktualnej dosłownej interpretacji” [123].

Podział na kolokacje ograniczone i idiomy, uważane za najłatwiejsze do zdefiniowania, opiera się na idiomach o jednorodnym znaczeniu, podczas gdy kolokacje mają charakter kompozycyjny – jeden element jest używany w pierwotnym znaczeniu, drugi jest ograniczony przez wcześniejsze znaczenie [65]. Kolokacje zastrzeżone

i kolokacje wolne dzielą się w istotny sposób, ponieważ nakładają ograniczenia na to, co można uznać za frazeologiczne. Howarth (1996) stwierdza, że *go to hospital / work / school* są swobodnymi kolokacjami, mimo że są to wyrażenia zinstytucjonalizowane (synonim terminu „frazy leksykalne” używane przez [194]), ze względu na dosłowne znaczenie obu elementów, co jest bezsprzeczne. Odwrotnie, *take a vacation* jest traktowana jako kolokacja ograniczona [123].

Spektrum frazeologiczne powinno obejmować wszystkie zinstytucjonalizowane wyrażenia, a „osobliwość utrwalonych kombinacji powinna być postrzegana jako normalna cecha, którą z kolei można uznać za motywowaną przez typowe konteksty sytuacji i ram” [123]. Frazeologię można rozumieć jako językowy rezerwuar praktyk kulturowych, które pozostają specyficzne dla języka. Stąd dość regularna jest analiza dykcji międzyjęzykowej, prowadząca do wyjaśnienia genezy wielu z tych zachowań językowych i kulturowych. Dziesiątki z nich mogą być stosunkowo ogólne, co ujawni fascynujące aspekty ludzkiego poznania [104].

W wielu językach uznano szeroko rozumiane wyrażenia zbiorowe [54]. Powszechnie uważa się, że tradycja frazeologiczna ma swoje podłoże w Rosji i Niemczech [285]. W związku z tym języki rosyjski i niemiecki były jednymi z pierwszych, które zyskały kompleksową perspektywę frazeologiczną, nawet jeśli przedsięwzięcie zostało następnie rozszerzone na angielski i francuski. Z czasem większość języków europejskich skorzystała z badań dotyczących wyrażeń zbiorowych [104].

Większość członków Europejskiego Towarzystwa Frazeologicznego to badacze niemieckojęzyczni, co stanowi podstawę twierdzenia, że główna część frazeologii międzyjęzykowej polegała na zestawieniu zbiorowych fraz z języka niemieckiego z idiomami występującymi w innych językach, w tym na przykład węgierskim [119], fińskim [154], słowackim [82], holenderskim [214], rosyjskim [79], francuskim [282], litewskim [52], rumuńskim [297], japońskim [281] oraz hiszpańskim [215], [104].

Zainteresowanie metodologią badania języka angielskiego jak i udziałem w procesie badawczym jego ewolucji, wzrasta. Dotychczas podejmowano próby poddania podobnej analizie języka niemieckiego [96], hebrajskiego [197], arabskiego [6], łotewskiego [284], francuskiego [96], hiszpańskiego [189] oraz niemieckiego i polskiego [208]. Wiodący

badacze Corpas Pastor (2016) oraz Granger i Meunier (2008) wybrali podejście polskie i rosyjskie jako reprezentatywne w grupy języków zachodnio-, południowo- i wschodniosłowiańskich [104].

### **4.3. Podejście do frazeologii - perspektywa polska**

Warto zauważyć, że frazeologia była zagadnieniem językowym omawianym już przed „erą korpusową”. Idiomy upowszechniły się dzięki narzędziom korpusowym, które umożliwiły szybką i precyzyjną analizę dużej liczby tekstów lub danych w krótkim czasie. Frazeologię można postrzegać jako lingwistyczne przechowywanie licznych konwencji kulturowych, które są charakterystyczne dla danego języka. Dlatego analiza frazeologii w różnych językach wydaje się dość instynktowna, ponieważ taka analiza prowadzi do wyjaśnienia początków wielu tradycji językowych i kulturowych. Corpas Pastor (2016) wykazał, że polska perspektywa lingwistyczna przyczyniła się w znaczącym stopniu do rozwoju leksykologii [61].

W Polsce pierwsze poważne badania z zakresu frazeologii, uważanej za podstawę frazeologii teoretycznej, prowadzili Skorupka (1967) oraz Doroszewski (1969 – 1978). Opracowali definicję frazeologii i wprowadzili podstawową klasyfikację formalną i semantyczną jednostek frazeologicznych, znacząco wpływając na późniejsze badania dotyczące stylistyki, leksykologii i leksykografii [81].

Najpopularniejszą i najważniejszą taksonomią polskich jednostek frazeologicznych jest ta przedstawiona przez Stanisława Skorupkę (1967). Opiera się na dwóch zasadach: formalnej i semantycznej. W tej typologii pierwsza kategoria odnosi się do gramatycznego charakteru ciągów leksykalnych tworzonych w jednostkach frazeologicznych, druga zaś do relacji semantycznych między prezentowanymi grupami leksykalnymi.

Klasyfikacja gramatyczna obejmuje trzy rodzaje polskich pozycji frazeologicznych:

- fraza nominalna (pl. wyrażenie) – centrum to zazwyczaj przymiotnik odrzeczownikowy, imiesłów przymiotnikowy lub przysówek – np. słowo „czerwone i czarne”,
- zwrot/wyrażenie czasownikowe (pl. zwrot) – są to często powtarzające się i bardzo złożone w znaczeniu zdania w postaci pełnych zdań lub ich odpowiedników, np. przysłów lub maksym,
- zdanie (pl. fraza) – środek to zwykle czasownik lub imiesłów przysłówkowy – np. „kochać kogoś na zabój”, „biorąc pod uwagę”.

Ponadto Skorupka (1950) wymienił kilka podtypów wyżej wymienionych kategorii frazeologicznych:

- wyrażenia,
- wyrażenia porównawcze,
- wyrażenia figuratywne,
- wyrażenia rymowane.

Biorąc pod uwagę semantykę jednostek frazeologicznych, Skorupka (1967) zaproponował następującą typologię:

- neutralne (jednostki frazeologiczne; związki frazeologiczne/ wyrazowe łączliwe) – każdy element ma swoje znaczenie. Są to frazy (zwykle kolokacje), których znaczenie wynika z jego składowych i może być zmieniane w zależności od kontekstu, np.: „drewniany dom”;
- stałe (stałe jednostki frazeologiczne; związki frazeologiczne/wyrazowe stałe) – elementy, których składowe nie mogą być zmieniane. W razie jakiegokolwiek zmiany zmieni się też znaczenie, np. „drzeć z kimś koty”, „patrzeć przez różowe okulary”;

- zmienne (luźne jednostki frazeologiczne; związki frazeologiczne/ wyrazowe luźne) – elementy leksykalne powiązane semantycznie. Ich znaczenie zależy od ich użycia w kontekście i częstotliwości występowania w danym dyskursie, np.: „dobić targu”.<sup>3</sup>

Niektóre z wyrażen zaklasyfikowanych, przede wszystkim jako zwroty wolne lub kolokacyjne, mogą później zostać przekształcone w frazy stałe lub te same frazy mogą być stałe lub dowolne w zależności od kontekstu, w którym występują. Skorupka reprezentuje tzw. analityczne podejście do frazeologii, które podkreśla semantyczne aspekty jednostek frazeologicznych; jednak wprowadzając związki luźne rozszerzył kategorię jednostek frazeologicznych na więcej swobodnych zwrotów [254].

Skorupka skupiał się na semantycznych aspektach fraz, natomiast Andrzej Bogusławski (1976) zaproponował włączenie do swoich badań nad frazami jeszcze więcej pozycji, m.in. następujące wiązki słów: wszystko rozumiem, ale; no tak, ale itp. Przywołanym wtrąceniom brakuje semantycznej stałości (spójności), co oznacza, że elementy konstytutywne frazy nie wykazują wysokiego stopnia semantycznej łączliwości. Ponadto są niekompletne składniowo. Tym samym typologia jednostek frazeologicznych Bogusławskiego odbiega od podejścia analitycznego i zbliża się do tego, co później nazwano podejściem dystrybucyjnym. Co ciekawe, wykazują one podobieństwo do koncepcji wiązki leksykalnej opisanej przez Bibera (2003). Lewicki (1983) zaproponował podobną typologię polskich jednostek frazeologicznych, która zawiera wyrażenia, frazy słowne i znaczniki frazeologiczne będące albo jednostkami gramatycznie kompletnymi (jednostki znakowe gramatycznie), albo jednostkami fragmentarycznymi (jednostki znakowe fragmentaryczne). Włączył więc sekwencje syntaktyczne (które mogły być niekompletne) w jednostki frazeologiczne.

---

<sup>3</sup> Więcej informacji na temat taksonomii polskiej frazeologii:

Awdiejew A., Problemy wyodrębnienia i klasyfikacji frazemów, w: *Frazeologia słowiańska. Księga pamiątkowa poświęcona prof.dr hab. Halinie A. Lilicz*, red. Bałowski M., Chlebda W. Opole 2001

Bąba S., Liberek J., Kilka uwag o współczesnym uzusie frazeologicznym w: *Norma językowa w polszczyźnie*, red. Bugajski M., Zielona Góra 1995

Another author who uses this term is Igor Melčuk (1988). He uses it for non-free phrases, that is, such phrases whose meaning is completely dependent on those phrases components' semantic and syntactic features. This term is equivalent with the concept of lexical bundles.

Wojciech Chlebda (1991) zaproponował w swojej frazematyce pojęcie frazemu. Według Chlebdy, co też podkreśla również Bączkowska (2005a, 2005b), frazem jest jednostką frazeologiczną powtarzaną w różnych kontekstach sytuacyjnych, a następnie odtwarzaną przez nadawcę z kilku gotowych fraz, niezmienną, ale modyfikowalną. Współwystępowanie słów we frazach nie opiera się na kryterium spójności semantycznej (wewnętrznej spójności), ale raczej na kryterium odtwarzalności w danej sytuacji [59]. Teoria Chlebdy przypomina więc dystrybucyjne podejście do jednostek frazeologicznych, typowe dla badań opartych na danych korpusowych, gdzie głównym kryterium klasyfikacji jest liczenie częstotliwości ich stosowania (a nie jednoznaczność semantyczna) [18], [59].

W wyniku kategoryzacji opartej na frekwencji, wyrażenia takie jak *the example of* lub *as well as the* (stanowiące przykład tego, co Biber i in., 2004, nazywają wiązkami leksykalnymi) również uzyskały status jednostek frazeologicznych, pomimo ich niekompletności składniowej i wzajemnego oddziaływania na poziomie semantycznym. Wiązki leksykalne zwykle zaczynają się od przyimka lub przedimka. Podobnie mogą kończyć się innym przyimkiem lub przedimkiem. Stopień łączności pomiędzy poszczególnymi słowami, które wchodzi w skład wiązki, jest zwykle analizowany za pomocą narzędzi statystycznych, takich jak parametr MI (mutual information) [125] lub MI3 [199], [18], które zostaną opisane w dalszej części rozdziału. Ogólnie rzecz biorąc, wiązki o wysokiej liczbie łączników są zwykle przedmiotem badań, ale niektórzy autorzy niekoniecznie uważają je za adekwatne [236], [102].

#### **4.4. Podejście do frazeologii - perspektywa rosyjska**

Frazeologia to dział leksykologii stanowiący rdzeń znaczących wkładów dotyczących ustalania terminologii, a także typologii fraz, zdań i całych tekstów [54], [96], [68]. Tradycja frazeologiczna wywodzi się z Niemiec i Rosji [285]. Jak stwierdził Cowie (1998), rosyjskie podejście do frazeologii stanowiło frazotwórczy fundament, na którym w następstwie rozwoju języka oparto kolejne założenia. W konsekwencji rosyjski i niemiecki stały się pierwszymi językami, które zostały wzbogacone o wyrażenia



frazeologiczne, podnoszące poziom komunikacji w kontekście mowy symbolicznej, alegorycznej czy kontekstualnej, stwarzające możliwość wypowiedzenia się na lingwistycznym metapoziomie. Badanie języków niemieckiego i rosyjskiego wpłynęło na potrzebę ewolucji zagadnienia w języku angielskim i francuskim. Proces objął większość języków używanych w Europie. Z tego powodu należy wnikliwie przyrzeć się zarówno naukowej perspektywie rozwoju polskiej frazeologii, jak i przedstawić spektrum zagadnienia w kontekście frazeologii obecnej w rosyjskiej rodzinie językowej.

#### **4.4.1. Klasyfikacja jednostek frazeologicznych Winogradowa**

Pierwszą kompleksową w historii kategoryzacją frazeologii, którą przyjęli badacze różnych języków, była ta, którą przedstawił Winogradow (1947). Klasyfikacja ta opiera się na zasadzie semantycznej, która podkreśla jej spójność w obrębie ciągów frazeologicznych i charakteryzuje się złożoną strukturą (tzw. podejście analityczne). Klasyfikacja Winogradowa dokładnie interpretuje związek między znaczeniem zawartym w całej jednostce frazeologicznej a jej częściami składowymi. Niepodzielność, sztywność wyrażenia, bezpośrednio odpowiadają interpretacji i wyjaśnianiu, choć nie we wszystkich przypadkach, możliwość dostosowania kolejności lub formy składników i zastąpienia całości jednym słowem.

Klasyfikacja jednostek frazeologicznych Winogradowa (1974) dzieli je na cztery sekcje: fuzje frazeologiczne, jednostki frazeologiczne, kombinacje frazeologiczne i wyrażenia frazeologiczne (ang. phraseological fusions, phraseological units, phraseological combinations, and phraseological expressions).

Fuzja frazeologiczna jest konstruktem semantycznie niepodzielnym, a jego składniki nigdy nie wpływają na znaczenie jednostki frazeologicznej. Jako takie, fuzje reprezentują najwyższą formę mieszania się ze znaczeniem jego składników wchłoniętych przez całość. Poniższe przykłady doskonale oddają istotę frazeologicznej fuzji: *once in a blue moon* – co oznacza, że coś zdarza się rzadko; *under the rose* – oznaczać, że coś się wydarzy bardzo cicho; i wreszcie *to cry for the moon* – co oznacza żądanie

niemożliwego. Emocjonalne i ekspresyjne właściwości wspomnianych fuzji frazeologicznych spowodowały, że kategoria ta była często określana mianem idiomów. Językoznawcy rozumieją idiomy jako wyrażenia, które całkowicie utraciły swoją wewnętrzną formę i rzadko można wyjaśnić znaczenie ich składników.

W przeciwieństwie do frazeologicznej fuzji, jednostki frazeologiczne są uważane za semantycznie niepodzielne - w tym przypadku znaczenie jest uwarunkowane przez jego składniki. Składniki nie są obliczane w celu uchwycenia istoty; to raczej znaczenie z nich się wywodzi, ponieważ ich cel w jednostkach frazeologicznych nie jest dramatycznie odmienny od celu i znaczenia, którym zwykle służą.

Jednostki frazeologiczne są rozumiane jako wynik figuratywnego znaczenia ich składników. Odosobnione kombinacje słów, które stanowią ich składniki, mogą być rozumiane w ich bezpośrednim znaczeniu. Niektóre z przykładów to: *to come to one's senses* – znaczy zmienić zdanie; *to come home* – czyli trafić w sedno; i *to fall into a rage* – czyli wpaść w złość.

Co więcej, kombinacje frazeologiczne (kolokacje) to konstrukcje lub wyrażenia, w których każde słowo ma jasne i niezależne znaczenie, ale jeden ze składników ma również znaczenie ograniczone. W rezultacie ten ostatni ma znaczenie przenośne, podczas gdy pozostałe elementy są używane dosłownie. Przykładami tego rodzaju są: *to make an attempt, to make haste, or to offer an apology*.

Ostatnia kategoria kolokacyjna obejmuje wyrażenia frazeologiczne, które są semantycznie podzielnymi konstrukcjami, których składniki mają dowolne znaczenia. Powszechnie określane jako przysłowia, aforyzmy lub powiedzenia, są zazwyczaj przypisywane znanym osobistościom.

#### 4.5. Wiązki leksykalne w ramach kategorii frazeologicznych

Zainteresowanie badaniem wyrażen i fraz wzrosło wraz z rozwojem językoznawstwa korpusowego. Jak zauważa Bączkowska (2008), jednym z najczęściej przywoływanych tematów badawczych okazuje się być frazeologia w aspekcie kolokacji. Jednym z najrzadziej wymienianych tematów jest zagadnienie typów jednostek frazeologicznych, zwanych wiązkami leksykalnymi.

Wiązki leksykalne były przedmiotem badań wielu badaczy analizujących różne typy dyskursu. Do tej pory pakiety leksykalne były omawiane w tekstach akademickich [30], [127], [247] oraz w tekstach pisanych przez nierodzimych użytkowników języka angielskiego w kontekst różnych języków [63], [148], [240], [137].

Oprócz dyskursu akademickiego wiązki leksykalne były badane także w innych typach tekstów. Na przykład w dyskursie politycznym [206], w tekstach związanych z telekomunikacją [203], w dokumentach unijnych [130], w dyskursie prawnym [100], [35], [36], [37], Kopaczyk, 2016), a także w dyskursie farmaceutycznym [102] oraz, co najcenniejsze dla tej rozprawy, analizie dyskursu medycznego [18].

Wiązki leksykalne były również analizowane porównawczo w tekstach reprezentujących dyskurs z różnych dziedzin. Jak donosi Bączkowska (2018), dyskurs historyczny był analizowany na tle dyskursu biologicznego. Dyskurs biologiczny porównano z dyskursem o biznesie, elektrotechnice i lingwistyce stosowanej [127] natomiast dyskurs nauk biologicznych porównano z dyskursem biochemicznym [235] w [18]. Wiązki leksykalne w tekstach medycznych zostały przeanalizowane ilościowo, a także za pomocą narzędzi korpusowych [18].

Przeprowadzone do tej pory badania wyszczególniły różne typy kombinacji słów i wykorzystywały szeroki wachlarz metod badawczych. Przedmiotem niniejszego badania jest rodzaj kombinacji słów zwanych „wiązkami leksykalnymi” lub czasami „klasterami”. Zostały one po raz pierwszy zdefiniowane w Longman Grammar of Spoken and Written English [27].

Wiązki leksykalne były badane i eksplorowane przez różnych uczonych, w tym Leecha, Bibera, Conrada, Johanssona i Finegana, którzy dokonali wszechstronnego, korpusowego badania gramatyki angielskiej. Podali poprawną definicję wiązek leksykalnych, opisując je jako:

„wiązki słów, które wykazują statystyczną tendencję do współwystępowania” [27] oraz jako „powracające wyrażenia, niezależnie od ich idiomatyczności i niezależnie od ich statusu strukturalnego” [27].

Inną definicją jest ta wskazana przez Cortesa (2004), który identyfikuje wiązki leksykalne jako:

rozszerzone kolokacje, sekwencje trzech lub więcej słów, które statystycznie współwystępują w rejestrze, np. *as a result of, on the other hand, in the case of the, the context of the, it is likely to* [62].

Podobną definicję podają McEnery i Hardie (2012), którzy twierdzą, że:

Metodologicznie i technicznie „wiązki leksykalne” to po prostu n-gramy/klastery – powtarzające się sekwencje n słów [187].

W dobie rozwoju lingwistyki korpusowej dostęp do pokaźnych baz tekstowych pozwolił zauważyć, że niektóre frazeologizmy są nie tylko spójne semantycznie, ale także powtarzalne i spójne składniowo. Chlebda (1991) wskazywał już, że czynnik stopnia powiązania jest w istocie centralnym elementem analizy i klasyfikacji, jeśli chodzi o frazeologię analityczną. Przykład tego można znaleźć w pracach Winogradowa, o których mowa w paragrafie 4.4.1. W korpusowo-językowym podejściu do frazeologii częstotliwość występowania jest ważniejsza niż stopień powiązania między elementami jednostek frazeologicznych, o czym wspomina Chlebda (1991). Znaczenie odtwarzalności jest istotne w przypadku wiązek leksykalnych, które są na ogół luźnymi jednostkami, tworzonymi ad hoc i semantycznie rozłącznymi, a także odtwarzalnymi w różnych kontekstach [18], [59].

Liczni badacze zwracali uwagę w pracach analitycznych [127] [247], że najniższa częstotliwość wiązek leksykalnych jest widoczna w wiązках długich. Ta obserwacja jest oparta na przykładach znalezionych w konwersacyjnych i akademickich sekcjach Korpusu Języka Angielskiego / Longman Spoken and Written English Corpus, w których czterowyrazowe wiązki leksykalne są około dziesięć razy mniej powszechne niż trzywyrazowe, a pięciowyrazowe wiązki leksykalne są około stu razy rzadziej występujące niż trzywyrazowe: „Klaster trójwyrazowy można znaleźć nawet 80 000 razy na milion słów w rozmowie i ponad 60 000 razy na milion słów w prozie akademickiej” [206]. Z drugiej strony „klaster czterech słów pojawia się tylko około 8500 razy na milion słów w rozmowie i ponad 5000 razy na milion słów w prozie akademickiej” [206].

Biber i wsp. (1999) podają, że tylko 15% wiązek leksykalnych wykorzystywanych w komunikacji ustnej i 5% wykorzystywanych w prozie akademickiej można uznać za kompletne jednostki strukturalne. Dodatkowo Biber i wsp. (1999) stwierdzają, że wiązki leksykalne można podzielić na kilka typów. Wiązki leksykalne używane w dyskursie mówionym są generalnie uważane za zwykłe i należą do typu zaimek + czasownik + dopełnienie (np. *it's going to be, I want you to*). Według Bibera (1999) większość wiązek leksykalnych to części fraz przyimkowych lub fraz rzeczownikowych w dyskursie akademickim (np. *as a result of, on the other hand*). Na podstawie typowych kolokacji gramatycznych zaproponowano następującą kategoryzację wiązek leksykalnych. Kategorie strukturalne dotyczące prozy akademickiej przedstawia Tabela I.

Tab. I. Strukturalna klasyfikacja wiązek leksykalnych w prozie akademickiej zaadaptowana przez Bibera [27]

STRUKTURA:	PRZYKŁADY:
Fraza rzeczownikowa <i>of</i>	<i>the end of the, the beginning of the, the base of the, the point of view of</i>
Fraza rzeczownikowa z innymi fragmentami post-modyfikatora	<i>the way in which, the relationship between the, such a way as to</i>
Fraza przyimkowa z osadzonym <i>of</i>	<i>about the nature of, as a function of, as a result of the, from the point of view of</i>
Inne frazy przyimkowe	<i>as in the case, at the same time as, in such a way as to</i>

Antycypacyjne <i>it</i> + fraza czasownikowa / fraza przymiotnikowa	<i>it is possible to, it may be necessary to, it can be seen, it should be noted that, it is interesting to note that</i>
Czasownik bierny + fragment frazy przymkowej	<i>is shown in figure/ fig., is based on the, is to be found in</i>
Copula <i>be</i> + rzeczownik fraza/ przymiotnik fraza	<i>is one of the, may be due to, is one of the most</i>
(faza czasownikowa +) fragment <i>that</i>	<i>has been shown that, that there is a, studies have shown that</i>
(Czasownik/przymiotnik +) fragment z <i>to</i>	<i>are likely to be, has been shown to, to be able to</i>
Fragment zdania przysłówkowego	<i>as shown in figure/ fig., as we have seen</i>
Zaimek/rzeczownik fraza + <i>be</i> (+...)	<i>this is not the, there was no significant, this did not mean that, this is not to say that</i>
Inne wyrażenia	<i>as well as the, may or may not, the presence or absence</i>

Dodatkowo krótkie wiązki leksykalne są zwykle częścią dłuższych, np. czterowyrazowy klaster, *it should be explained*, służy jako część pięciosłowu – *it should be explained that*, który z kolei jest używany do stworzenia sześciosłowowego wiązku, *it should be explained that the*.

Biber, Conrad i Cortes (2003, 2004) również podjęli próbę stworzenia klasyfikacji wiązek leksykalnych według ich funkcji. Stworzyli oni wstępną taksonomię, która uwzględnia cele i znaczenia wiązek leksykalnych we właściwym kontekście oraz rozróżnia cztery podstawowe funkcje przedstawione w tabeli II:

- 1) Wyrażenia postaw- ang. stance expressions,
- 2) Organizatory dyskursu- ang. discourse organisers,
- 3) Wyrażenia referencyjne- ang. referential bundles,
- 4) Specjalne funkcje konwersacyjne- ang. special conversational functions.

Biber, Conrad i Cortes (2003, 2004) podają definicje dla każdej kategorii. Stance expressions przedstawiają oceny lub postawy danej wypowiedzi. Discourse organisers przedstawiają relacje, jakie istnieją między poprzednim a nowym dyskursem. Natomiast referential bundles są bezpośrednio połączone z kontekstem tekstowym lub fizycznym

w celu zidentyfikowania lub podkreślenia znaczenia danej wypowiedzi. Wreszcie, specjalne funkcje konwersacyjne polegają na wskazaniu celu i charakteru wypowiedzi, czy ma ona funkcję informacyjną, sprawozdawczą, czy uprzejmą [30], [137].

Tab. II. Klasyfikacja funkcjonalna wiązek leksykalnych [30]

<b>I. Stance expressions</b>  Wyrażają postawy lub oceny dyskursu	<b>II. Discourse organisers</b>  Odzwierciedlają relacje między dyskursami	<b>III. Referential bundles</b>  Odwołują się do relacji między dyskursami	<b>IV. Special conversational functions</b>  Specjalne funkcje konwersacyjne
a. Epistemic stance <i>I don't know if, I think it was, are more likely to, the fact that the</i> b. Attitudinal/ modality stance b1. Desire <i>if you want to, I don't want to</i> b2. Obligation/ directive <i>you might want to, it is important to</i> b3. Intention/ prediction <i>I'm not going to, it's going to be</i> b4. Ability <i>to be able to, can be used to</i>	a. Topic introduction/ focus <i>what do you think, if you look at</i> b. Topic elaboration/ clarification <i>I mean you know, on the other hand</i>	a. Identification/ focus <i>that's one of the, of the things that</i> b. Imprecision <i>or something like that, and stuff like that</i> c. Specification of attributes c1. Quantity specification <i>there's a lot of, how many of you</i> c2. Tangible framing attributes <i>the size of the, in the form of</i> c3. Intangible framing attributes <i>the nature of the, in the case of</i> d. Time/ place/ text reference d1. Place reference <i>in the United States</i> d2. Time reference <i>at the same time, at the time of</i>	a. Politeness <i>thank you very much</i> b. Simple inquiry <i>what are you doing</i> c. Reporting <i>I said to him/her that</i>

		d3. Text deixis <i>shown in figure N, as  shown in figure</i> d4. Multifunctional reference <i>the end of the, the  beginning of the</i>	
--	--	--	--

Wspomniana klasyfikacja funkcjonalna wiązek leksykalnych, zaproponowana przez Bibera, została szeroko zaakceptowana przez Conrada i Cortesa [29], [30], a następnie rozwinięta i zmodyfikowana przez innych badaczy, przede wszystkim [128]. Biber i wsp. (2004) zbadali formy, funkcje i frekwencje wiązek leksykalnych w korpusie składającym się z artykułów naukowych i rozpraw reprezentujących cztery różne dyscypliny. Na podstawie uzyskanych wyników Hyland (2008b) zmodyfikował klasyfikację Bibera i wsp. (2004) i stworzył kategorie mające na celu dokładniejsze odzwierciedlenie funkcji wiązek leksykalnych. Opracowana taksonomia zapewnia powiązanie między każdym klasterem a jedną z trzech ogólnych kategorii: tekstem, badaniami i uczestnikami. Kategorie te są następnie podzielone na podkategorie przedstawione w tabeli III.

Tab. III. Klasyfikacja funkcjonalna wiązek leksykalnych w piśmiennictwie akademickim [127]

<b>Wiązki zorientowane na badania</b>	<b>Wiązki zorientowane na tekst</b>	<b>Wiązki zorientowane na odbiorcę</b>
Pomaga pisarzom uporządkować ich działania i doświadczenia w prawdziwym świecie	Organizacja tekstu i jego znaczenia jako przesłania lub argumentu	Koncentruje się na pisarzu lub czytelniku
<b>Położenie</b> Wskazując czas/miejsce <i>At the beginning of, at the same time, in the present study</i> <b>Wiązki proceduralne</b> <i>the use of the, the role of</i>	<b>Sygnaty przejścia</b> Ustalenie addytywnych lub kontrastowych powiązań między elementami <i>on the other hand, in addition to the, in contrast to the</i>	<b>Cechy stance</b> Postawy i oceny nadawcy <i>are likely to be, may be due to, it is possible that</i> <b>Cechy zaangażowania</b> <i>it should be noted that, as can be seen</i>



<p><i>the, the purpose of the, the operation of the</i></p> <p><b>Ujęcie ilościowe</b> <i>the magnitude of the, a wide range of, one of the most</i></p> <p><b>Opis</b> <i>the structure of the, the size of the, the surface of the</i></p> <p><b>Temat</b> <i>Związane z dziedziną badań in Hong Kong, the currency board system</i></p>	<p><b>Wyniki</b> Zaznacz inferencyjne lub przyczynowe relacje między elementami <i>as a result of, it was found that, these results suggest that</i></p> <p><b>Struktury</b> Markery tekstowo-refleksyjne, które organizują fragmenty dyskursu lub kierują czytelnika w inne miejsce tekstu <i>in the present study, in the next section, as shown in figure</i></p> <p><b>Obramowanie</b> <i>in the case of, with respect to the, on the basis of, in the presence of, with the exception of</i></p>	
--	---	--

Przegląd danych powyższej klasyfikacji pozwala stwierdzić, że wiązki leksykalne w istotny sposób wpłynęły na przegląd wielowrazowych ciągów znaczeniowych, a tym samym okazały się przydatne w badaniu organizacji tekstu i rozróżnień między dyskursami, biorąc pod uwagę nierodzimych i rodzimych użytkowników języka [232].

#### 4.6. Ekscerpowanie wiązek leksykalnych

W procesie badawczym, w wyniku ewolucji zagadnienia poddanego analizie, pojawiła się potrzeba poszukiwania coraz efektywniejszych narzędzi korpusowych. Te stawały się coraz bardziej skomplikowane, ale z czasem, wraz z rozwojem form oraz metod badawczych, znacznie łatwiej było identyfikować frazy, które mogły okazać się przydatne w procesie rozwoju języka. Jednym z podstawowych parametrów analizy

tekstu i wyodrębniania wiązek leksykalnych jest frekwencja związana z tendencją do pojawiania się elementów leksykalnych. Warto zauważyć, że powszechnie istnieją różnice między tym, co ludzie wiedzą na temat frekwencji, a frekwencją ustaloną za pomocą narzędzi statystycznych. Kjellmer (1994) na początku Dictionary of English Collocations (DEC) wspomina, że jedną z ich kluczowych cech jest fakt, że wiązki leksykalne są często używane w języku:

Rozpoznajemy słowo jako wiązki leksykalne, ponieważ wielokrotnie je słyszeliśmy lub widzieliśmy. Jest więc naturalne, że frekwencja ich występowania powinna być jednym z decydujących czynników. Jednocześnie często uderza fakt, że nawet dobrze ugruntowane w języku kolokacje mogą pojawiać się stosunkowo rzadko w sporych próbkach pisma [151].

Dlatego w DEC Kjellmer zdecydował się potraktować cechę powtarzalności jako warunek konieczny do uznania przedmiotu językowego za kolokację [151]. Z kolei Gablasov, Brezina i McEnery (2017) podkreślają, że przy identyfikacji wiązek leksykalnych należy również brać pod uwagę odległość między współwystępującymi słowami oraz pożądaną bliskość jednostek. Można to osiągnąć za pomocą trzech podejść opartych na sieciach kolokacji, kolokacjach i wreszcie n-gramach (z uwzględnieniem wiązek leksykalnych, koncgramów, collgramów i klasterów) [49], [94].

Sieci kolokacji [49], [213], [289] łączą skojarzenia identyfikowane za pomocą wspomnianego wyżej podejścia w celu uwypuklenia powiązań między słowami, które istnieją nie tylko w całym języku, ale również i w obrębie konkretnego dyskursu. Ostateczne podejście, podejście n-gramowe, może zidentyfikować sąsiednie kombinacje, takie jak bigramy (kombinacje dwuwyrzowe) – *of the, minor changes, I think* – lub te z możliwą zmiennością wewnętrzną – *minor but important/significant/observable changes* [49].

Częstotliwość występowania struktur językowych jest bez wątpienia zmienną o dużym znaczeniu w LLR (Language Learning Research) i ma istotne powiązania z procesami psycholingwistycznymi oraz neuropsychologicznymi, które odgrywają główną rolę w nauce zarówno języka ojczystego, jak i języka obcego [84], [40], [226]. Jeśli chodzi o przyswajanie i tworzenie języka, mówcy L1 i L2 są wrażliwi na kombinacje słów, które występują w określonej częstotliwości [84], [98], [257], [294].

Niezbędny jest jasny obraz roli frekwencji występowania wiązek i aby to osiągnąć, należy również uwzględnić dyspersję tych wiązek w danym korpusie [49], [106], [107]. Do opisu wiązek leksykalnych można wymienić różne narzędzia statystyczne. Jednym z nich jest tzw. mutual information (MI). MI score – pomiar siły frekwencji wiązek [125] jest związany ze spójnością [85], ścisłością [98] oraz trafnością [253] kombinacji słów.

Według Hunstona (2002) wynik MI jest znormalizowanym wynikiem, który można porównywać w różnych korpusach językowych [125]. Nie działa w skali, która faktycznie ma teoretyczne minimum lub maksimum, więc nie jest skalowana do określonego przedziału wartości, ale im większa wartość, tym bardziej ściśle powiązane są ze sobą dwa słowa, a tym samym rzadsza kombinacja. Stwierdzono, że to narzędzie zdecydowanie wymienia terminy, nazwy oraz specjalistyczne lub techniczne kombinacje słów o niskiej częstotliwości. Z tego powodu podkreśla wiązki, które nie są równomiernie rozmieszczone w języku, ponieważ te elementy o niskiej częstotliwości są zwykle ograniczone do określonych gatunków i tekstów (Gablasov i in., 2017).

T-score jest kolejną wartością statystyczną, która jest obliczana jako skorygowana wartość frekwencji występowania wiązki i od której następnie odejmuje się losową częstotliwość współwystępowania. Ta wartość jest następnie dzielona przez pierwiastek kwadratowy z częstotliwości pierwotnej, aby ostatecznie uzyskać wynik T-score. Głównym problemem związanym z wynikiem T-score jest to, że nie działa on na znormalizowanej skali i jako taki nie może być naprawdę stosowany do bezpośredniego porównywania wiązek leksykalnych między różnymi korpusami [125] lub nawet do ustalenia weryfikowalnego punktu odcięcia wartości dla wyników (Gablasov i in., 2017).

Gablasov (2017) zauważa, że początkowe zliczenia częstotliwości i T-score zależą bezpośrednio od wielkości korpusu. Ponieważ początkowe zliczenia częstotliwości i T-score działają w różnych skalach, nie są one porównywalne między korpusami o różnej wielkości. Bardziej zaawansowany poziom obejmuje normalizację, która wymagałaby dostosowania wartości do jednej skali, wspólnej dla różnych korpusów, a zatem bezpośrednio porównywalnej. Taką znormalizowaną skalę stanowiłyby na przykład wartości procentowe lub względne częstotliwości na milion słów. Najbardziej złożony poziom opiera się na skalowaniu wartości i przekształcaniu ich

w skalę z ustalonym zakresem wartości: „Na przykład współczynnik korelacji ( $r$ ) działa w skali -1 do 1” (Gablasov i in., 2017 : 24). Wyniki takich eksperymentów były obserwowane przez kilku badaczy, w szczególności Hunstona (2002), Conklina i Schmitta (2008) oraz Li i Schmitta (2009) [179].

Następna figura statystyczna jest znana jako Log Dice i uwzględnia średnią harmoniczną dwóch proporcji, które wyrażają tendencję dwóch słów do współwystępowania w stosunku do ich indywidualnej częstotliwości w korpusie [255], [86], [87]. Log Dice jest standaryzowaną miarą i działa na skali ze wstępnie zdefiniowaną wartością maksymalną 14 i jako taka jest bezpośrednio porównywalna w różnych korpusach i podobna do wyniku MI ze względu na to, jak ujawnia wyszukane, ale nie zawsze rzadkie kombinacje słów (Gablasov i in., 2017).

Określony system rankingu wiązek lub klasterów, różniący się od surowego/początkowego rankingu częstotliwości, priorytetuje pewne aspekty, takie jak dostosowana częstotliwość (za pomocą T-score), poziom rzadkości występowania (MI-score) i wyłączność (Log Dice).

## KONKLUZJE:

Rozdział czwarty rozpoczyna krótkie wprowadzenie do frazeologii. Dział językoznawstwa, o którym mowa, uczeni definiują różnie, w zależności od poddanych analizie jednostek frazeologicznych. Hyland (2008) nazywa je na przykład wyrażeniami skonwencjonalizowanymi, O’Keeffe i wsp. (2007) nazywają je kolokacjami, Sinclair i Mauranen (2006) kawałkami (chunks), a Biber i wsp. (2004) formułami, prefabrykowanymi wzorami, prefabrykatami i procedurami, jednostkami wielowyrazowymi, frazami leksykalnymi, blokami tekstu i wreszcie wiązkami leksykalnymi/ klasterami. Ostatni z wymienionych terminów jest używany jako podstawowy w niniejszej rozprawie.

Biber i wsp. (2005) twierdzą, że istnieje wiele różnic dotyczących kombinacji słów: cele, wymagania dotyczące identyfikacji, cechy, ilość użytych przykładów, przepisy dotyczące rejestrów. Cowie i Howarth (1996) wskazują na istotną różnicę w podejściu do frazeologii, ponieważ wykazuje ona tendencję do skupiania się na jednym elemencie mieszczącym się w zakresie spektrum frazeologicznego, takim jak idiomy, kolokacje czy inne części mowy [67].

Pierwsza kategoria to jednostki wielowyrazowe, które były różnie definiowane. Różni badacze proponowali różne nazewnictwo, zaczynając od “prefabricated patterns” [112]; “routine formulae” [64]; “lexical phrases” [194]; “chunks” [75]; “lexicalized stems” [209] or “formulaic sequences” [294], [239], i wreszcie bardziej popularne nazwy, takie jak “fixed expressions”, “(restricted) collocations”; or “multi-word units/expressions” [201].

Granger (1998) oraz Pawley i Syder (1983) opisali jednostki wielowyrazowe jako wyzwanie nawet dla uczniów z wysokim poziomem rozumienia języka, jeśli chcą brzmieć naturalnie, być jak natywni mieszkańcy. Drugim elementem są kolokacje, które określa się jako częściowo kompozycyjne. Taką definicję podali Cowie i Howarth (1996), którzy zgadzają się, że znaczenie każdego elementu wchodzącego w skład kolokacji jest określone przez inny element. Jest też kilka typowych połączeń. Hausmann i Blumenthal

(2006) zauważyli na przykład, że wyrażenia rzeczownikowe odnoszą się do zjawisk opisujących świat, podczas gdy czasowniki i przymiotniki zwykle łączą się z przysłówkami [104], [117].

Bardziej współczesne podejście, zaprezentowane przez L'Homme i Bertranda (2000), zapewnia kategoryzację kolokacji na trzy grupy według ich znaczenia: grupa osłabiona, grupa figuratywna i grupa wyspecjalizowana [171]. Taksonomia L'Homme i Bertranda jest podobna do stosowanej przez Cowie i wsp. (1983) - z jedną różnicą. Kryterium Cowie zakłada kolokację, w której o znaczeniu jednego elementu kolokacji nie decyduje znaczenie drugiego elementu. Z kolei Howarth (1998) przedstawił klasyfikację kolokacji według luźnych kryteriów, dzieląc je na kolokacje niezmiennie, kolokacje z ograniczonym wyborem w jednym punkcie i dwóch punktach oraz kolokacje nakładające się [124].

Podkategoria Howartha (1996) nie jest jedyną prezentowaną przez niego. Kolejnym podziałem jest ścisły i sugerowany podział kolokacji według stopnia idiomatyczności. W tej sytuacji można wymienić cztery kategorie: kolokacje swobodne, kolokacje ograniczone, idiomy figuratywne oraz idiomy czyste [123].

Mówiąc o frazeologii i wiązkach wielowyrazowych należy przedstawić pewne podejścia do frazeologii. Często mówi się, że tradycja frazeologiczna rozpoczęła się od frazeologii rosyjskiej i niemieckiej [104]. Z biegiem czasu analizie poddano większość języków europejskich, w tym na przykład węgierski [119], fiński [154], słowacki [82], niderlandzki [214], rosyjski [79], francuski [282], litewski [52], rumuński [297], japoński [281] i hiszpański [215]. Corpas Pastor (2016) oraz Granger i Meunier (2008) wybierają polskie i rosyjskie podejście jako przedstawicieli grup zachodnio-, południowo- i wschodniosłowiańskich tworzących podstawy frazeologii. W rezultacie na potrzeby niniejszej rozprawy wybrano perspektywę polską i rosyjską, które stanowią podstawę dla późniejszych założeń frazeologii jako działu lingwistyki stosowanej [61].

Jeśli chodzi o polską perspektywę w kontekście omawianego zagadnienia, za istotną w zakresie frazeologii uważa się analizę wstępną przeprowadzoną przez Stanisława Skorupkę (1967) i Witolda Doroszewskiego (1969-1978). Uчени ci ustalili definicje frazeologii i wprowadzili prostą taksonomię formalną i semantyczną jednostek

frazeologicznych, co znacząco wpłynęło na późniejsze badania leksykografii, leksykologii i stylistyki [81].

Skorupka (1967) zaproponował klasyfikację gramatyczną dzielącą polskie pozycje frazeologiczne na frazy nominalne, frazy czasownikowe i zdania. Ponadto wymieniono kilka podtypów tych kategorii: wyrażenia szeregowo, wyrażenia porównawcze, wyrażenia figuratywne i wyrażenia rymowane [254]. Skorupka reprezentuje tzw. analityczne podejście do frazeologii, które podkreśla semantyczne aspekty jednostek frazeologicznych, co skutkuje wskazaniem innej typologii dzielącej jednostki frazeologiczne na neutralne, stałe i zmienne. Pierwsza kategoria odnosi się do jednostek (zwykle kolokacji), w których każdy element ma swoje unikalne znaczenie, a znaczenie to może się zmieniać wraz ze zmianą jednego z elementów. Stałe jednostki frazeologiczne, jak sama nazwa wskazuje, zawierają składniki, których nie można zmienić ani wymienić. Ostatnia kategoria, luźne jednostki frazeologiczne, reprezentuje semantyczne powiązania elementów leksykalnych. Niektóre jednostki, określane głównie jako zwroty wolne lub kolokacyjne, mogą być dalej przekształcane w frazy stałe. Alternatywnie te same frazy mogą być stałe lub wolne, w zależności od współwystępującego kontekstu.

Skorupka skoncentrował się na semantycznych aspektach fraz, co potwierdziły opracowania jednostek frazeologicznych Andrzeja Bogusławskiego (1976). Frazom analizowanym przez Bogusławskiego często brakuje semantycznej stałości i często są niekompletne składniowo. Jego podejście jest bardziej dystrybucyjne niż analityczne, polegające na frekwencji występowania wiązek leksykalnych. Warty uwagi jest fakt, że wykazuje ono podobieństwo do koncepcji wiązek opisanych przez Bibera i wsp. (2003). W niniejszym rozdziale przedstawiono również typologię polskich jednostek frazeologicznych Lewickiego (1983) z podziałem na wyrażenia, frazy werbalne i markery frazeologiczne. Wszystkie mogą być jednostkami gramatycznie kompletnymi lub fragmentarycznymi [167].

W swojej frazematyce Wojciech Chlebda (1991) zaproponował pojęcie frazemów. Jak podkreśla Bączkowska (2009), Chlebda (1991) opisał frazem jako jednostkę frazeologiczną powtarzaną w wielu kontekstach sytuacyjnych i dalej przebudowywaną z gotowych już fraz, które pozostają niezmiennie, ale modyfikowalne. Słowa, które

współwystępują we frazach, nie zależą od paradygmatu łączności semantycznej, ale od paradygmatu odtwarzalności w konkretnej sytuacji. Tym samym teoria Chlebudy skłania się ku rozkładowi jednostek frazeologicznych, co jest typowe dla badań opartych na danych korpusowych, gdzie naczelną zasadą klasyfikacji jest liczenie częstości występowania konkretnych wiązek słów [18], [59].

W konsekwencji powstała kategoryzacja na podstawie częstotliwości występowania, która ukazuje zjawiska oznaczone przez Biber et al. (2004) jako wiązki leksykalne - taką nazwę nadano również jednostkom frazeologicznym. Niekompletność syntaktyczna i wzajemne oddziaływanie na poziomie semantycznym, jaki wykazują, pozostają na bardzo niskim poziomie. Z reguły wiązki leksykalne zaczynają się od przyimka lub przedimka. Podobnie mogą kończyć się innym przyimkiem lub przedimkiem. Ogólnie rzecz biorąc, wiązki o wysokiej konektywności wydają się być rdzeniem badań, ale nie wszyscy autorzy uważają je za reprezentatywne [236], [101].

Historyczna perspektywa odniesienia do polskiej obejmuje rosyjską perspektywę frazeologiczną, w tym klasyfikację jednostek frazeologicznych Winogradowa. Podejściem, które jest postrzegane jako pierwsza kompleksowa klasyfikacja w historii frazeologii, jest podejście Winogradowa. Układ ten charakteryzuje się spójnością semantyczną w obrębie jednostek frazeologicznych i charakteryzuje się wieloaspektową konstrukcją, określaną również mianem podejścia analitycznego. Organizacja Winogradowa precyzyjnie konstruuje relację między znaczeniem wyznaczonym w obrębie całego ciągu frazeologicznego a jego częściami składowymi. Grupowanie jednostek frazeologicznych Winogradowa (1974) dzieli je na cztery sekcje: fuzje frazeologiczne, jednostki frazeologiczne, kombinacje frazeologiczne i wyrażenia frazeologiczne.

Kolejny aspekt opisany w tym rozdziale dotyczy wiązek leksykalnych. Wraz z rozwojem językoznawstwa korpusowego wzrosło zainteresowanie przeglądami wyrażen i fraz. Obecnie jednym z najczęściej wymienianych tematów badawczych dotyczących frazeologii pozostaje jeden typ wielowrazowych typów jednostek frazeologicznych, identyfikowanych jako wiązki leksykalne. Wiazki leksykalne były przedmiotem badań wielu badaczy zajmujących się kwestią kilku rodzajów dyskursu.



Do tej pory wiązki leksykalne były badane w pismach naukowych [30], [127], [247], [18] oraz tekstach pisanych przez nierodzimych użytkowników języka angielskiego w kontekście różnych języków [63], [148], [240], [137].

Oprócz dyskursu akademickiego wiązki leksykalne były również badane w innych typach tekstów. Analizie poddano różne obszary językowe, w tym dyskurs polityczny ([206] dokumenty UE [130] teksty związane z telekomunikacją [204], dyskurs prawniczy [100], [37], [153] wraz z dyskursem farmaceutycznym [102]. Najcenniejsza dla tej rozprawy jest jednak analiza dyskursu medycznego [18].

W niniejszym rozdziale przedstawiono liczne definicje wiązek leksykalnych. Pierwsza została przedstawiona przez Bibera (1999) w *Longman Grammar of Spoken and Written English*. Zjawisko to było później badane przez wielu uczonych, na przykład Conrada, Bibera, Leech, Finegana i Johanssona, którzy zaoferowali swoje szeroko zakrojone i korpusowe badania gramatyki angielskiej [162]. Definicja wiązek leksykalnych ustalona przez Bibera i wsp. (1999) wskazywała na wysoką tendencję do współwystępowania. Cortes (2004) oraz McEnery i Hardie (2012) zgodzili się z tym stwierdzeniem, dodając, że wiązki leksykalne mogą być współwystępującymi jednostkami frazeologicznymi w obrębie określonego kontekstu. W epoce językoznawstwa korpusowego dostęp do obszernych tekstowych baz danych ujawnił, że pewne frazeologizmy pozostają semantycznie spójne, a także powtarzalne i składnie dające się łączyć. Jeśli chodzi o frazeologię analityczną, jak wskazał Chlebda (1991), czynnik stopnia powiązania pozostaje w istocie podstawowym elementem analizy i klasyfikacji. Zjawisko to ilustrują wspomniane w niniejszym rozdziale prace Winogradowa, które są cenną informacją dla celów niniejszej rozprawy [59], [63].

Poza definicją wymieniono bardzo ważny aspekt analizy, głównie frekwencję wiązek leksykalnych. Badane przez wielu badaczy [127], [247] najmniejsze występowanie wiązek leksykalnych obserwuje się w wiązках rozszerzonych. Uwagi w tej obserwacji oparte są na ilustracjach znalezionych w konwersacjach akademickich w *Longman Spoken and Written English Corpus*, w którym czterowyrazowe wiązki leksykalne występują około dziesięciokrotnie rzadziej niż trzywyrazowe wiązki. Podobnie wiązki

leksykalne składające się z pięciu wyrazów są około sto razy rzadziej spotykane niż wiązki składające się z trzech wyrazów.

Biber i wsp. (1999) stwierdzają, że zaledwie 15% wiązek leksykalnych używanych w rozmowie i tylko 5% stosowanych w dyskursie akademickim może być postrzeganych jako kompletne jednostki strukturalne. Ponadto Biber i wsp. (1999) twierdzą, że wiązki leksykalne można podzielić na kilka kategorii. Wiązki leksykalne stosowane w dyskursie mówionym są powszechnie uważane za przypadek i należą do kategorii zaimiek + czasownik + dopełnienie. Co więcej, krótkie wiązki leksykalne są zwykle częściami dłuższych ciągów. Jak stwierdza Biber (1999), wiązki leksykalne w prozie akademickiej stanowią z reguły fragmenty fraz przyimkowych lub fraz rzeczownikowych.

Biber, Conrad i Cortes (2003, 2004) również dokonali kategoryzacji wiązek leksykalnych według ich funkcji, dzieląc je na wyrażenia postaw- ang. stance expressions, organizatorów dyskursu- ang. discourse organisers, wyrażenia referencyjne- ang. referential expressions, specjalne funkcje konwersacyjne- ang. special conversational functions. Ta funkcjonalna klasyfikacja wiązek leksykalnych została następnie rozwinięta przez innych badaczy, m.in. Hylanda (2008b), który zmodyfikował kategoryzację Bibera (2004, 2004).

Po zakończeniu podziału i wdrożeniu typologii należy wziąć pod uwagę wyodrębnianie wiązek leksykalnych. Wraz ze wzrostem złożoności urządzeń korpusowych, identyfikacja użyteczności tych fraz dla badaczy zajmujących się analizą języka zdecydowanie się zmniejszyła. Warto wspomnieć o pracy badawczej Gablasova, Breziny i McEnery (2017), którzy zaproponowali trzy podejścia niezbędne do ekstrakcji wiązek leksykalnych: sieci kolokacyjne, kolokacje i n-gramy [94].

Bez wątplenia występowanie struktur językowych jest zmienną o istotnym znaczeniu w LLR (Language Learning Research). Wiąże się z istotnymi skojarzeniami z procesami psycholingwistycznymi i neuropsychologicznymi, które przewodzą w nauce języka [84], [40], [226]. Badanie korpusów ujawnia, że często nawet dość powtarzające się wiązki leksykalne istnieją tylko w określonym kontekście lub są tworzone przez niewielką liczbę osób [49], [107]. Drugi czynnik dotyczy wyłączności wiązek leksykalnych, czyli stopnia, w jakim współwystępuje para wyrazów. Jest to powszechnie określane jako

związek między liczbą ich wspólnych wystąpień zestawionych z liczbą przypadków, w których działają osobno w danym korpusie, co można zmierzyć stosując wskaźnik [MI score]. Ostatnie założenie sugerowane przez Brezinę i wsp. (2015) i Gries (2013) to kierunkowość wiązek leksykalnych, która wymaga, aby elementy wiązki leksykalnej nie przyciągały się nawzajem w równym stopniu, ale raczej by przyciąganie było nieproporcjonalne [107].

W celu wyjaśnienia jednostek wielowyrazowych lub pakietów leksykalnych można zastosować różne narzędzia statystyczne: MI score, T score, którego można zastosować do prostego porównania wiązek leksykalnych w różnych korpusach [125] oraz LogDice, które pomagają zidentyfikować skłonność dwóch słów do współwystępowania w odniesieniu do ich częstotliwości w korpusie.

Wszystkie przedstawione w tym rozdziale badania porównawcze [27] wykazały, że w dyskursie akademickim o różnych specjalnościach dominują wiązki leksykalne. W przypadku wiązek identyfikowanych w tekstach akademickich można zauważyć przewagę tych, które składają się z dwóch elementów: frazy rzeczownikowej lub części frazy przyimkowej. Dlatego większość z tych wiązek ma na końcu słowo funkcyjne, które zwykle jest przyimkiem lub przedimkiem. Kolejnymi wiązkami są frazy przyimkowe pełniące funkcję markerów dyskursu, co znacząco poprawia styl wypowiedzi, zarówno w komunikacji pisemnej, jak i mówionej, czego dowodzą Flores (2016) oraz Shakarami i wsp. (2016) [246]. O występowaniu wiązek leksykalnych w dyskursie akademickim wspominają Biber i wsp. (1999) oraz Hyland (2008b), który ze swojego korpusu akademickiego wydobyl 240 wiązek leksykalnych z artykułów naukowych, prac magisterskich i doktorskich z czterech dyscyplin, w celu uzyskania wiedzy o dyscyplinarnych zróżnicowaniach ich częstotliwości oraz preferowanych zastosowaniach. Badanie dowodzi, że wiązki te mają nie tylko fundamentalne znaczenie dla tworzenia gatunków akademickich, ale oferują również istotny sposób różnicowania tekstów pisanych według dyscypliny naukowej. Wyniki badań potwierdzili także inni badacze Biber, Conrad i Cortes (2003) oraz Biber, Conrad i Cortes (2004).

## ROZDZIAŁ V

### CEL BADAŃ I HIPOTWEZY BADAWCZE

#### 5.1. Cel badań

Celem pracy jest analiza wyników testów fluencji słownej (fonologicznej i kategoryalnej) u osób z zaburzeniami funkcji poznawczych na poziomie łagodnych zaburzeń poznawczych i wczesnego stadium choroby otępiennej (prawdopodobnej choroby Alzheimera) w kontekście analizy korpusowej według Narodowego Korpusu Języka Polskiego.

Analizie zostaną poddane wyniki badań testu fluencji słownej fonologicznej (kryterium literowe) i kategoryalnej (kryterium „zwierzęta” i „przedmioty ostre”) w tej grupie badanych w kontekście ogólnego funkcjonowania poznawczego ocenianego skalą przesiewową dla otępienia (MMSE). Za pomocą narzędzi do analizy danych korpusowych zbadane zostaną najczęściej występujące słowa wyekscerpowane na podstawie wykonanego przez pacjentów testu fluencji słownej oraz analiza ubytku słownego w różnych stadiach zaburzeń poznawczych.

Dotychczasowe badania analizy fluencji słownej dotyczyły głównie grup chorych z różnymi dysfunkcjami, np. z uszkodzeniem prawej, lewej lub obu półkul mózgu [270], stosowane są w diagnozie różnicowej zaburzeń neurologicznych u dzieci i młodzieży (Biechowska, Kaczmarek, Witkowska, Steinborn 2012), jak również mają zastosowanie w diagnostyce neuropsychologicznej [44], [216], [217]. Pomimo, że ten obszar badawczy jest od kilku lat systematycznie eksplorowany, nie została przeprowadzona wnikliwa analiza oparta o narzędzia korpusowe uwzględniająca dwa punkty widzenia: neuropsychologiczny oraz lingwistyczny. Niniejsza praca może zatem przyczynić się do częściowego uzupełnienia brakujących informacji.

## **5.2. Hipotezy badawcze**

W pracy przyjęto następujące hipotezy badawcze:

- 1) Stopień dysfunkcji fluencji słownej oraz globalnego osłabienia funkcji poznawczych może wiązać się z czynnikami demograficznymi (wiekiem i płcią badanych) oraz czynnikami klinicznymi (nasilenie zmian chorobowych, czas trwania zaburzeń poznawczych).
- 2) Nieprawidłowości w zakresie fluencji słownej mogą wiązać się z nasileniem zaburzeń poznawczych w różnych obszarach.
- 3) Osłabienie fluencji słownej u badanych może być relatywne do wyników analiz statystyki korpusowej, słowa występujące częściej w języku polskim według Narodowego Korpusu Języka Polskiego mogą mieć charakter bardziej trwałe u osób z zaburzeniami poznawczymi.

## **5.3. Szczegółowe cele badawcze**

1. Ocena zgodności wewnętrznej i analiza składowych głównych stosowanego narzędzia – testu fluencji słownej.
2. Ocena wyników fluencji słownej fonologicznej i kategoryjnej w badanej grupie chorych.
3. Dokonanie porównań wyników kobiet i mężczyzn uzyskanych w poszczególnych podtestach fluencji słownej.
4. Dokonanie analizy korelacyjnej pomiędzy wynikami testu fluencji słownej a wynikami MMSE oceniającego globalne sprawności poznawcze.
5. Dokonanie analizy korelacji wyników fluencji słownej uzyskanych przez badanych z wiekiem i czasem trwania zaburzeń poznawczych.

6. Określenie frekwencji najczęściej występujących słów w poszczególnych podtestach fluencji słownej.
7. Dokonanie analizy korpusowej wyników testu fluencji semantycznej i kategoryalnej.

## ROZDZIAŁ VI

### METODOLOGIA BADAŃ

#### 6.1. Osoby badane: opis badanej grupy i zgoda komisji

Do badania włączono 279 osób , w wieku 50-91 lat, średnia wieku 72,00 + 10,47 lat. W grupie badanej znalazło się 177 kobiet (średnia wieku 72,18+10,52 lat) i 102 mężczyzn (średnia wieku 71,67+ 10,41 lat). Wiek badanych oraz wynik testu MMSE nie różnicowały grupy kobiet i mężczyzn. Analizie lingwistycznej został poddany retrospektywnie materiał naukowy zgromadzony w Katedrze Neuropsychologii Klinicznej w Collegium Medicum im. Ludwika Rydygiera w Bydgoszczy Uniwersytetu Mikołaja Kopernika w Toruniu na które wyraziła zgodę Komisja Bioetyki. Ocenie poddany został Test Fluencji Słownej – wersja z zastosowaniem kryterium formalnego i kategoryjnego, który wykonali pacjenci chorzy na schizofrenię, z MCI i rozpoznaną chorobą Alzheimera.

Najczęściej występujące dysfunkcje to persewacje, które są związane z uszkodzeniem płatów czołowych, najbardziej powszechne zaburzenie przy otępieniu. Kobiety i mężczyźni nie różnili się pod względem wieku, liczbowo było więcej kobiet niż mężczyzn, lecz mężczyźni mieli większą rozpiętość wiekową. Kobiety generowały więcej wyrazów na literę K oraz zwierząt, natomiast mężczyźni mieli znacznie więcej przedmiotów ostrych, lecz istotna różnica dotyczyła pierwszej kategorii wymienienia jak największej ilości zwierząt. Różnice w nazwach własnych wystąpiły tylko w teście literowym. Neologizmy występowały tylko w teście literowym co było statystycznie istotne. U kobiet wystąpiło więcej nazw własnych niż u mężczyzn. W pozostałych parametrach ich wyniki nie różniły się istotnie.

## **6.2. Metody badań**

Badanie procesów werbalnych stanowi istotne kryterium oceny neuropsychologicznej kondycji pacjentów stosowanym w psychiatrii i neurologii. Istnieje szereg testów bazujących na badaniu funkcji językowych. Jednym z najbardziej znanych i powszechnie stosowanych testów jest Verbal Fluency Test (test fluencji słownej), będący przedmiotem niniejszych rozważań. VFT nie narzuca użycia narzędzi ani procedur, które wymagałyby specjalistycznej wiedzy lub umiejętności i można go stosować w badaniu przyłóżkowym.

### **6.2.1. Opis testu fluencji słownej**

Test fluencji słownej jest najczęściej postrzegany przez pacjentów jako mało inwazyjny [259]. Z psychometrycznego punktu widzenia test fluencji słownej umożliwia diagnozę zaburzeń dotyczących całej sfery poznawczej. Służy ewaluacji wielu funkcji takich jak płynność słowna, procesy uwagi, szybkość przetwarzania informacji, pamięć operacyjna i funkcje wykonawcze [234].

Test występuje w wielu wersjach. Thurstone stworzył jego pierwowzór w pierwszej połowie XX wieku. Thurstone's Word Fluency Test polega na generowaniu w formie pisemnej, w określonym czasie, jak największej ilości słów na wskazaną literę i służy do oceny produktywności umysłowej osób z uszkodzeniami ośrodkowego układu nerwowego (OUN) [275].

Zakłada się, że do wykonania testu włączone są funkcje mowy, pamięci i procesy wykonawcze. Test stanowi również ważny indykator funkcjonowania głównie płatów skroniowych i czołowych. Literatura anglojęzyczna omawia szereg istotnych badań dotyczących omawianego zagadnienia. Deficyt polskich norm i certyfikowanych kryteriów dotyczących sposobu oceny wyników przeprowadzanych badań skutkuje zmniejszeniem wartości narzędzia badawczego, w związku z powyższym wymaga szczególnej ostrożności oraz sumienności i precyzji w ocenie i ich opisie.



Test fluencyjny polega na wymienieniu jak największej ilości słów zgodnie z określonym kryterium, w określonym czasie (najczęściej przyjmuje się czas 60 sekund). Test fluencji słownej pozwala na zdiagnozowanie i zbadanie płynności słownej semantycznej, zwanej również kategoryalną, lub też literowej, zwanej fonemiczną, o których mowa w sekcji 2.1 w rozdziale II.

Jak wspomina Piskunowicz, Bieliński, Zgliński oraz Borkowska (2013) płynność słowna semantyczna (kategoryalna, ang. semantic fluency) określa zdolność przystosowania słów z danej kategorii semantycznej, która może być określana jako obszerna (jak w przypadku kategorii „zwierzęta”) lub wąska (przykładem może być kategoria „przedmioty ostre”). Badanie fluencji literowej (formalnej, ang. phonemic fluency) natomiast odnosi się do generowania słów rozpoczynających się na określoną literę lub głoskę, jak ma to miejsce w przypadku fluencji fonemicznej. Zdaniem Piskunowicz, Bieliński, Zglińskiego oraz Borkowskiej (2013) do najczęściej stosowanych liter w teście fluencji literowej w literaturze polskiej należą: F, A i S.

Wykorzystanie konkretnych liter potrzebnych do przeprowadzenia testu fluencji wiąże się z niską lub wysoką liczbą ich okurencji w obrębie jednego języka. Czas przeznaczony na wykonanie testu wynosi najczęściej 60 sekund. Można wyróżnić kilka fonetycznych wariacji testu, składających się z różnych zestawów liter. Do najbardziej powszechnie używanych zestawów należy wersja oryginalna Thurstone’a zakładająca litery: F, A, S w czy P, R, W oraz C, F, L w wersji Controlled Word Fluency Test (Mitrushina et al., 2005). Inne przykłady zestawień to: P, D, S, T w Test of Verbal Conceptualization and Fluency czy F, A, S oraz B, H, R w Delis–Kaplan Executive Function System [12]. W badaniach na gruncie polskim często wskazuje się litery K, M, P, S [132], [269].

Jak podaje Piskunowicz, Bieliński, Zgliński oraz Borkowska (2013) jednym z najpopularniejszych testów na badanie płynności słownej literowej jest anglojęzyczny COWAT (Controlled Oral Word Association Test). Wymienić można również inne wersje testów, które nie są tak powszechnie stosowane. Jedną z nich jest fluencja asocjacyjna, która polega na swobodnych skojarzeniach oraz wymaga podawania czasowników lub innych słów o nasyceniu emocjonalnym. Kolejną wersją testu jest fluencja afektywna

polegająca na wymienieniu egzemplarzy należących do kategorii słów z podziałem na przyjemne lub nieprzyjemne [216].

Standardowe rozpoznanie i interpretacja wyników ma najczęściej charakter ilościowy, w związku z tym charakteryzuje słowa zgodne z przyjętym kryterium oraz błędy. Za błędne uznaje się odpowiedzi nieodpowiednie do danej kategorii, neologizmy oraz persewacje. Istotny okazuje się współczynnik wykonania zadania, jest to czas w którym respondent generuje swoją pierwszą odpowiedź oraz długość przerw pomiędzy odpowiedziami. Bousfield oraz Sedgewick (1944) zaobserwowali, że badani nie produkują jednorodnych słów, ale wykazują tendencję do sylabizowania słów złożonych z różnej ilości elementów najczęściej odseparowanych pauzami [46].

Troyer i wsp. (1997) zaproponowali analizę wyników biorących pod uwagę wielkość wiązek leksykalnych oraz przełączenia między nimi. Powyższe kategorie uznano za osobne przedmioty badań, które podlegają połączeniom z innymi funkcjami poznawczymi, jak również strukturami OUN. Klaster lub wiązki leksykalne definiowane są jako słowa powiązane ze sobą semantycznie lub fonetycznie w przypadkach, gdy „1) rozpoczynają się tą samą zbitką dwóch spółgłosek (np. frotka, fraż), 2) różnią się tylko samogłoską (np. sok, syk), 3) rymują się (np. bułka, półka), 4) są homofonami (np. spać, spadać)” [279].

Jak twierdzi Gruen i wsp. (1990), kolejna typologia stworzona na podstawie badań pacjentów z urazami czaszkowo-mózgowymi i traktuje o trzech fundamentalnych mechanizmach wiązania słów: „według pierwszej sylaby, przez skojarzenie semantyczne oraz poprzez zlepek pierwszych spółgłosek” [109].

Według Abwendera i wsp. (2001) w teście fluencji w kryterium literowym występują klaster lub wiązki leksykalne zgodne lub niezgodne semantycznie. Według Schwartz i wsp. (2003) pacjenci generując odpowiedzi zgodnie z kryterium literowym stosują głównie strategie semantyczne. Założenie to zostało potwierdzone w polskich badaniach przeprowadzonych przez Szepietowską (2011) na grupie zdrowych pacjentów. Wykazano, że bez względu na wiek respondentów i niezależnie od przyjętego kryterium wykonawczego, strategie semantyczne są znacznie częściej stosowane wśród pacjentów niż fonemiczna organizacja materiału. Klaster

semantyczne okazały się być zdecydowanie większe niż klasterzy fonemiczne, które zawierały mniejszą ilość słów oraz współwystępowały z większą ilością przełączeń [269].

### **6.2.2. Metoda analizy korpusowej**

W niniejszym badaniu została zastosowana korpusowa metoda analizy lingwistycznej polegająca na automatycznej ekscerpcji wyżej wymienionych nazw i pogrupowaniu ich pod względem częstotliwości występowania. Taka forma elektronicznych konkordancji pozwala na szybkie wykrycie słów i wyrażen kluczowych, charakterystycznych dla badanego dyskursu.

W miarę jak narzędzia korpusowe stawały się coraz bardziej wyrafinowane, znacznie łatwiej było identyfikować frazy, klasterzy czy wiązki leksykalne. Jednym z podstawowych parametrów analizy tekstu i wyodrębniania wiązek leksykalnych lub pojedynczych egzemplarzy jest częstotliwość związana z tendencją do pojawiania się elementów leksykalnych. Warto zauważyć, że powszechnie istnieją różnice między tym, co ludzie wierzą na temat częstotliwości, a częstotliwością ustaloną za pomocą narzędzi korpusu statystycznego (opisanych w rozdziale 3).

Kjellmer (1994) wspomina na początku Dictionary of English Collocations (DEC) iż należy potraktować cechę powtarzalności jako warunek konieczny do uznania przedmiotu językowego za kolokację [151]. Krenn (2000) stwierdza, że wyniki badań potwierdzają „słabość miar statystycznych w zakresie identyfikacji wiązek leksykalnych z danych o wysokim udziale egzemplarzy o niskiej częstotliwości występowania” i uważa, że prowadzi to do wniosku, że należy opracować metody które pozwalają na identyfikację wiązek na podstawie danych o niskiej częstotliwości występowania, ponieważ okurencja tej grupy słów przeważa w korpusach językowych [156].

Z kolei Gablasov, Brezina i McEnery (2017) podkreślają, że przy identyfikacji wiązek leksykalnych należy również brać pod uwagę odległość między współwystępującymi słowami oraz pożądaną bliskość jednostek. Można to osiągnąć za pomocą identyfikacji wiązek leksykalnych, lub też n-gramach, czyli pojedynczych egzemplarzy.

Pierwsza metoda ekscerpcji to metoda okna kolokacyjnego (ang. collocation window approach), wyszukuje koincydencje w określonym przedziale okiennym, na przykład pięć słów na lewo od badanego słowa i pięć na prawo – w skróconym do 5L 5R. Pozwala ona na identyfikację luźniejszych skojarzeń słownych niż podejście n-gramowe opisane w niniejszej dysertacji [94].

Sieci kolokacji (ang. collocation networks) [49], [213], [289] łączą skojarzenia identyfikowane za pomocą wspomnianego wyżej podejścia okna kolokacyjnego w celu uwypuklenia powiązań między słowami, które istnieją w konkretnym języku i dyskursie. Ostatnia metoda opisuje podejście n-gramowe, która polega na identyfikacji sąsiednich kombinacji, takich jak bigramy (kombinacje dwuwyrzowe) czy też n-gramy (pojedyncze jednostki znaczeniowe) [94].

Częstotliwość struktur językowych jest bez wątpienia zmienną o dużym znaczeniu w badaniach nad językiem LLR (ang. Language Learning Research) i ma istotne powiązania z procesami psycholingwistycznymi i neurolingwistycznymi, które odgrywają główną rolę w nauce języka oraz jego tworzeniem języka [84], [40], [226].

Częstotliwość surowa (bezwzględna) jest rozsądną miarą ogólnej powtarzalności języka i została zastosowana w LLR opartym na korpusie. Jednak może nie być najdoskonalszą metodą przewidywania regularności i przewidywalności w użyciu języka. Analizy korpusów pokazują, że często nawet dość częste kookurencje występują tylko w bardzo konkretnym kontekście lub są tworzone przez stosunkowo niewielką liczbę respondentów.

Niezbędnym jest, aby uwzględnić dyspersję wiązek w danym korpusie [49], [106], [107]. Rozproszenie cechy językowej odnosi się do częstotliwości występowania danego egzemplarza w korpusie językowym. Im rozproszenie jest większe tym rzadsze jest jego występowanie. Jako takie, rozproszenie jest kluczowym predyktorem w nauce języka, ponieważ okurencja bardziej ogólnych wiązek (tj. tych, które występują w różnych kontekstach) jest bardziej prawdopodobna.

Drugim ważnym czynnikiem jest łączliwość wiązek, czyli stopień, w jakim dwa słowa występują w swoim towarzystwie. Jest to zazwyczaj wyrażane jako stosunek między liczbą współwystępowania egzemplarzy, a liczbą przypadków, w których funkcjonują one oddzielnie w danym korpusie (np. wynik wzajemnej informacji [MI]). Istnieje duże prawdopodobieństwo, że łączliwość jest silnie powiązana z przewidywalnością współwystępowania, to znaczy, że pojawienie się jednej części wiązki wywołuje okurencję drugiej części. Przykłady takich wiązek zaczerpnięte z Wray (2002) obejmują okey dokey czy annus mirabilis [294].

Trzecim wymiarem wartym rozważenia jest kierunkowość [49], [107], która zakłada, że elementy wiązki nie przyciągają się jednakowo, ale przyciąganie jest asymetryczne. W związku z tym możemy być w stanie przewidzieć jedno słowo na podstawie drugiego, które tworzy parę, ale nie na odwrót. Doskonałym przykładem jest słowo dekoracje, które skłania rozmówców do myślenia o Bożym Narodzeniu (w BNC 11% przypadków dekoracji jest w rzeczywistości poprzedzonych Bożym Narodzeniem, w NKJP jest to 8%) [107].

Do opisu wiązek leksykalnych czy klasterów można wymienić różne narzędzia statystyczne. Jednym z narzędzi jest tzw. wzajemna informacja (ang. MI score). Wzajemna informacja (MI) jest związana ze spójnością [85], ścisłością [98] oraz trafnością [253] kombinacji słów. Bestegen i Granger (2014) zauważyli, że faworyzuje on pozycje o niskiej częstotliwości i został skonstrastowany z wynikiem T-score w celu pomiaru pozycji o wysokiej częstotliwości [22], [104], (Gablasov et al., 2017).

Według Hunstona (2002) wynik MI jest znormalizowanym wynikiem, który można porównywać w różnych korpusach językowych [125]. Nie działa w skali, która faktycznie ma teoretyczne minimum lub maksimum, więc nie jest skalowana do określonego przedziału wartości, ale im większa wartość, tym bardziej ściśle powiązane są ze sobą dwa słowa, a tym samym rzadsza kombinacja. Stwierdzono, że ten wynik zdecydowanie faworyzuje terminy, nazwy oraz specjalistyczne lub techniczne kombinacje o niskiej częstotliwości. Z tego powodu podkreśla pakiety, które nie są równomiernie rozmieszczone w języku, ponieważ te elementy o niskiej częstotliwości są zwykle ograniczone do określonych gatunków i tekstów (Gablasov i in., 2017).

Gablasov (2017) zauważa, że surowe zliczenia częstotliwości i T-score zależą bezpośrednio od wielkości korpusu. Ponieważ surowe zliczenia częstotliwości okurencji i T-score działają w różnych skalach, nie są one porównywalne między korpusami o różnej wielkości. Bardziej zaawansowany poziom obejmuje normalizację, która wymagałaby dostosowania wartości do jednej skali, wspólnej dla różnych korpusów, a zatem bezpośrednio porównywalnej. Taką znormalizowaną skalę stanowiłyby na przykład wartości procentowe lub względne częstotliwości na milion słów. Najbardziej złożony poziom opiera się na skalowaniu wartości i przekształcaniu ich w skalę z ustalonym zakresem wartości: „Na przykład współczynnik korelacji (r) działa w skali od -1 do 1” (Gablasov i in., 2017 : 24). Wyniki takich eksperymentów były obserwowane przez kilku badaczy, w szczególności Hunstona (2002), Conklina i Schmitta (2008) oraz Li i Schmitta (2009) [172].

Z przedstawionej powyżej analizy jasno wynika, że im wyższa częstotliwość, tym silniejsza wiązka leksykalna i jej odmiany. W kolejnej części rozprawy skupiono się na częstości występowania n-gramów.

### **6.3. Metody analizy statystycznej**

Charakter rozkładu zmiennych oceniano za pomocą testy Shapiro Wilka. W zależności od charakteru rozkładu zmiennych stosowano testy parametryczne lub nieparametryczne. Ponieważ większość zmiennych miało rozkład odbiegający od normalnego, w analizach stosowano testy nieparametryczne. W ocenie zgodności wewnętrznej stosowanych narzędzi obliczano wskaźnik alpha-Cronbacha, wykonano analizę czynnikową składowych głównych. W porównaniu wyników dwóch grup stosowano test U-Manna Whitneya. Za znamienność statystyczną różnic pomiędzy grupami przyjęto  $p < 0,05$ . W analizie korelacji użyto testu R-Spearmana. W analizie predykcji wykorzystano test jednoczynnikowej ANOVA.

## ROZDZIAŁ VII

### WYNIKI BADAŃ

#### 7.1. Analiza spójności wewnętrznej stosowanych narzędzi

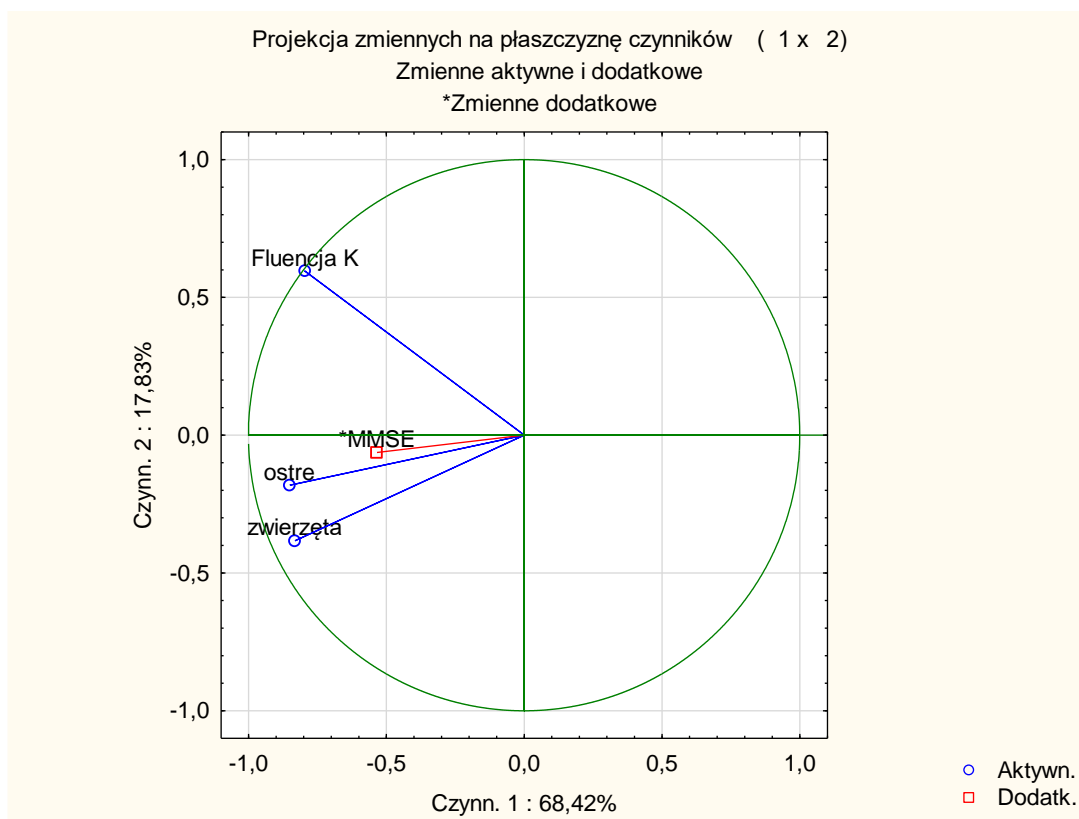
W pierwszej fazie analiz dokonano analizy spójności wewnętrznej stosowanej skali Fluencji słownej. W tabeli IV przedstawiono interkorelacje wewnętrzne poszczególnych podskal testu fluencji słownej. Tabela V prezentuje wyniki analizy czynnikowej składowych głównych testu fluencji słownej oraz testu MMSE, a rycina 1 projekcję przestrzenną składowych głównych testu fluencji. Jak wynika z tabeli V, wszystkie wymiary skali fluencji oraz MMSE wyjaśnia jeden czynnik. Oznacza to, że badane parametry mają podobny mechanizm etiologiczny.

Tab. IV. Interkorelacje podskal testu Fluencji Słownej

	Test literowy	zwierzęta	ostre
test literowy	1	0,48	0,52
zwierzęta		1	0,56
ostre			1

Tab. V. Analiza czynnikowa składowych głównych testu fluencji słownej.

	Oznaczone składowe główne istotne na poziomie alpha >0,7
Test literowy	0,80
zwierzęta	0,82
Przed ostre	0,82
MMSE	0,73
Wariancja wyj.	2,46
Udział	0,61



Ryc. 1. Projekcja przestrzenna czynników głównych testu Fluencji Słownej i MMSE

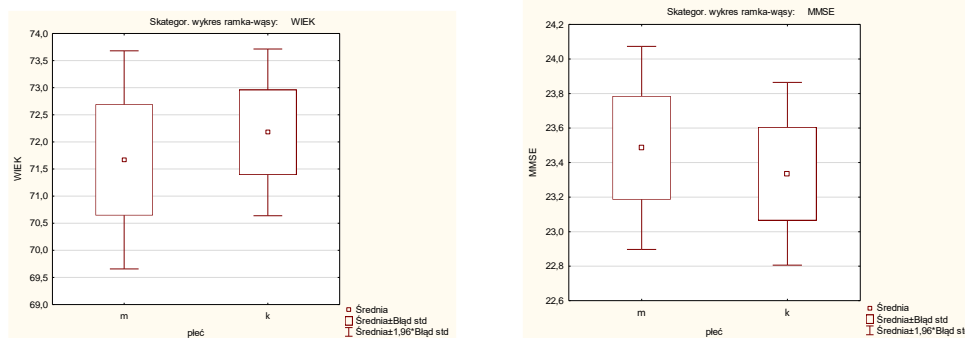


## 7.2. Wyniki testów MMSE i fluencji słownej

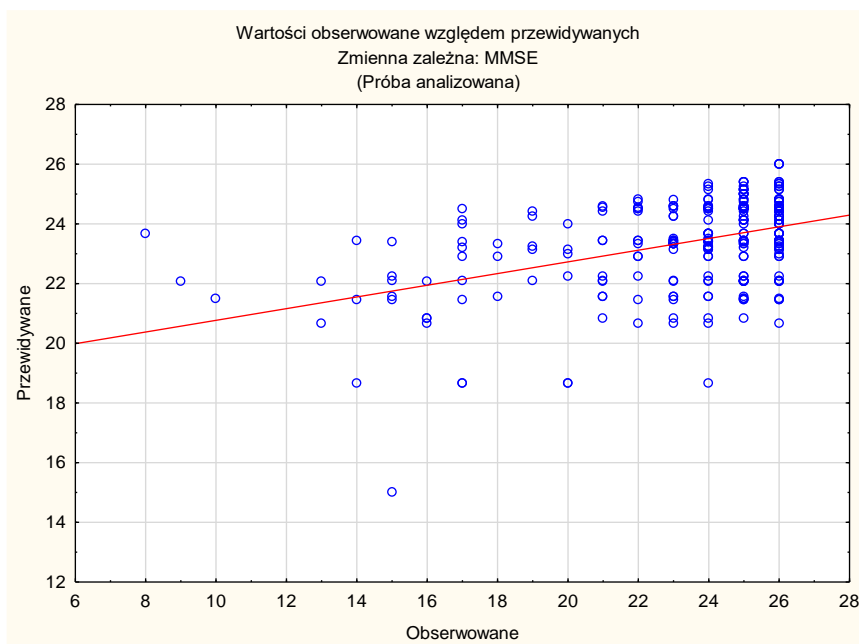
Tabela VI zawiera dane dotyczące wieku oraz wyniku MMSE w badanej grupie osób.

Tab. VI. Wiek i wyniki testu MMSE badanych osób. Porównanie wyników kobiet i mężczyzn.

	Cała grupa	Kobiety	Mężczyźni	Stat. Z (z popr. Lifersona)	Test U Manna Whitneya
wiek	72,00± 10,47	72,18 ± 10,52	71,67 +10,41	-0,38	0,71
MMSE	23,39±3,41	23,34± 3,61	23,49+ 3,04	-0,07	0,94



Ryc. 2. Wiek i wyniki testu MMSE badanych osób. Porównanie wyników kobiet i mężczyzn. Wartości średnie, SD, min-max

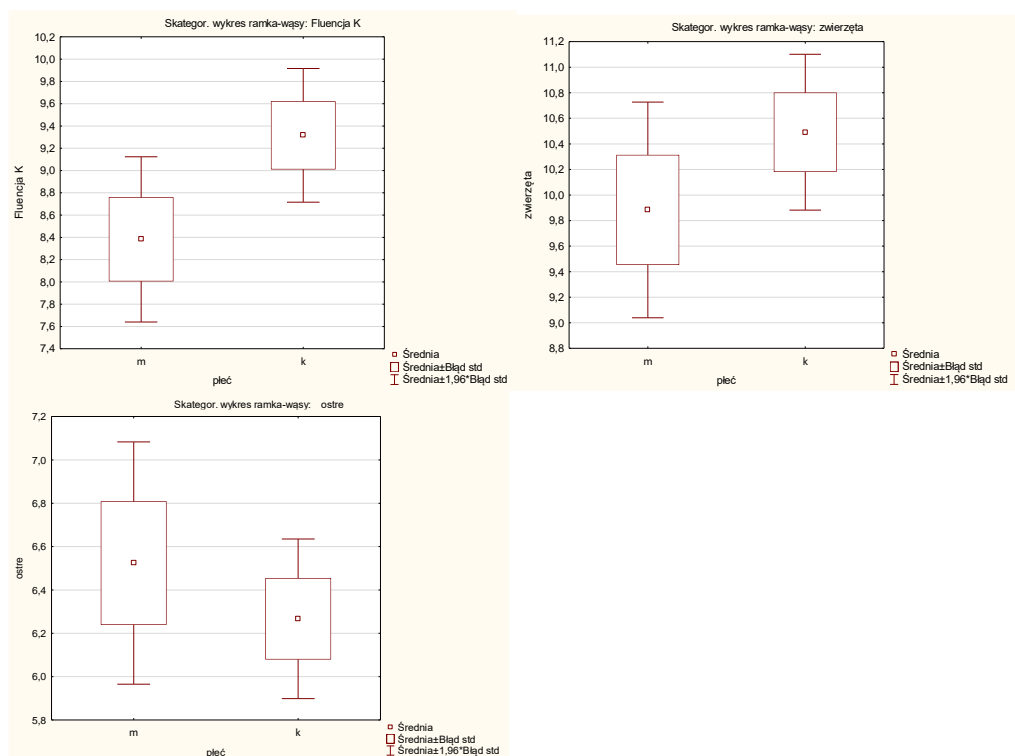


Ryc.3. Rozkład wartości obserwowanych względem oczekiwanych wyników testu MMSE w badanej grupie.

Tabela VII prezentuje wyniki testu fluencji słownej oraz porównanie wyników uzyskanych przez kobiety i mężczyzn, a rycina 4 graficzny rozkład wyników testu fluencji słownej.

Tab. VII. Wyniki testów fluencji słownej. Porównanie wyników kobiet i mężczyzn.

Test Fluencji	Cała grupa		Kobiety N=177	Mężczyźni N=102	Stat. Z (z popr. Lifersona)	Test U Manna Whitneya
	Średnia ±SD	Mediana (25-75%)	Średnia ±SD	Średnia ±SD		
semantyczny - literowy	8,97± 4,00	9,00 6,00-12,00	9,32 ± 2,52	8,38±2,89	-2,18	0,029
kategorialny- zwierzęta	10,27 ± 4,24	10,00 8,00-13,00	10,49 ± 4,16	9,88 ± 4,37	-1,52	0,15
kategorialny przedmioty ostre	6,36 ± 2,66	6,00 5,00-8,00	6,27 ± 2,52	6,52 ± 2,89	0,55	0,58



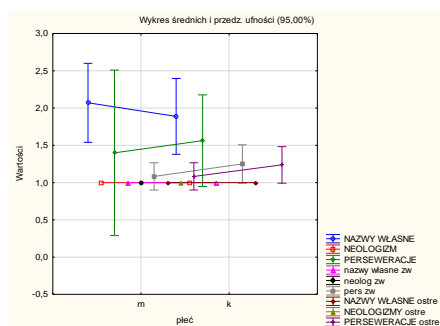
Ryc. 4. Rozkład wyników testów fluencji słownej w badanej grupie. Wartości średnie, SD, min-max

W tabeli VIII i na rycinie 5 zaprezentowano frekwencję nieprawidłowo włączanych słów w poszczególnych testach fluencji, czyli słów niezgodnych z instrukcją, jak nazwy własne, neologizmy i persewacje, które są wskaźnikami patologii w zakresie fluencji słownej. Dokonano również porównania tych parametrów pomiędzy kobietami i mężczyznami. Okazało się, że kobiety, częściej niż mężczyźni, używały nazw własnych, natomiast w pozostałych parametrach grupy nie różniły się istotnie.

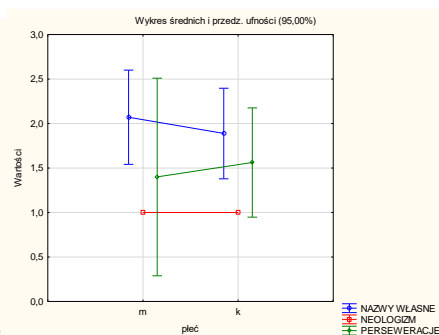
Tab. VIII. Frekwencja nieprawidłowo włączonych słów w poszczególnych testach fluencji słownej

Test fluencji	kategoria	Cała grupa	Kobiety N=177	Mężczyźni N=102	Stat. Z (z popr. Lifersona)	Test U Manna Whitneya
		Średnia $\pm$ SD	Średnia $\pm$ SD	Średnia $\pm$ SD		
test literowy	Nazwy własne	8,97 $\pm$ 4,00	9,32 $\pm$ 2,52	8,38 $\pm$ 2,89	-2,18	0,029
	neologizmy	10,27 $\pm$ 4,24	10,49 $\pm$ 4,16	9,88 $\pm$ 4,37	-1,52	0,15
	persewercje	6,36 $\pm$ 2,66	6,27 $\pm$ 2,52	6,52 $\pm$ 2,89	0,55	0,58
test kateg. zwierzęta	persewercje	1,52 $\pm$ 1,08	1,56 $\pm$ 1,16	1,40 $\pm$ 0,89	-0,17	0,94
test kateg. przedmioty ostre	persewercje	1,18 $\pm$ 0,46	1,08 $\pm$ 0,29	1,08 $\pm$ 0,29	-0,81	0,61

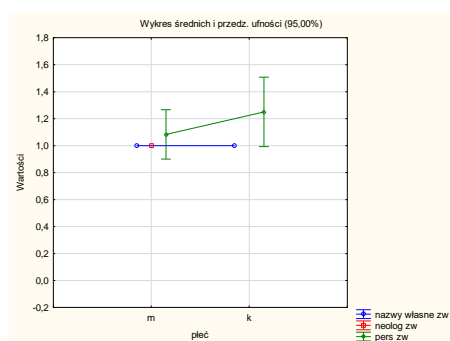
A



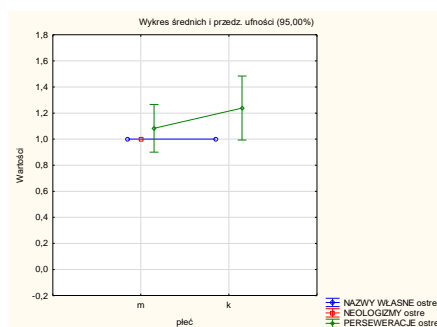
B



C



D



Ryc. 5. Słowa nieprawidłowo włączone w poszczególnych testach fluencji słownej

A: test fluencji ogółem

B: Test fluencji fonologicznej – literowy

C: Test fluencji kategorialnej- zwierzęta

D: Test fluencji kategorialnej – przedmioty ostre

### 7.3. Analiza korelacji

Tabela IX prezentuje wyniki analizy korelacji R-Spearmana pomiędzy wiekiem badanych i czasem trwania zaburzeń poznawczych a wynikami testu MMSE i wykonaniem przez nich testu fluencji słownej.

Tab. IX. Korelacje rang Spearmana pomiędzy wiekiem i czasem trwania zaburzeń poznawczych a wykonaniem testu MMSE i testów fluencji słownej

test	wiek	Czas trwania zaburzeń poznawczych
MMSE	-0,12	-0,12
TF literowy	0,00	-0,36*
TF zwierzęta	-0,10	-0,34*
TF przedmioty ostre	0,02	-0,21*
TF literowy nazwy własne	0,19	0,10
Test literowy persewercje	0,08	0,05
Test kat. zwierzęta persewercje	-0,15	0,01
Test kat. przedmioty ostre persewercje	-0,15	0,00

Nie stwierdzono istotnych korelacji pomiędzy wiekiem badanych a wykonaniem testu MMSE oraz wszystkich trzech części testu fluencji słownej. Podobnie nie stwierdzono korelacji pomiędzy wiekiem a frekwencją występowania nieprawidłowo włączonych słów w teście fluencji. Czas trwania zaburzeń poznawczych korelował z wykonaniem testów fluencji słownej – liczbą generowanych słów. Im dłuższy był czas trwania zaburzeń poznawczych tym mniej słów generowały osoby badane w danej części testu fluencji. Nie stwierdzono natomiast istotnych korelacji pomiędzy czasem trwania zaburzeń a liczbą neologizmów i persewercji, które były sporadyczne.

W tabeli X zaprezentowano korelacje R-Spearmana pomiędzy wynikami testu MMSE a wynikami testów fluencji słownej u badanych.

Tab. X. Korelacje rang Spearmana pomiędzy wynikami testu MMSE i testów fluencji słownej

Test & MMSE	Cała grupa	Kobiety	Mężczyźni
TF literowy	0,29*	0,29*	0,26*
TF zwierzęta	0,44*	0,41*	0,49*
TF przedmioty ostre	0,35*	0,36*	0,33*

\*Korelacje istotne  $p < 0,05$

Jak wynika z powyższej tabeli, stwierdzono istotne dodatnie korelacje pomiędzy wykonaniem testu MMSE i testu fluencji słownej. Oznacza to, że im lepsza była sprawność funkcji poznawczych badanych, ocenianych testem MMSE, tym lepiej jest wykonywany był zarówno test fluencji słownej fonologicznej, jak i kategoryalnej.

#### 7.4. Analiza predykcji – wyniki analizy czynnikowej

W poniższych tabelach (XI-XIII) przedstawiono wyniki analizy czynnikowej, przewidywanie wykonania testów fluencji słownej w zależności od uzyskanego przez badanych wyniku w skali MMSE.

Tab. XI. Wyniki analizy jednoczynnikowej ANOVA – przewidywanie wykonania testu fluencji fonologicznej (test literowy) na podstawie wyniku MMSE w badanej grupie

Jednowymiarowe testy istotności dla Fluencja K (Sheet1 w monika dane) Parametryzacja z sigma-ograniczeniami Dekompozycja efektywnych hipotez; Błąd standardowy oceny:3.6700					
Efekt	SS	Stopnie swobody	MS	F	p
Wyraz wolny	2405,257	1	2405,257	178,5817	0,000000
MMSE	858,664	16	53,666	3,9845	0,000001
Błąd	3515,322	261	13,469		

Tab. XII. Wyniki analizy jednoczynnikowej ANOVA – przewidywanie wykonania testu fluencji kategoryjnej (zwierzęta) na podstawie wyniku MMSE w badanej grupie

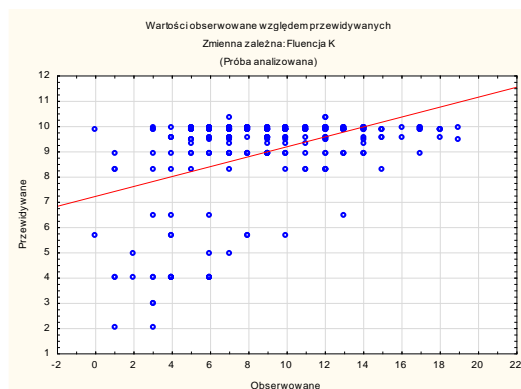
Jednowymiarowe testy istotności dla zwierzęta (Sheet1 w monika dane) Parametryzacja z sigma-ograniczeniami Dekompozycja efektywnych hipotez; Błąd standardowy oceny:3.7337					
Efekt	SS	Stopnie swobody	MS	F	p
Wyraz wolny	2874,688	1	2874,688	206,2087	0,000000
MMSE	1273,342	16	79,584	5,7088	0,000000
Błąd	3680,337	264	13,941		

Tab. XIII. Wyniki analizy jednoczynnikowej ANOVA – przewidywanie wykonania testu fluencji kategoryjnej (przedmioty ostre) na podstawie wyniku MMSE w badanej grupie

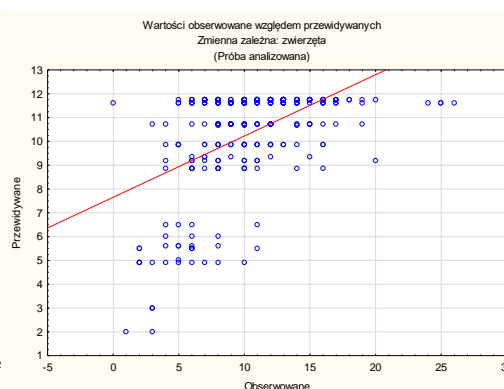
Jednowymiarowe testy istotności dla ostre (Sheet1 w monika dane) Parametryzacja z sigma-ograniczeniami Dekompozycja efektywnych hipotez; Błąd standardowy oceny:2.3603					
Efekt	SS	Stopnie swobody	MS	F	p
Wyraz wolny	1093,497	1	1093,497	196,2785	0,000000
MMSE	478,284	16	29,893	5,3656	0,000000
Błąd	1476,355	265	5,571		

Na rycinie 6 przedstawiono rozkład normalności reszt przewidywanych względem oczekiwanych w teście fluencji słownej – test literowy i kategoryalny (zwierzęta i przedmioty ostre).

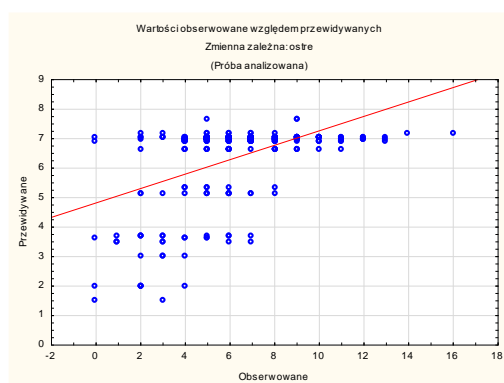
A



B



C



Rycina 6. Rozkład reszt przewidywanych względem oczekiwanych w teście fluencji słownej

A: test fluencji fonologicznej

B: test fluencji kategoryalnej – zwierzęta

C: test fluencji kategoryalnej – przedmioty ostre



W dalszych analizach postanowiono sprawdzić, czy dla poziomu wyników uzyskiwanych przez badanych w teście fluencji słownej ma znaczenie wartość uzyskana przez badanych w teście MMSE. Zastosowano tu analizę współczynników współliniowości z sigma – ograniczeniami, zaś wyniki tych analiz przedstawiono w tabelach XIV, XV i VI.

Jak wynika z danych w tabeli XIV dla sprawności wykonanie testu fluencji fonologicznej najbardziej istotne znaczenie miały wyniki początkowego osłabienia funkcji poznawczych na poziomie 25-23p w teście MMSE ( $p < 0.000$ ), następnie wynik 21 p w MMSE, zaś wyraźne osłabienie wyników odnotowane było of 17 p w MMSE.

Jak wynika z tabeli XV, dla sprawności wykonania testu fluencji kategoryalnej w kategorii „zwierzęta”, sprawność wykonania testu wiązała się z wynikami w granicach 25-23 p w MMSE ( $p < 0,000$ ), zaś krytycznym momentem pogorszenia było osłabienie poznawcze obserwowane w MMSE na poziomie 17 punktów. Dla wyników w teście kategoryalnym „przedmioty ostre” podobnie największe znaczenie miały wyniki 25-23 p w MMSE, a także wynik na poziomie 21 i 18 punktów.

Wyniki te świadczą o tym, że osłabienie fluencji słownej u badanych osób z zaburzeniami poznawczymi nie występuje jednostajnie, a są prawdopodobnie krytyczne momenty przekładające się na pogorszenie wyników fluencji słownej w trakcie postępowania zaburzeń poznawczych. Mechanizm ten wymaga jednak szczegółowego wyjaśnienia, być może na gruncie neurobiologicznym.

Tab. XIV. Zależność pomiędzy wynikami testu MMSE a sprawnością wykonania testu fluencji fonologicznej – test literowy w badanej grupie

Efekt	Statystyki współliniowości								
	Parametryzacja z sigma-ograniczeniami								
	Poziom Efekt	Toleranc.	Wariancja cz.infl.	R <sup>2</sup>	Fluencja K Beta (w)	Fluencja K Cząstk.	Fluencja K Semi-cz	Fluencja K t	Fluencja K p
MMSE	8	0,02	54,95	0,98	-0,43	-0,06	-0,06	-1,05	0,30
MMSE	9	0,02	54,95	0,98	-0,43	-0,06	-0,06	-1,05	0,30
MMSE	10	0,02	54,95	0,98	-0,31	-0,05	-0,04	-0,76	0,45
MMSE	13	0,03	28,78	0,97	-0,56	-0,11	-0,10	-1,87	0,06
MMSE	14	0,05	20,07	0,95	-0,20	-0,05	-0,04	-0,80	0,42
MMSE	15	0,09	11,41	0,91	-0,12	-0,04	-0,04	-0,65	0,51
MMSE	16	0,06	15,73	0,94	-0,02	0,00	0,00	-0,08	0,94
MMSE	17	0,12	8,02	0,88	-0,35	-0,14	-0,12	-2,21	0,03
MMSE	18	0,05	20,07	0,95	0,45	0,11	0,10	1,80	0,07
MMSE	19	0,08	13,13	0,92	-0,33	-0,10	-0,09	-1,63	0,10
MMSE	20	0,09	11,41	0,91	0,34	0,11	0,10	1,80	0,07
MMSE	21	0,13	7,57	0,87	0,37	0,15	0,14	2,44	0,02
MMSE	22	0,15	6,62	0,85	0,22	0,10	0,09	1,57	0,12
MMSE	23	0,20	5,13	0,80	0,45	0,21	0,20	3,56	0,00
MMSE	24	0,21	4,83	0,79	0,37	0,19	0,17	3,05	0,00
MMSE	25	0,21	4,75	0,79	0,60	0,29	0,27	4,95	0,00

Tab. XV. Zależność pomiędzy wynikami testu MMSE a sprawnością wykonania testu fluencji kategorialnej – zwierzęta w badanej grupie

Efekt	Statystyki współliniowości									
	Parametryzacja z sigma-ograniczeniami									
	Poziom Efekt	Toleranc.	Wariancja cz.infl.	R <sup>2</sup>	Zwierzęta Beta (w)	Zwierzęta Cząstk.	Zwierzęta Semi-cz	Zwierzęta t	Zwierzęta p	
MMSE	8	0,02	55,20	0,98	-0,47	-0,07	-0,06	-1,20	0,23	
MMSE	9	0,02	55,20	0,98	-0,47	-0,07	-0,06	-1,20	0,23	
MMSE	10	0,02	55,20	0,98	-0,25	-0,04	-0,03	-0,64	0,52	
MMSE	13	0,03	28,90	0,97	-0,59	-0,13	-0,11	-2,08	0,04	
MMSE	14	0,05	20,15	0,95	-0,14	-0,04	-0,03	-0,61	0,55	
MMSE	15	0,09	11,45	0,91	-0,21	-0,07	-0,06	-1,16	0,25	
MMSE	16	0,06	15,79	0,94	-0,09	-0,03	-0,02	-0,42	0,68	
MMSE	17	0,12	8,04	0,88	-0,29	-0,12	-0,10	-1,94	0,05	
MMSE	18	0,05	20,15	0,95	0,39	0,10	0,09	1,63	0,10	
MMSE	19	0,08	13,18	0,92	-0,19	-0,06	-0,05	-1,01	0,31	
MMSE	20	0,09	11,45	0,91	0,24	0,08	0,07	1,36	0,17	
MMSE	21	0,13	7,59	0,87	0,24	0,10	0,09	1,64	0,10	
MMSE	22	0,15	6,64	0,85	0,20	0,09	0,08	1,50	0,14	
MMSE	23	0,19	5,13	0,81	0,37	0,19	0,16	3,07	0,00	
MMSE	24	0,21	4,80	0,79	0,53	0,27	0,24	4,57	0,00	
MMSE	25	0,21	4,76	0,79	0,75	0,37	0,34	6,45	0,00	

Tab. XVI. Zależność pomiędzy wynikami testu MMSE a sprawnością wykonania testu fluencji kategoryalnej – przedmioty ostre- w badanej grupie

Efekt	Statystyki współliniowości								
	Parametryzacja z sigma-ograniczeniami								
	Poziom Efekt	Toleranc.	Wariancja cz.infl.	R <sup>2</sup>	ostre Beta (w)	ostre Cząstk.	ostre Semi-cz	ostre t	ostre p
MMSE	8	0,02	55,64	0,98	-0,44	-0,07	-0,06	-1,11	0,27
MMSE	9	0,02	55,64	0,98	-0,44	-0,07	-0,06	-1,11	0,27
MMSE	10	0,02	55,64	0,98	-0,44	-0,07	-0,06	-1,11	0,27
MMSE	13	0,03	29,12	0,97	-0,54	-0,11	-0,10	-1,86	0,06
MMSE	14	0,05	20,30	0,95	-0,45	-0,11	-0,10	-1,88	0,06
MMSE	15	0,09	11,53	0,91	-0,18	-0,06	-0,05	-1,02	0,31
MMSE	16	0,06	15,91	0,94	-0,27	-0,08	-0,07	-1,28	0,20
MMSE	17	0,12	8,10	0,88	-0,15	-0,06	-0,05	-1,01	0,31
MMSE	18	0,05	20,30	0,95	0,58	0,15	0,13	2,42	0,02
MMSE	19	0,08	13,28	0,92	-0,16	-0,05	-0,05	-0,84	0,40
MMSE	20	0,09	11,53	0,91	0,16	0,05	0,05	0,89	0,37
MMSE	21	0,13	7,64	0,87	0,54	0,22	0,19	3,65	0,00
MMSE	22	0,15	6,68	0,85	0,14	0,06	0,05	0,99	0,33
MMSE	23	0,19	5,17	0,81	0,49	0,24	0,22	4,03	0,00
MMSE	24	0,21	4,83	0,79	0,60	0,30	0,27	5,12	0,00
MMSE	25	0,21	4,78	0,79	0,67	0,33	0,31	5,77	0,00

Podobną analizę zależności przeprowadzono w przypadku znaczenia czasu trwania zaburzeń poznawczych dla sprawności utrzymania zdolności językowych. Wyniki analiz jednowymiarowych przedstawiono w poniższych tabelach XVII-XIX. Na rycinie 7 przedstawiono graficzny obraz wartości obserwowanych względem oczekiwanych w teście fluencji słownej.

Tab. XVII. Wyniki analizy jednoczynnikowej ANOVA – przewidywanie wykonania testu fluencji fonologicznej – test literowy na podstawie czasu trwania zaburzeń poznawczych.

Jednowymiarowe testy istotności dla Fluencja K (Sheet1 w monika dane) Parametryzacja z sigma-ograniczeniami Dekompozycja efektywnych hipotez; Błąd standardowy oceny:3.7518					
Efekt	SS	Stopnie swobody	MS	F	p
Wyraz wolny	7503,989	1	7503,989	533,1191	0,000000
czas trwania zaburzeń	992,218	32	31,007	2,2029	0,000417
Błąd	3462,606	246	14,076		

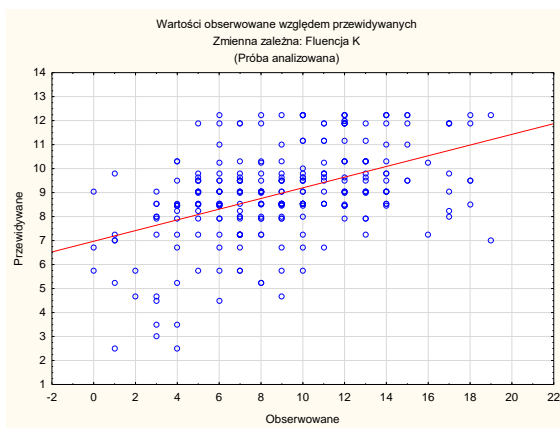
Tab. XVIII. Wyniki analizy jednoczynnikowej ANOVA – przewidywanie wykonania testu fluencji kategoryjnej (zwierzęta) na podstawie czasu trwania zaburzeń poznawczych

Jednowymiarowe testy istotności dla zwierzęta (Sheet1 w monika dane) Parametryzacja z sigma-ograniczeniami Dekompozycja efektywnych hipotez; Błąd standardowy oceny:3.9009					
Efekt	SS	Stopnie swobody	MS	F	p
Wyraz wolny	9420,337	1	9420,337	619,0734	0,000000
czas trwania zaburzeń	1270,526	32	39,704	2,6092	0,000018
Błąd	3788,991	249	15,217		

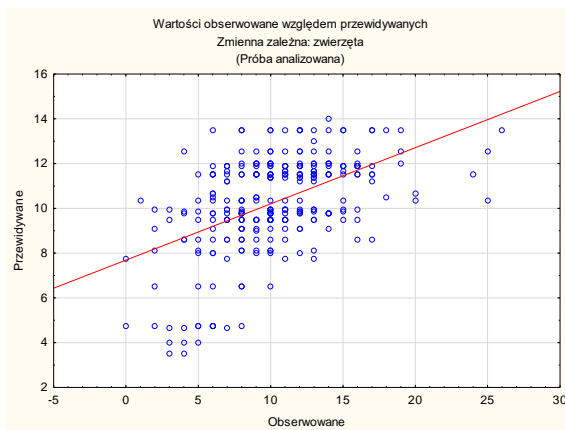
Tab. XIX. Wyniki analizy jednoczynnikowej ANOVA – przewidywanie wykonania testu fluencji kategoryjnej (przedmioty ostre) na podstawie czasu trwania zaburzeń poznawczych

Jednowymiarowe testy istotności dla ostre (Sheet1 w monika dane) Parametryzacja z sigma-ograniczeniami Dekompozycja efektywnych hipotez; Błąd standardowy oceny:2.4739					
Efekt	SS	Stopnie swobody	MS	F	p
Wyraz wolny	3788,171	1	3788,171	618,9611	0,000000
czas trwania zaburzeń	465,185	32	14,537	2,3752	0,000111
Błąd	1530,052	250	6,120		

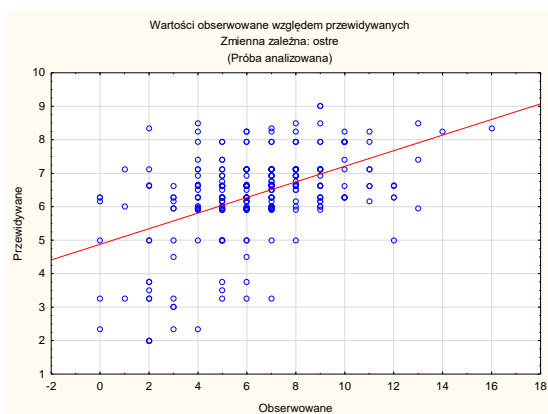
A



B



C



Ryc. 7. Rozkład wartości obserwowanych względem oczekiwanych w teście fluencji słownej w zależności od czasu trwania zaburzeń poznawczych.

Jak wynika z przedstawionych wyżej danych, czas trwania zaburzeń miał znaczenie w sprawności wykonania wszystkich rodzajów testów fluencji słownej. Wykonano więc analizę, jaki czas trwania zaburzeń poznawczych jest najbardziej istotny w obserwowanych zmianach wykonania testu fluencji. Poniżej w tabelach XX-XXII przedstawiono współczynniki współliniowości z sigma-0ograniczeniami dla konkretnych miesięcy czasu trwania zaburzeń poznawczych. Wyniki te wskazują, że dla fluencji fonologicznej najbardziej znamienne były 6, 8 i 9 miesiąc, dla fluencji kategoryalnej w kategorii „zwierzęta” tych punktów czasowych było znacznie więcej i były to miesiące: 6, 8-9, 12, 26, 31-32 i 36. Dla fluencji kategoryalnej w kategorii „przedmioty ostre” były to miesiące: 5-6,9, 21-22, 26, 31-32, 35-36.

Analizując wszystkie znaczące okresy najbardziej istotne dla zmian poznawczych, to dla wszystkich kategorii były to miesiące 6 i 9, a dla kategorii fluencji kategoryalnej miesiące 6,9, a także 26, 31-32 i 36.

Tab. XX. Czas trwania zaburzeń poznawczych a wyniki testu fluencji słownej fonologicznej w badanej grupie

Efekt	Statystyki współliniowości								
	Parametryzacja z sigma-ograniczeniami								
	Poziom Efekt	Toleranc.	Wariancja cz.infl.	R <sup>2</sup>	Fluencja K Beta (w)	Fluencja K Cząstk.	Fluencja K Semi-cz	Fluencja K t	Fluencja K p
Czas trwania zaburzeń	3	0,35	2,84	0,65	0,10	0,07	0,06	1,04	0,30
Czas trwania zaburzeń	4	0,35	2,84	0,65	0,05	0,03	0,03	0,49	0,62
Czas trwania zaburzeń	5	0,80	1,25	0,20	0,07	0,07	0,06	1,07	0,28
Czas trwania zaburzeń	6	0,86	1,17	0,14	0,24	0,25	0,22	3,98	0,00
Czas trwania zaburzeń	7	0,85	1,18	0,15	0,09	0,10	0,09	1,54	0,12
Czas trwania zaburzeń	8	0,85	1,18	0,15	0,12	0,13	0,11	1,98	0,05
Czas trwania zaburzeń	9	0,86	1,17	0,14	0,26	0,26	0,24	4,25	0,00
Czas trwania zaburzeń	10	0,85	1,18	0,15	0,02	0,02	0,02	0,32	0,75
Czas trwania zaburzeń	11	0,76	1,31	0,24	0,13	0,12	0,11	1,95	0,05
Czas trwania zaburzeń	12	0,84	1,18	0,16	0,07	0,07	0,07	1,18	0,24
Czas trwania zaburzeń	13	0,80	1,25	0,20	0,01	0,01	0,01	0,23	0,82
Czas trwania zaburzeń	14	0,80	1,25	0,20	-0,01	-0,01	-0,01	-0,14	0,89
Czas trwania zaburzeń	15	0,85	1,18	0,15	-0,02	-0,02	-0,01	-0,25	0,80
Czas trwania zaburzeń	16	0,86	1,17	0,14	0,10	0,10	0,09	1,64	0,10
Czas trwania zaburzeń	17	0,73	1,37	0,27	0,03	0,03	0,03	0,48	0,63
Czas trwania zaburzeń	18	0,86	1,17	0,14	0,02	0,02	0,02	0,37	0,71
Czas trwania zaburzeń	19	0,83	1,20	0,17	0,01	0,01	0,01	0,23	0,82
Czas trwania zaburzeń	20	0,69	1,46	0,31	0,08	0,07	0,06	1,11	0,27
Czas trwania zaburzeń	21	0,85	1,17	0,15	-0,06	-0,06	-0,05	-0,97	0,33
Czas trwania zaburzeń	22	0,69	1,46	0,31	0,10	0,10	0,09	1,52	0,13
Czas trwania zaburzeń	23	0,69	1,46	0,31	0,00	0,00	0,00	0,03	0,98
Czas trwania zaburzeń	24	0,78	1,28	0,22	-0,07	-0,07	-0,06	-1,04	0,30
Czas trwania zaburzeń	26	0,69	1,46	0,31	-0,11	-0,10	-0,09	-1,59	0,11
Czas trwania zaburzeń	27	0,52	1,91	0,48	0,01	0,01	0,01	0,12	0,91
Czas trwania zaburzeń	28	0,69	1,46	0,31	-0,09	-0,08	-0,07	-1,32	0,19
Czas trwania zaburzeń	29	0,35	2,84	0,65	0,10	0,07	0,06	1,04	0,30
Czas trwania zaburzeń	30	0,35	2,84	0,65	-0,13	-0,09	-0,08	-1,42	0,16
Czas trwania zaburzeń	31	0,62	1,61	0,38	-0,12	-0,11	-0,09	-1,66	0,10
Czas trwania zaburzeń	32	0,69	1,46	0,31	-0,09	-0,08	-0,07	-1,32	0,19
Czas trwania zaburzeń	34	0,62	1,61	0,38	-0,04	-0,04	-0,03	-0,56	0,58
Czas trwania zaburzeń	35	0,52	1,91	0,48	-0,14	-0,11	-0,10	-1,81	0,07
Czas trwania zaburzeń	36	0,52	1,91	0,48	-0,11	-0,09	-0,08	-1,42	0,16



Tab. XXI. Czas trwania zaburzeń poznawczych a wykonanie testu fluencji kategoryjnej - zwierzęta

Efekt	Statystyki współliniowości									
	Parametryzacja z sigma-ograniczeniami									
	Poziom	Kolumna	Toleranc.	Wariancja	R <sup>2</sup>	zwierzęta	zwierzęta	zwierzęta	zwierzęta	zwierzęta
Czas trwania zaburzeń	3	2	0,35	2,84	0,65	0,09	0,06	0,06	1,01	0,31
Czas trwania zaburzeń	4	3	0,35	2,84	0,65	0,00	0,00	0,00	-0,04	0,97
Czas trwania zaburzeń	5	4	0,80	1,25	0,20	0,10	0,10	0,09	1,59	0,11
Czas trwania zaburzeń	6	5	0,86	1,17	0,14	0,17	0,18	0,15	2,82	0,01
Czas trwania zaburzeń	7	6	0,85	1,18	0,15	0,07	0,07	0,06	1,11	0,27
Czas trwania zaburzeń	8	7	0,85	1,17	0,15	0,16	0,17	0,15	2,72	0,01
Czas trwania zaburzeń	9	8	0,86	1,17	0,14	0,27	0,27	0,25	4,50	0,00
Czas trwania zaburzeń	10	9	0,85	1,18	0,15	-0,03	-0,03	-0,03	-0,49	0,62
Czas trwania zaburzeń	11	10	0,76	1,31	0,24	0,10	0,10	0,09	1,58	0,12
Czas trwania zaburzeń	12	11	0,85	1,18	0,15	0,19	0,20	0,17	3,17	0,00
Czas trwania zaburzeń	13	12	0,80	1,25	0,20	0,12	0,12	0,11	1,95	0,05
Czas trwania zaburzeń	14	13	0,80	1,25	0,20	0,03	0,03	0,03	0,51	0,61
Czas trwania zaburzeń	15	14	0,85	0,01	0,15	0,18	0,19	0,17	0,03	0,00
Czas trwania zaburzeń	16	15	0,86	1,17	0,14	0,02	0,03	0,02	0,40	0,69
Czas trwania zaburzeń	17	16	0,73	1,37	0,27	0,08	0,07	0,06	1,18	0,24
Czas trwania zaburzeń	18	17	0,01	1,17	0,14	0,04	0,04	0,04	0,66	0,51
Czas trwania zaburzeń	19	18	0,83	1,20	0,17	0,00	0,00	0,00	-0,06	0,95
Czas trwania zaburzeń	20	19	0,69	1,46	0,31	-0,05	-0,05	-0,04	-0,73	0,46
Czas trwania zaburzeń	21	20	0,85	1,17	0,15	0,05	0,05	0,04	0,76	0,45
Czas trwania zaburzeń	22	21	0,69	1,46	0,31	0,05	0,04	0,04	0,01	0,49
Czas trwania zaburzeń	23	22	0,69	1,46	0,31	-0,04	-0,04	-0,03	-0,60	0,55
Czas trwania zaburzeń	24	23	0,80	1,25	0,20	-0,05	0,00	-0,04	-0,75	0,45
Czas trwania zaburzeń	26	24	0,69	0,01	0,31	-0,15	-0,14	0,00	-2,29	0,02
Czas trwania zaburzeń	27	25	0,01	1,91	0,48	-0,03	-0,03	-0,02	-0,43	0,67
Czas trwania zaburzeń	28	26	0,69	0,01	0,31	-0,09	-0,09	-0,08	-1,38	0,17
Czas trwania zaburzeń	29	27	0,35	2,84	0,65	0,12	0,08	0,07	0,01	0,20
Czas trwania zaburzeń	30	28	0,35	2,84	0,65	-0,13	-0,09	-0,07	-1,36	0,18
Czas trwania zaburzeń	31	29	0,62	1,61	0,38	-0,14	-0,13	-0,11	-2,03	0,04
Czas trwania zaburzeń	32	30	0,69	1,46	0,31	-0,15	-0,14	0,00	-2,29	0,02
Czas trwania zaburzeń	34	31	0,62	1,61	0,38	0,00	0,04	0,04	0,68	0,50
Czas trwania zaburzeń	35	32	0,01	1,91	0,48	-0,15	-0,12	-0,10	-1,91	0,06
Czas trwania zaburzeń	36	33	0,01	1,91	0,48	-0,16	-0,13	-0,12	-2,10	0,04

Tab. XXII. Czas trwania zaburzeń a wyniki testu fluencji kategoryjnej - przedmioty ostre

Efekt	Statystyki współliniowości									
	Parametryzacja z sigma-ograniczeniami									
	Poziom Efekt	Kolumna	Toleranc.	Wariancja cz.infl.	R <sup>2</sup>	ostre Beta (w)	ostre Cząstk.	ostre Semi-cz	ostre t	ostre p
Czas trwania zaburzeń	3	2	0,35	2,84	0,65	-0,11	-0,07	-0,06	-1,17	0,24
Czas trwania zaburzeń	4	3	0,35	2,84	0,65	0,12	0,08	0,07	1,32	0,19
Czas trwania zaburzeń	5	4	0,80	1,25	0,20	0,17	0,17	0,15	2,77	0,01
Czas trwania zaburzeń	6	5	0,86	1,17	0,14	0,21	0,22	0,19	3,49	0,00
Czas trwania zaburzeń	7	6	0,85	1,18	0,15	-0,07	-0,08	-0,07	-1,19	0,24
Czas trwania zaburzeń	8	7	0,85	1,17	0,15	0,10	0,11	0,09	1,70	0,09
Czas trwania zaburzeń	9	8	0,86	1,17	0,14	0,13	0,13	0,12	2,12	0,03
Czas trwania zaburzeń	10	9	0,85	1,18	0,15	0,01	0,01	0,01	0,16	0,87
Czas trwania zaburzeń	11	10	0,76	1,31	0,24	0,05	0,05	0,05	0,85	0,40
Czas trwania zaburzeń	12	11	0,84	1,19	0,16	0,06	0,06	0,06	1,01	0,31
Czas trwania zaburzeń	13	12	0,80	1,25	0,20	0,09	0,09	0,08	1,49	0,14
Czas trwania zaburzeń	14	13	0,80	1,25	0,20	0,05	0,05	0,04	0,78	0,43
Czas trwania zaburzeń	15	14	0,85	1,18	0,15	0,03	0,03	0,03	0,49	0,63
Czas trwania zaburzeń	16	15	0,85	1,17	0,15	0,09	0,10	0,08	1,53	0,13
Czas trwania zaburzeń	17	16	0,73	1,37	0,27	0,09	0,09	0,08	1,45	0,15
Czas trwania zaburzeń	18	17	0,86	1,17	0,14	0,01	0,01	0,01	0,23	0,82
Czas trwania zaburzeń	19	18	0,83	1,20	0,17	0,07	0,07	0,06	1,09	0,28
Czas trwania zaburzeń	20	19	0,69	1,46	0,31	0,05	0,05	0,04	0,77	0,44
Czas trwania zaburzeń	21	20	0,85	1,17	0,15	0,12	0,13	0,11	2,04	0,04
Czas trwania zaburzeń	22	21	0,69	1,46	0,31	0,15	0,14	0,12	2,20	0,03
Czas trwania zaburzeń	23	22	0,69	1,46	0,31	0,01	0,01	0,01	0,15	0,88
Czas trwania zaburzeń	24	23	0,80	1,25	0,20	0,01	0,01	0,01	0,21	0,83
Czas trwania zaburzeń	26	24	0,69	1,46	0,31	-0,14	-0,13	-0,12	-2,10	0,04
Czas trwania zaburzeń	27	25	0,52	1,91	0,48	-0,06	-0,05	-0,04	-0,77	0,44
Czas trwania zaburzeń	28	26	0,69	1,46	0,31	-0,11	-0,11	-0,09	-1,69	0,09
Czas trwania zaburzeń	29	27	0,35	2,84	0,65	0,12	0,08	0,07	1,32	0,19
Czas trwania zaburzeń	30	28	0,35	2,84	0,65	-0,11	-0,07	-0,06	-1,17	0,24
Czas trwania zaburzeń	31	29	0,62	1,61	0,38	-0,17	-0,15	-0,14	-2,48	0,01
Czas trwania zaburzeń	32	30	0,69	1,46	0,31	-0,14	-0,13	-0,12	-2,10	0,04
Czas trwania zaburzeń	34	31	0,62	1,61	0,38	0,13	0,11	0,10	1,80	0,07
Czas trwania zaburzeń	35	32	0,52	1,91	0,48	-0,17	-0,14	-0,12	-2,23	0,03
Czas trwania zaburzeń	36	33	0,52	1,91	0,48	-0,17	-0,14	-0,12	-2,23	0,03

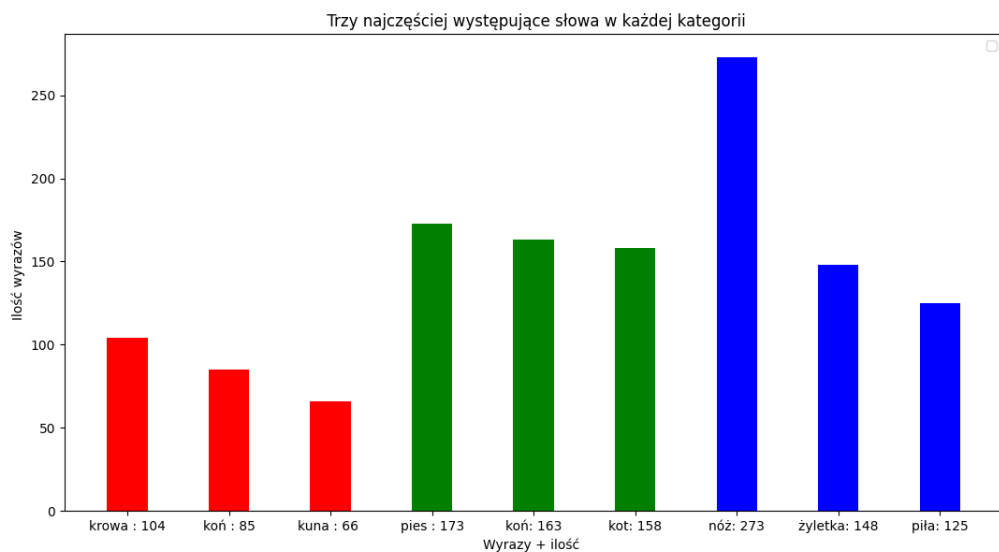
## 7.5. Analiza korpusowa

Analiza korpusowa wykazała znaczące różnice w nazwach własnych, które wystąpiły tylko w teście literowym. Zaobserwowano zubożenie użycia neologizmów przez pacjentów. Neologizmy występowały tylko w teście literowym, co było statystycznie istotne. U kobiet wystąpiło więcej nazw własnych niż u mężczyzn. W pozostałych parametrach ich wyniki nie różniły się istotnie.

Dalszej analizie empirycznej części niniejszej dysertacji zostały poddane wyekscerpowane słowa. Lista najczęściej występujących wiązek leksykalnych, czy też klasterów została skategoryzowana za pomocą pomiaru siły frekwencji wiązek (MI). Strukturalne i funkcjonalne cechy najczęściej występujących wiązek leksykalnych zostały uporządkowane w oparciu o drobiazgową analizę zgodności, która umożliwiła alokację wiązek za pomocą zmodyfikowanych ram strukturalnych zaproponowanych przez Bibera i wsp. (1999) wraz z uwzględnieniem taksonomi funkcjonalnej Hylanda (2008a).

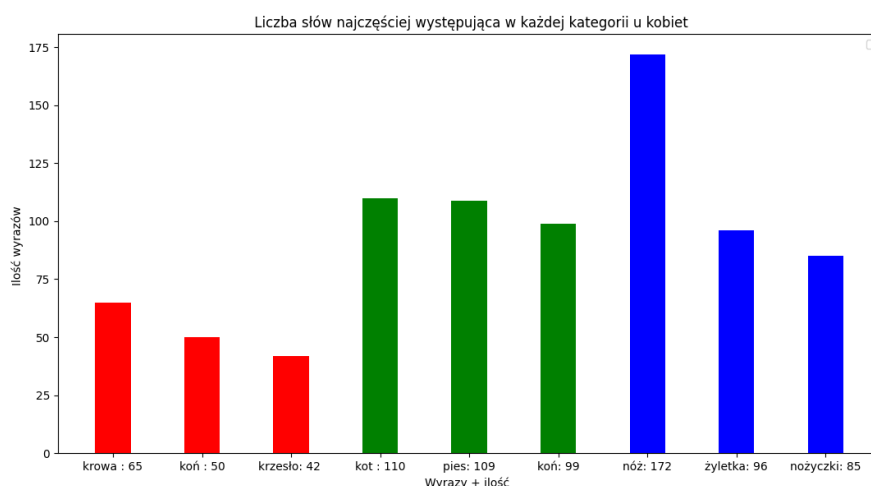
Niniejszy korpus językowy zawiera łączną liczbę słów 7669, w tym 7536 słów poprawnych i 133 słów niepoprawnych, wliczając perseweracje, neologizmy oraz słowa nienależące do danej kategorii i nazwy własne. Czasami słowami nieprawidłowymi były również słowa ogólnie przyjęte za wulgarne (np. kutas itp.). Nie jest to słowo literackie i występuje najczęściej w ujęciu pejoratywnym, lecz ze względu na niską frekwencję występowań, wulgaryzmy stanowią w badaniu wartość statystycznie nieistotną.

Poniższa rycina przedstawia trzy najczęściej występujące słowa w każdej z trzech kategorii (słowa na literę „k”, zwierzęta oraz przedmioty ostre) w teście fluencji słownej dotyczącej badanej grupy pacjentów z zaburzeniami otępiennymi.

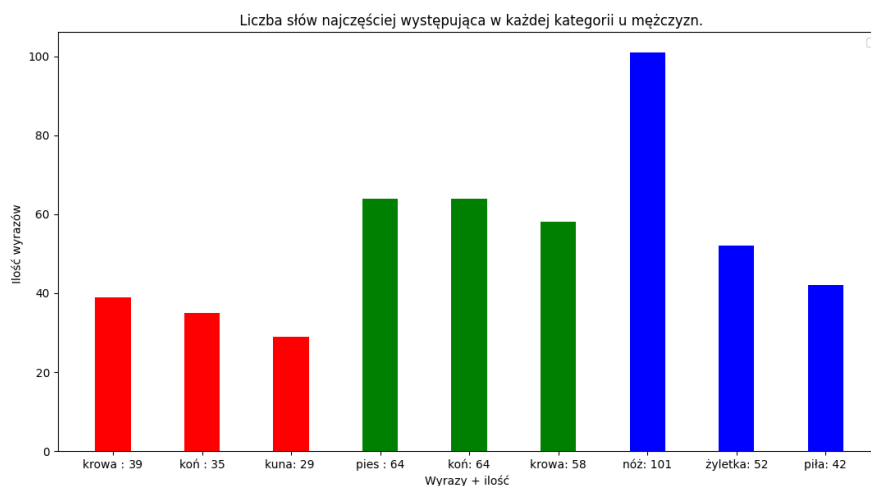


Ryc. 8. Trzy najczęściej występujące słowa w wykonaniu testu fluencji przez badanych w każdej kategorii

Co ciekawe, słowo *koń* występuje jako drugie najczęściej występujące słowo w pierwszych dwóch kategoriach. W kategorii słowa na literę „k” wystąpiło ono 85 razy, natomiast w kategorii narzędzia słowo *koń* wystąpiło 163 razy. Co więcej, jak widać na wykresach z podziałem na płeć, kobiety i mężczyźni wykazują wyraźne różnice. U mężczyzn słowo *krowa* wystąpiło jako najczęstsze słowo w dwóch kategoriach. W kategorii słowa na literę „k” wystąpiło 39 razy i plasuje się na miejscu pierwszym oraz w drugiej kategorii zwierzęta, gdzie z frekwencją występowania 58 zajmuje ono trzecie miejsce. Co ciekawe, u kobiet słowo *krowa* występuje 65 razy i tylko w kategorii słowa na literę „k”. Natomiast słowo *koń* występuje w dwóch kategoriach w przypadku obu płci. U kobiet w pierwszej kategorii słowa na literę „k” z frekwencją 50 zajmuje drugie miejsce, a w kategorii zwierzęta z frekwencją 99 zajmuje trzecie miejsce. U mężczyzn słowo *koń* wystąpiło 35 razy w pierwszej kategorii, słowa na literę „k” zajmując tym samym drugie miejsce oraz drugie miejsce z frekwencją 64 w kategorii zwierzęta.



Ryc. 9. Trzy najczęściej występujące słowa w wykonaniu testu fluencji w każdej kategorii u kobiet



Ryc. 10. . Trzy najczęściej występujące słowa w wykonaniu testu fluencji w każdej kategorii u mężczyzn

Porównując listy frekwencyjne wyekscerpowane z Narodowego Korpusu Języka Polskiego do trzech najczęściej występujących słów w każdej z trzech kategorii w teście fluencji w niniejszym badaniu można wyróżnić trzy najczęściej występujące słowa w każdej z tych kategorii (ryc. 8.) oraz najczęściej występujące słowa w każdej z tych kategorii z podziałem na płeć (ryc. 9., ryc. 10.). Przy pomocy wyszukiwarki PELCRA Narodowego Korpusu Języka Polskiego można dokonać przeszukania ponad 250

milionów słów tekstowych, czyli w przybliżeniu 300 milionów segmentów oraz z całej zawartości danych Narodowego Korpusu Języka Polskiego liczącego około 1500 milionów słów tekstowych, co daje 1800 milionów segmentów.

Słowo „krowa” wystąpiło 930 razy w zbalansowanym podkorpusie NKJP (liczącym około 300 milionów segmentów) oraz 555 razy w korpusie IPI PAN (liczącym około 250 milionów segmentów). W niniejszym badaniu słowo „krowa” wystąpiło 104 razy, co stanowi 3,8% wszystkich słów w danej kategorii (słowa na literę „k”). Następnie słowo najczęściej występujące w tej samej kategorii „koń” wystąpiło 1000 razy w zbalansowanym podkorpusie NKJP oraz 1000 razy w korpusie IPI PAN, natomiast w niniejszym badaniu wystąpiło 85 razy, co stanowi około 3% wszystkich słów w tej kategorii. Następująco słowo „kuna” wystąpiło 94 razy w zbalansowanym podkorpusie NKJP oraz 57 razy w korpusie IPI PAN, w danym badaniu wystąpiło 66 razy i stanowi 2,44% wszystkich słów w tej kategorii.

W kolejnej kategorii (zwierzęta) najczęściej występującym wariantem jest słowo „pies”, które wystąpiło 109 razy w zbalansowanym podkorpusie NKJP oraz 1000 razy w korpusie IPI PAN, natomiast w niniejszym badaniu wystąpiło 173 razy, co stanowi już tylko 5,3% wszystkich słów w danej kategorii. Na drugim miejscu pod względem frekwencji występowania znalazło się słowo „koń”, które wystąpiło 1000 razy w zbalansowanym podkorpusie NKJP oraz 1000 razy w korpusie IPI PAN, a w niniejszym badaniu pojawiło się 163 razy, co stanowi 5,2% wszystkich słów w tej kategorii. Należy nadmienić, że to słowo wystąpiło również jako jedno z najczęściej wymienianych słów przez pacjentów w pierwszej kategorii, słów na literę „k”. Trzecie najczęściej wymieniane słowo „kot”, które wystąpiło 881 razy w zbalansowanym podkorpusie NKJP oraz 717 razy w korpusie IPI PAN, w niniejszym badaniu wystąpiło 158 razy, co stanowi 5,1% wszystkich słów w tej kategorii.

Ostatnia kategoria (narzędzia ostre) zawiera słowo o najwyższej frekwencji „nóż”, które wystąpiło 755 razy w zbalansowanym podkorpusie NKJP oraz 1000 razy w korpusie IPI PAN. W niniejszym badaniu wystąpiło 273 razy, co stanowi 14,5% wszystkich słów w tej kategorii. Słowo „żyłetka” natomiast wystąpiło 39 razy w zbalansowanym podkorpusie NKJP oraz 14 razy w korpusie IPI PAN, plasując się na niższej pozycji jeśli

chodzi o frekwencję występowania w oparciu o dane w Narodowym Korpusie Języka Polskiego niż wcześniejsze słowa. W niniejszym badaniu zostało ono wymienione aż 148 razy, co stanowi 7,8% wszystkich słów w tej kategorii, czyli charakteryzuje się dość wysoką frekwencją. A słowo „piła”, które wystąpiło 358 razy w zbalansowanym podkorpusie NKJP oraz 252 razy w korpusie IPI PAN, wskazując raczej na wysoką frekwencję występowania w Narodowym Korpusie Języka Polskiego, w niniejszym badaniu zostało wymienione 125 razy, co stanowi tylko 0,66% wszystkich słów w danej kategorii.

Poniższa tabelka opiera się na danych zebranych przez Bodę i wsp. (2011) pokazujących procentowy udział poszczególnych liter w najczęściej wymienianych słowach w Narodowym Korpusie Języka Polskiego w wydaniu podkorpusu IPI PAN liczącego ponad 30 milionów segmentów oraz w przybliżeniu 25 milionów słów ortograficznych. W korpusie tym zaobserwowano występowanie również rzadszych liter pochodzących z innych alfabetów, jednak ich frekwencja nie była wysoka, więc stanowiły wartość statystycznie nieistotną.

Tab. XXIII. Procentowy udział poszczególnych liter w najczęściej wymienianych słowach w NKJP w podkorpusie IPI PAN [51]

a	8.91%	w	4.65%	p	3.13%	g	1.42%	ć	0.40%
i	8.21%	s	4.32%	m	2.80%	ę	1.11%	f	0.30%
o	7.75%	t	3.98%	u	2.50%	h	1.08%	ń	0.20%
e	7.66%	c	3.96%	j	2.28%	ą	0.99%	q	0.14%
z	5.64%	y	3.76%	l	2.10%	ó	0.85%	ź	0.06%
n	5.52%	k	3.51%	ł	1.82%	ż	0.83%	v	0.04%
r	4.69%	d	3.25%	b	1.47%	ś	0.66%	x	0.02%

O doborze liter decyduje niska bądź wysoka frekwencja występowania wyrazów w konkretnym języku. W rozdziale 2. zostało wyróżnionych kilka wariantów testu, w skład których wchodzi różne zestawy liter, np. F, A, S w wersji oryginalnej Thurstone'a, co stanowi odpowiednio F- 0,30%, A- 8, 91%, natomiast frekwencja występowania słów na literę S była na tyle niska iż jej wartość jest statystycznie nieistotna, więc nie została ujęta w tabeli. Warto zauważyć, iż w odróżnieniu od litery S, frekwencja występowania słów na literę Ś stanowi aż 0,66% podkorpusu IPI PAN.

W Controlled Word Fluency Test najczęściej występujące słowa są na litery P (3,13%), R (4,69%), W (wartość statystycznie nieistotna) oraz C (wartość statystycznie nieistotna), F (0,30%), L (2,10%). W Test of Verbal Conceptualization and Fluency są to odpowiednio: P (3,13%), D (3,25%), S (wartość statystycznie nieistotna), oraz T (3,98%). Z kolei, w Delis–Kaplan Executive Function System najczęściej występujące słowa są na litery F (0,30%), A (8,91%), S (wartość nieistotna statystycznie) oraz B (1,47%), H (1,08%), R (4,69%).

W badaniach polskich na fluencję fonemiczną/ literową często wykorzystywane litery według Jodzio (2006) to K (3,51%), M (2,80%), P (3,13), S (wartość statystycznie nieistotna), a według Szepietowskiej oraz Gawdy (2011) są to K i F, gdzie słowa na literę K według Broda i wsp. (2011) stanowią 3,51% podkorpusu IPI PAN, natomiast słowa na literę F stanowią tylko 0,30% tego właśnie korpusu.

Frekwencja występowania słów na konkretną literę w Narodowym Korpusie Języka Polskiego oraz w niniejszej grupie badawczej jest bardzo zróżnicowana. Każdy test na fluencję słowną zawiera litery, które występują stosunkowo często w uzusie języka polskiego oraz te które występują statystycznie rzadko, co również można zauważyć w porównaniu wyników badań na najczęściej występujące słowa w każdej z trzech kategorii.



## ROZDZIAŁ VIII

### DYSKUSJA

Niniejsza praca jest jedną z pierwszych prac naukowych stosujących w podejściu do analizy neurolingwistycznej analizę korpusową. Wybrano w niej analizę testu fluencji słownej stosowanego rutynowo w badaniu osób w starszym wieku. Wyniki odnoszono również do nasilenia zaburzeń poznawczych mierzonych testem przesiewowym Mini Mental State Examination.

#### 8.1. Wyniki testów poznawczych

W pierwszej fazie analiz, zgodnie z przyjętym postępowaniem w badaniach naukowych, dokonano oceny przydatności stosowanych narzędzi badawczych w grupie osób starszych. Z uwagi na charakter analizowanych zmiennych i spełnione warunki dla określonego rodzaju analizy, w analizie testu Fluencji Słownej oraz MMSE zastosowano wskaźnik Alfa-Cronbacha. Przyjęto, że wskaźnik  $\alpha > 0,70$  świadczy o wysokiej zgodności wewnętrznych analizowanych skal. Na podstawie tych analiz wykazano, że podskale Testu Fluencji Słownej oraz MMSE w grupie badanych osób starszych uzyskały wysoką zgodność wewnętrzną, na co wskazują bardzo wysokie wskaźniki alpha Cronbacha mieszczące się w granicach 0,73 (MMSE) do 0,82 (0,80- test Fluencji literowy, 0,82 Test Fluencji – „zwierzęta” i „przedmioty ostre”). W analizie czynnikowej stwierdzono natomiast, że wszystkie podtesty Testu Fluencji oraz wynik MMSE wyjaśnia jeden czynnik, co wskazuje na podobny mechanizm etiopatogenetyczny wyjaśniający zaburzenia poznawcze w badanych obszarach. Podobne podejście do analizy testów

neurokognitywnych, m.in. testów fluencji, stosowane jest w badaniach współczesnych u osób starszych z zaburzeniami poznawczymi [157].

Analizowana grupa osób starszych była jednorodna. Pod względem wieku i wyników testu MMSE grupa kobiet i mężczyzn się nie różniła. W teście Fluencji Słownej, kobiety uzyskały wyższy wynik, niż mężczyźni w teście literowym, natomiast w testach kategoryalnych (zwierzęta i przedmioty ostre) wyniki były podobne. Wynik ten należy jednak rozpatrywać wyłącznie w odniesieniu do badanej grupy, gdyż nie wszystkie badania światowe są w tym względzie jednoznaczne. U osób starszych w populacji hiszpańskojęzycznej nie stwierdzono istotnych różnic w wykonaniu testów fluencji słownej w zależności od płci, poza podtestem „zwierzęta” [170], a w innym badaniu tej grupy językowej takich istotnych różnic nie stwierdzono [237]. W badaniu autorów polskich stwierdzono natomiast istotne różnice w wykonaniu testów fluencji u osób młodych. Mężczyźni uzyskali w tym badaniu nieco lepsze wyniki w teście fluencji semantycznej, ale nie zauważono różnic w zakresie strategii werbalnej. Autorzy wysunęli hipotezę, że wyniki te mogą wskazywać na różnice w systemie leksykalnym w powiązaniu z rolami społecznymi. Jest to jednakże hipoteza mało przekonująca, zwłaszcza, że badanie nie było wykonane na grupie reprezentatywnej uprawniającej do uogólnień [256].

Jedna z ostatnich metaanaliz wykazała, że najbardziej istotną zmienną wpływającą na poziom wykonania testów fluencji u osób starszych jest wiek, natomiast płeć nie ma tu istotnego znaczenia [258]. Wyniki analizy korelacyjnej przeprowadzonej w niniejszej pracy nie do końca korespondują z tą obserwacją. Nie stwierdzono bowiem istotnych korelacji pomiędzy wiekiem badanych i wykonaniem testu fluencji słownej oraz testu MMSE, ani też pomiędzy wiekiem badanych osób a frekwencją nieprawidłowo włączonych słów w teście fluencji. Jedynie czas trwania dysfunkcji neurokognitywnych ujemnie korelował z liczbą generowanych słów w teście fluencji słownej. Można więc stwierdzić, że ogólny wynik testu fluencji pogarsza się wraz z czasem trwania zaburzeń poznawczych. Wynik ten jest zbieżny z wynikami badań amerykańskich, które wyraźnie wskazują, że pogarszanie się fluencji słownej, zwłaszcza w teście fluencji fonematycznej i semantycznej jest predyktorem ogólnego pogorszenia poznawczego w czasie [182]. Podobnie w niniejszej pracy stwierdzono wyraźny związek pomiędzy sprawnością funkcji

poznawczych mierzonych testem MMSE a wykonaniem testów fluencji słownej. Im większe obniżenie funkcji poznawczych obserwowano w MMSE, tym gorsze wyniki we wszystkich testach fluencji uzyskiwały badane osoby.

W niniejszej pracy analizowano również frekwencję nieprawidłowo włączanych słów w stosowanych próbach fluencji słownej. Analizowano tu częstość występowania nazw własnych, neologizmów i perseweracji w całej badanej grupie oraz oceniano różnice w tym względzie pomiędzy wykonaniem testów fluencji przez kobiety i mężczyzn. Różnicę zaobserwowano w odniesieniu do jednego z parametrów, mianowicie kobiety częściej używały nazw własnych. Warto zaznaczyć, że słowa niezgodne z kryterium są często istotnym wskaźnikiem patologii mózgowej, w szczególności zaś perseweracje mogą wskazywać na patologię okolic czołowych mózgu i są ważnym elementem w ocenie neuropsychologicznej. W jednej z prac wykazano, że częstość perseweracji w podteście fluencji w kategorii „zwierzęta” są wskaźnikiem wczesnego stadium zaburzeń poznawczych u osób starszych [202]. Inni badacze donoszą natomiast, że liczba intruzji może być wskaźnikiem patologii OUN w stwardnieniu zanikowym bocznym [212]. Dane te wskazują na potrzebę oceny charakteru słów nieprawidłowo włączanych w różnych jednostkach klinicznych, a także znaczenia klinicznego tego zjawiska.

Współczesne wyniki badań wskazują, że wyższa sprawność fluencji słownej jest związana z redukcją ryzyka demencji o 60% oraz o około 25% konwersji z etapu łagodnych zaburzeń poznawczych (MCI) do otępienia. Wykazano, że jest to niezależne od wieku, płci, rasy, czy grupy etnicznej oraz polimorfizmu APOE-4. Wskazuje to na ogromne znaczenie badań fluencji słownej w grupach osób zagrożonych ryzykiem otępienia [266].

Interesujące dane, stanowiące potwierdzenie powyżej wymienionych badań światowych, uzyskano dzięki analizie predykcji. Dokonano analizy predykcji wykonania testów fluencji słownej na podstawie uzyskanego przez badanych wyniku w skali MMSE. Uzyskano dane świadczące o bardzo dużej sile predykcyjnej wykonania MMSE na wynik wszystkich testów fluencji słownej (dane te przedstawiają tabele XI-XIII i rycina 6). Najciekawszą jednak analizą w tym zakresie było określenie, czy konkretny wynik w MMSE może być wiązany z nagłym pogorszeniem wyników w testach fluencji słownej,

a być może nawet świadczyć o nasileniu patologii UON warunkującej obniżenie fluencji słownej. Jest to oczywiście próba, której wyniki nie mogą być nadinterpretowane, jednakże mogą stanowić ciekawy obszar poszukiwań w interdyscyplinarnych badaniach neurolingwistycznych. Wyniki analiz (tabele XIV-XVI) wskazały, że pogorszenie wykonania testu fluencji fonologicznej (literowego z literą „K”) najbardziej istotne znaczenie miały wyniki początkowego osłabienia funkcji poznawczych na poziomie 25-23p, a w kolejności wynik 21p w MMSE, zaś wyraźne osłabienie wyników w tej próbie wiązało się z obniżeniem wyniku MMSE od 17 p. Podobnie w odniesieniu do testu fluencji kategoryjnej w kategorii „zwierzęta” lepsza sprawność wykonania testu wiązała się z wynikami w granicach 25-23 p w MMSE, a momentem nagłego pogorszenia wyników fluencji było również osłabienie poznawcze na poziomie 17 punktów w MMSE. W wykonaniu testu kategoryjnego „przedmioty ostre” największe znaczenie miały wyniki 25-23 p w MMSE, a także, nieco odmiennie niż w poprzednich dwóch próbach, wynik na poziomie 21 i 18 punktów. Wyniki te wskazują na potrzebę analiz na dużych grupach osób starszych z zaburzeniami poznawczymi, aby można było dokonać szerszej interpretacji znaczenia klinicznego tego zjawiska. Obniżanie się sprawności fluencji słownej prawdopodobnie występuje skokowo, a wyniki uzyskane w tej analizie mogą świadczyć o istnieniu krytycznych etapów w trakcie postępujących zmian poznawczych. Interesujące byłoby poznanie mechanizmów neurobiologicznych leżących u podstaw tego zjawiska, ale przekracza to możliwości niniejszej pracy.

Prace w tym nurcie są podejmowane na świecie, jedna z nich wskazuje, że istnieje relatywny związek pomiędzy zmianami naczyniowymi mózgu a osłabieniem poznawczym, w tym nasilaniem się zaburzeń fluencji słownej. Być może mechanizm naczyniowy lub nasilenie zmian neurodegeneracyjnych byłby możliwy do opisania na różnych etapach zaburzeń poznawczych, w tym zaburzeń OUN skutkujących osłabieniem fluencji słownej [2].

Potwierdzeniem tych przypuszczeń są wyniki analizy wskazującej na najbardziej istotny czas osłabienia fluencji słownej w trakcie trwania zaburzeń poznawczych (wyniki zaprezentowano w tabelach XX-XXII). Analiza dla znaczących okresów pogorszenia fluencji wyrażona w miesiącach wykazała, że obniżenie wyników w teście fluencji fonologicznej nastąpiło w miesiącach krytycznych, czyli 6, 8 i 9, w teście fluencji

kategorialnej „zwierzęta” zidentyfikowano więcej punktów czasowych i były to miesiące 6, 8-9, 12, 26, 31-32 i 36, a dla fluencji słownej w kategorii „przedmioty ostre” były to miesiące: 5-6,9, 21-22, 26, 31-32, 35-36. Podsumowanie tej analizy dla wszystkich testów wskazuje, że najbardziej istotne dla wszystkich kategorii fluencji były to miesiące 6 i 9, także 26, 31-32 i 36. Jest to bardzo interesujące, jakie zmiany w OUN nasilają się w tych okresach, tak znaczące dla nasilenia się dysfunkcji w obszarze funkcji werbalnych.

Nieco światła na ten problem rzucają wyniki najnowszych badań neuroobrazowych, które wskazują na istotne znaczenie zmian w zakresie struktur mózgowych i naczyń w predykcji nasilenia się zmian poznawczych. Z drugiej strony, po nasileniu się zmian poznawczych można wnioskować o występującym nasileniu zmian mózgowych. Jest to podejście bardzo istotne w poszukiwaniu markerów kognitywnych otępienia. Jokinen i wsp. (2020) opisali zmiany w obrębie małych naczyń mózgowych, które mogą być uznane za markery neuroobrazowe osłabienia poznawczego [133]. Inni autorzy opisali zmiany biochemiczne (metaboliczne) jako markery osłabienia poznawczego w chorobie Alzheimera [126].

Jednocześnie rozwija się nurt badań wzmacniający podejście traktujące dysfunkcje poznawcze jako biomarkery uszkodzenia mózgu w chorobach otępiennych [211]. Wyniki uzyskane w tej pracy wpisują się więc w nurt badań światowych poruszający znaczenie zaburzeń poznawczych w relacji do różnych zmian mózgowych w chorobach przebiegających z otępieniem. Prawdopodobnie identyfikacja punktów czasowych osłabienia fluencji słownej dokonana w niniejszej dysertacji może wskazywać na zaburzenia funkcjonowania połączeń neuronalnych odpowiedzialnych za sprawność fluencji słownej (tzw. verbal fluency brain network). Badania opublikowane w 2021 roku wskazują, że system połączeń neuronalnych związanych z fluencją semantyczną obejmuje struktury mózgu związane z procesami językowymi, które są zlokalizowane w okolicach czołowych (m.in. pola Brodmana 6, 9, 45 i 47) i przedniej części kory zakrętu obręczy. Związane są one z przebiegiem procesów semantycznych, generowaniem słów i programowaniem mowy, a połączenia funkcjonalne pomiędzy tymi strukturami tworzą główną sieć neuronalną dostępną w badaniach neuroobrazowych. Zmiany w tych obszarach mogą zaowocować zaburzeniami czynności sieci neuronalnej, zaburzeniami

procesów werbalnych, a także zaburzeniami fluencji słownej mierzonych testami neuropsychologicznymi [207].

Wyniki uzyskane w niniejszej pracy są więc bardzo istotne z punktu widzenia klinicznego i poznawczego, gdyż prawdopodobnie zaburzenia fluencji słownej są w pewnym sensie markerami kognitywnymi głębszych zmian strukturalnych i funkcjonalnych mózgu.

## **8.2. Wyniki analizy korpusowej**

Analiza korpusowa, wykorzystana w badaniu struktur językowych, wykazała istotne różnice w zakresie generowanych nazw własnych. Znacząca jest informacja, że dyferencje wystąpiły tylko w teście literowym. Podczas badania danych testów korpusowych zaobserwowano zubożenie użycia neologizmów przez respondentów. Proporcjonalnie odwrotna prawidłowość wystąpiła w przypadku pacjentów z zaburzeniami schizofrenicznymi, u których pojawia się mnogość zastosowania neologizmów kolokacyjnych. Zaburzenia te w schizofrenii są często związane z dysfunkcjami czołowymi, głównie w zakresie funkcji wykonawczych [20]. Prawdopodobnie w zaburzeniach poznawczych w różnych stadiach chorób otępiennych dochodzi raczej do zmniejszenia możliwości generowania słów, zarówno poprawnych, jak i niepoprawnych, co jest związane z wielopłaszczyznowym procesem uszkodzenia OUN. W przeciwieństwie do chorych na schizofrenię, u których zmniejszona produkcja słów, ale zwiększenie włączenia słów niepoprawnych, jak perswewracje i neologizmy, może wiązać się z występowaniem dysfunkcji czołowej.

W badanej grupie osób neologizmy były rzadkie i występowały tylko w teście literowym, co wskazuje, że ten test był dla chorych najtrudniejszy. W werbalizacji kobiet pojawiło się więcej nazw własnych, co nie stanowiło domeny językowej wśród mężczyzn. W kontekście pozostałych badanych parametrów korpusowych nie wystąpiły znaczące różnice językowe.

W dalszej części niniejszej dysertacji analizie empirycznej poddano wyekscerpowane słowa. Lista najczęściej występujących wiązek leksykalnych, czy też klasterów, została skategoryzowana na podstawie konkluzji wynikających z pomiaru siły frekwencji wiązek (MI). Strukturalne i funkcjonalne cechy najczęściej występujących wiązek leksykalnych uporządkowano w oparciu o drobiazgową analizę zgodności. Kongruencja umożliwiła alokację wiązek za pomocą zmodyfikowanych ram strukturalnych zaproponowanych przez Bibera i wsp. (1999) wraz z uwzględnieniem taksonomii funkcjonalnej Hylanda (2008a).

Łączna liczba słów wykorzystanych do niniejszego badania wynosi 7669, w tym 7536 słów poprawnych i 133 słów niepoprawnych. Łączna liczba słów uwzględnia persewercje, neologizmy oraz słowa nienależące do danej kategorii jak również nazwy własne. Do kategorii słów nieprawidłowych należą słowa stosowane sporadycznie, ale ogólnie uznawane za wulgaryzmy (np. kutas itp.). Ze względu na niską frekwencję występowania, wulgaryzmy stanowią w badaniu wartość statystycznie nieistotną.

Analizując trzy najczęściej występujące słowa w każdej z trzech wskazanych kategorii (słowa na literę „k”, zwierzęta oraz przedmioty ostre). Wyniki określają domenę językową jaka wystąpiła w teście fluencji słownej przeprowadzonym w grupie pacjentów z zaburzeniami otępiennymi. Zauważono, że słowo *koń* to drugie najczęściej występujące wyrażenie w pierwszych dwóch kategoriach. W kategorii słowa na literę „k” wystąpiło ono 85 razy, natomiast w kategorii narzędzia słowo *koń* wystąpiło 163 razy. Przeprowadzone badania wykazują wyraźne zróżnicowanie językowe występujące wśród kobiet i mężczyzn. Hiatus okazuje się być dość znaczący. U mężczyzn słowo *krowa* wystąpiło jako najczęstsze w dwóch kategoriach. W kategorii słowa na literę „k” występuje 39 razy i plasuje się na miejscu pierwszym oraz w drugiej kategorii “zwierzęta”, z frekwencją występowania 58, zajmuje trzecie miejsce. Co ciekawe, u kobiet słowo *krowa* występuje 65 razy, tylko w kategorii słowa na literę „k”. Natomiast słowo *koń* występuje w dwóch kategoriach w przypadku obu płci. U kobiet w pierwszej kategorii słowa na literę „k” z frekwencją 50 zajmuje drugie miejsce, a w kategorii “zwierzęta”, z frekwencją 99, zajmuje trzecie miejsce. U mężczyzn słowo *koń* wystąpiło 35 razy w pierwszej kategorii - słowa na literę „k”, zajmując tym samym drugie miejsce oraz drugie miejsce, z frekwencją 64, w kategorii “zwierzęta”.

W przeprowadzanych badaniach polskich dotyczących fluencji fonemicznej i literowej zauważono wysoką częstotliwość użycia słów na litery K (3,51%), M (2,80%), P (3,13), S [132]. Zgodnie z teorią i badaniami Szepietowskiej oraz Gawdy (2011) to K i F są najczęściej pojawiającymi się literami podczas przywoływania swobodnych skojarzeń werbalnych, gdzie słowa na literę K, według Broda i wsp. (2011), stanowią 3,51% podkorpusu IPI PAN, natomiast słowa na literę F stanowią tylko 0,30% tego właśnie korpusu.

Frekwencja występowania słów na konkretną literę w Narodowym Korpusie Języka Polskiego oraz w niniejszej grupie badawczej, jest bardzo zróżnicowana. Każdy test na fluencję słowną zawiera litery występujące często w uzusie języka polskiego, oraz takie, które występują statystycznie rzadko. Prawidłowość tę wyraźnie dostrzeżemy zestawiając wyniki badań dotyczących najczęściej występujących słów w każdej z trzech kategorii, przywołanych chociażby w niniejszej dysertacji.

W dotychczasowych analizach dotyczących badań z zakresu neuropsychologii wybranych zespołów otępiennych w tym zaburzeń typu alzheimerowskiego skupiano się głównie na etiologii zaburzeń, jak również diagnostyce i ich charakterystyce [39], Biechowska, Orłowska, 2012), zaburzeniach funkcji poznawczych i depresji w chorobie Alzheimera [43] czy też badaniach fluencji słownej [216], [218].

Niewielu jest autorów, którzy podjęli próbę analizy leksykalno-gramatycznej. Do jednych z nich należy Antczak-Kujawin (2021), która omówiła i scharakteryzowała zaburzenia sprawności leksykalno-semantycznych u pacjentów w otępieniu alzheimerowskim ze szczególnym uwzględnieniem starzenia się fizjologicznego. Badania przedstawione w jej monografii mają charakter longitudinalny, a wykryte zaburzenia struktur związków frazeologicznych mają charakter progresywny co jest jednoznaczne z postępowaniem zaburzeń funkcji poznawczych. Autorka tej monografii wykorzystowała nowatorskie narzędzie badawcze do oceny frekwencji użytych związków frazeologicznych w języku polskim. Respondenci byli zobligowani do wykonania następujących zadań:

„1) rozpoznawanie zilustrowanych nazw desygnatów, 2) nazywanie tzw. konfrontacyjne z wykorzystaniem ilustracji przedstawiających różne przedmioty, miejsca, osoby, organizmy, zjawiska, czynności, cechy istot żywych, rzeczy, 3) nazywanie tzw. oralne na



podstawie usłyszanej definicji wyrazu, 4) etykietowanie, czyli dopasowywanie związku frazeologicznego do opisu jego znaczenia, 5) aktualizowanie struktury frazeologizmu poprzez uzupełnianie zdań z tzw. „luką” właściwym składnikiem frazeologizmu, 6) wyjaśnienie znaczenia frazeologizmu” [3].

Wyniki badań pokazały jasno, tak samo jak założono w niniejszej dysertacji, iż zdiagnozowane dysfunkcje gramatyczno-semantyczne w stosowanych frazeologizmach u pacjentów cierpiących na chorobę Alzheimera mają charakter iście rozwojowy, co ma bezpośredni i pośredni wpływ rozwój zaburzeń funkcji poznawczych.

Renata Cuprych (2020) w swojej pracy doktorskiej dokonuje analizy cech biolektalnych osób starszych, skupiając się w znacznej mierze na leksykalno-semantycznej analizie języka respondentów [72]. Członkami grupy badawczej były osoby zdrowe obu płci w wieku 60-90 lat, jak również osoby z typowymi zaburzeniami mogącymi być wynikiem procesu starzenia się, tj. zaburzeniami motoryki, osłabieniem wzroku, osłabieniem słuchu czy kłopotami z koncentracją. W charakterystyce profilu językowego badanych osób uwzględnione zostały kompetencje językowe, komunikacyjne oraz aksjologiczne. Wyniki badań wskazują jednoznacznie iż wiek, a nie płeć osób badanych mają bezpośredni wpływ na produktywność językową. Podobne wnioski wykazano w niniejszej dysertacji, gdzie płeć nie stanowi czynnika różnicującego pod względem dyferencji używanego słownictwa a jest jedynie czynnikiem drugorzędym. Co ma swoje potwierdzenie w obcojęzycznej literaturze przedmiotu. Karlekar i wsp. (2018) opracowując zbiór danych korpusowych *DementiaBank* w celu przewidzenia demencji Alzheimera u pacjentów, wykazali, że nie ma znaczącej różnicy między wynikami przewidywania, biorąc pod uwagę płeć [136].

Analiza neuropsychologiczno-lingwistyczna jest tematem eksplorowanym nie tylko na gruncie polskim. Bhatia i wsp. (2020) zauważa rosnące zainteresowanie tym specyficznym obszarem zastosowań, do czego nawiązuje w swoich warsztatach na temat tworzenia korpusów dialogów medycznych. Godnym uwagi przykładem jest również korpus *Carolina Conversations Collection* (CCC), który jest jednym z nielicznych baz danych dotyczących medycznego dyskursu mówionego dostępnego w języku angielskim w kontekście badań AD. Jest to korpus dyskursu medycznego, który został skompilowany i udostępniony przez Uniwersytet Medyczny Karoliny Południowej [21], [219].

Badania języka pacjentów z zaburzeniami otępiennymi oraz z AD przy pomocy korpusowych metod obliczeniowych są stosunkowo nowe, niemniej jednak wiele prac obcojęzycznych wykazało zastosowanie algorytmów korpusowych i analiz statystycznych do predykowania chorób otępiennych oraz demencji starczej i innych chorób otępiennych [136], Mirheidari i in., 2019; Kong i in., 2019).

W literaturze przedmiotu wykazano również obcojęzyczne badania stawiające zaburzenia pamięci, jak również zaburzenia językowe jako główny marker choroby Alzheimera (AD). Stosunkowo niewiele badań nad językiem w przypadku chorych na AD określanych jest za pomocą analiz ilościowych zaburzeń mowy przy użyciu technik obliczeniowych korpusowych. Fraser, Rudzicz oraz Meltzer (2015) wykorzystali w swoich badaniach krótkie próbki narracyjne w języku angielskim, polegające na opisie obrazów, w celu wskazania istotnych czynników językowych mających znaczący wpływ na ocenę identyfikacji choroby Alzheimera [93]. Użyta została statystyczna analiza czynników, a próbki leksykalne zostały zaczerpnięte z korpusu *DementiaBank*, z którego 167 pacjentów ze zdiagnozowaną prawdopodobną chorobą Alzheimera dostarczyło 240 próbek narracyjnych, a grupa referencyjna 97 pacjentów zapewniła dodatkowe 233 próbki narracyjne. Co stanowi relewantnie podobną wielkościowo bazę korpusową do bazy danych analizowanej w niniejszej dysertacji. Wyniki badań Fraser, Rudzicz oraz Meltzer (2015) polegających na analizie zebranych próbek narracyjnych pozwalają na uzyskanie 81% dokładności w klasyfikacji chorych cierpiących na AD [93]. Do najczęstszych zaburzeń należą zaburzenia semantyczne, akustyczne, składniowe, jak również zaburzenia informacji. Wnioski wynikające z tych badań na gruncie angielskim pokrywają się z wnioskami wynikającymi z niniejszej dysertacji, stanowiąc o ich uniwersalności oraz potwierdzając przydatność badań neuropsychologiczno-lingwistycznych z wykorzystaniem korpusów językowych w diagnozie i klasyfikacji zaburzeń otępiennych.

Podobne badania przeprowadzili Benvenuti, Bolioli, Bosca oraz Mazzei et Pietro Vigorelli (2020) na gruncie włoskim. Celem tych badań było stworzenie pierwszego włoskiego korpusu swobodnych rozmów między pracownikami służby zdrowia a osobami cierpiącymi na demencję starczą, w celu zbadania określonych zjawisk językowych charakterystycznych dla tego typu zaburzeń poznawczych z obliczeniowego

punktu widzenia. Większość dotychczasowych badań dotyczących zaburzeń mowy osób z demencją opierała się na analizie jakościowej lub na badaniu kilkudziesięciu przypadków przeprowadzonych w warunkach laboratoryjnych. Stworzenie *Corpus Anchise 320* miało na celu zbadanie języka używanego przez pacjentów z demencją. Automatyczna korpusowa analiza językowa ma na celu pomoc pracownikom służby zdrowia zrozumieć niektóre cechy języka używanego przez pacjentów i wdrożyć skuteczne strategie dialogu [19].

Niewiele jest jednak badań neuropsychologiczno-językoznawczych na gruncie polskim w badanym obszarze respondentów. Same badania metodą korpusową są cały czas eksplorowane i doskonalone. Majewska-Tworek oraz Zaśko-Zielińska (2021) w swoich badaniach nad mową najstarszego pokolenia użytkowników polszczyzny wykorzystwały metody badań korpusowych opisując najistotniejsze cechy mowy rodzimych użytkowników języka polskiego w podeszłym wieku [175]. W korpusowym badaniu gerontolingwistycznym autorki skupiły się głównie na charakterystyce mowy seniorów z uwzględnieniem procesów Tip-of-the Tongue (TOT), Off Target Verbosity (OTV), Off-Topic Speech (OTS). Pierwszym zjawiskiem jest syndrom TOT, który przejawia się w trudnościach przywołania nazw własnych oraz odczuciu, że potrzeba czasu, aby móc zrealizować to zadanie. Badania wykazały, iż problemy tego typu zachodzą również u osób zdrowych. Zjawiska OTV lub też OTS dotyczą respondentów w wieku 60-90 lat, i charakteryzują się nadmierną produkcją słów wysoce nierelevantnych do tematu wypowiedzi. Zjawiska tego typu nasilają się wraz z wiekiem i mogą prowadzić do poważnych zaburzeń otępiennych.

Niniejsza dysertacja prezentuje innowatorskie podejście do analizy neuropsychologiczno-lingwistycznej, a implikacje z niej wynikające mogą posłużyć jako źródło informacji do dalszych badań w zakresie fluencji słownej u osób z dysfunkcjami poznawczymi.

## ROZDZIAŁ IX

### WNIOSKI

Na podstawie uzyskanych wyników badań sformułowano następujące wnioski:

1. Większe nasilenie zaburzeń fluencji słownej fonologicznej i kategoryalnej wiąże się z wyższym nasileniem zaburzeń poznawczych mierzonych testem MMSE i dłuższym czasem trwania zaburzeń poznawczych. Wiek i płeć nie mają istotnego znaczenia w tym zakresie w badanej grupie osób starszych z dysfunkcjami poznawczymi.
2. Biorąc pod uwagę czas trwania zaburzeń poznawczych istnieją krytyczne punkty czasowe związane z nasileniem się zaburzeń fluencji słownej, zaś na podstawie wyników testu MMSE można przewidywać pogorszenie sprawności fluencji słownej fonologicznej i kategoryalnej.
3. Wraz z postępującym osłabieniem fluencji słownej u osób z zaburzeniami poznawczymi najbardziej trwałe są słowa najczęściej występujące w korpusie Języka Polskiego. Najbardziej czułym testem w zakresie nieprawidłowo włączanych słów jest test fluencji fonologicznej, szczególnie w grupie kobiet. Dłuższy czas zaburzeń poznawczych wiąże się z większym ubytkiem generacji słów najczęściej występujących w języku polskim.
4. Testy fluencji słownej i MMSE mają wysoką zgodność wewnętrzną, co wskazuje na ich podobny mechanizm etiologiczny oraz wysoką przydatność w badaniach grupy osób starszych z zaburzeniami poznawczymi.

## ROZDZIAŁ X

### STRESZCZENIE W JĘZYKU POLSKIM

#### 10.1. Streszczenie

Celem pracy była analiza wyników testów fluencji słownej (fonologicznej i kategoryalnej) u osób z zaburzeniami funkcji w kontekście analizy korpusowej według Narodowego Korpusu Języka Polskiego. Przeanalizowano dane archiwalne 279 osób, 177 kobiet i 102 mężczyzn, z zaburzeniami poznawczymi, w wieku 50-91 (średnia u 72,00 + 10,47 ) lat. Wykorzystano wyniki testu fluencji słownej fonologicznej (test literowy) i kategoryalnej (kategorie „zwierzęta” i „przedmioty ostre”) oraz wyniki testu przesiewowego dla otępienia MMSE. Zastosowano metodę lingwistycznej analizy korpusowej polegającej na automatycznej ekscerpacji słów z testu fluencji werbalnej i pogrupowaniu ich pod względem częstotliwości występowania w języku polskim. W analizie statystycznej wykorzystano testy nieparametryczne, do porównań wyników pomiędzy dwiema grupami test U-Manna Whitneya, w analizie korelacji test R-Spearmana, w ocenie zgodności wewnętrznej stosowanych narzędzi obliczano wskaźnik alpha-Cronbacha i analizę czynnikową składowych głównych. W analizie predykcji wykorzystano test ANOVA.

Większe nasilenie zaburzeń fluencji słownej fonologicznej i kategoryalnej wiązało się z wyższym nasileniem zaburzeń poznawczych w MMSE i dłuższym czasem trwania zaburzeń poznawczych. Zidentyfikowano krytyczne punkty czasowe związane z nasileniem się zaburzeń fluencji słownej. Analiza predykcji wykazała, że na podstawie wyników testu MMSE można przewidywać pogorszenie sprawności fluencji słownej. Wraz z postępującym osłabieniem fluencji słownej w grupie osób z zaburzeniami poznawczymi najbardziej trwałe są słowa najczęściej występujące w Korpusie Języka Polskiego.

Najbardziej czułym testem w zakresie nieprawidłowo włączanych słów jest test fluencji fonologicznej, szczególnie w grupie kobiet. Dłuższy czas zaburzeń poznawczych wiązał się z większym ubytkiem generacji słów najczęściej występujących w języku polskim. Testy fluencji słownej i MMSE mają wysoką zgodność wewnętrzną, co wskazuje na ich podobny mechanizm etiologiczny oraz wysoką przydatność w badaniach grupy osób starszych z zaburzeniami poznawczymi.

## ROZDZIAŁ XI

### STRESZCZENIE W JĘZYKU ANGIELSKIM

#### 10.1. Summary

The aim of the study was to analyse the results of verbal fluency tests (phonological and categorical) in people with functional disorders in the context of corpus analysis according to the Polish National Corpus (PNC). Archival data of 279 people, 177 women and 102 men, with cognitive impairment, aged 50-91 (mean in 72.00 + 10.47) years underwent a deep analysis. The results of the verbal phonological (letter test) and categorical (categories "animals" and "sharp objects") and the results of the screening test for MMSE dementia were used. The method of linguistic corpus analysis consisting in the automatic extraction of words from the verbal fluency test and grouping them according to the frequency of their occurrence in Polish was implemented. Non-parametric tests were used in a statistical analysis, the Mann-Whitney U test was used to compare the results between the two groups, in the analysis of the R-Spearman test, in the assessment of internal compatibility of the applied tools, the alpha-Cronbach index and factor analysis of the principal components were calculated. Furthermore, the ANOVA test was used in the prediction analysis.

The greater severity of phonological and categorical verbal fluency disorders was associated with a higher severity of cognitive disorders in MMSE and a longer duration of cognitive disorders. Critical time points related to the intensification of verbal fluency disorders were identified. The analysis of predictions showed that, based on the results of the MMSE test, it is possible to predict a deterioration in the efficiency of verbal fluency. Along with the progressive weakening of verbal fluency in the group of people with cognitive disorders, the most persistent words are the most commonly used bundles according to the Polish National Corpus.

The most sensitive test of misspelled words is the phonological fluency test, especially in the group of women. A longer period of cognitive impairment was associated with a greater loss in the generation of words most commonly found in the Polish language. Verbal fluency and MMSE tests have a high internal consistency, which indicates their similar etiological mechanism and a high usefulness in the study of the group of elderly people with cognitive disorders.



## ROZDZIAŁXII

### BIBLIOGRAFIA

- [1] Abwender, D.A., Swan, J.G., Bowerman, J.T., Connolly, S.W. (2001) *Qualitative analysis of verbal fluency output: Review and comparison of several scoring methods*. *Assessment* 2001; 8: 323–336.
- [2] Akoudad, S., Wolters, F.J., Viswanathan, A., de Bruijn, R.F., van der Lugt, A., Hofman, A., Koudstaal, P.J., Ikram, M.A., Vernooij, M.W. (2016) *Association of Cerebral Microbleeds With Cognitive Decline and Dementia*. *JAMA Neurol.* 2016;73(8):934-43.
- [3] Antczak-Kujawin, J. (2021) *Zaburzenia sprawności leksykalno-semantycznej w otępieniu alzheimerowskim na tle starzenia się fizjologicznego*. Wydawnictwo Uniwersytetu Łódzkiego.
- [4] Anthony, L. *AntConc*, (2014) A freeware corpus analysis toolkit for concordancing and text analysis.
- [5] Auriacombe, S., Fabrigoule, C., Lafont, S. (2001) *Letter and category fluency in normal elderly participants: a population based study*. *Aging Neuropsychol Cogn* 2001; 8: 98–108.
- [6] Awwad, M. (1990) Equivalence and translatability of English and Arabic idioms. *Papers and Studies in Contrastive Linguistics* 26: 57–67.
- [7] Bäckman, L., Wahlin, A., Small, B.J., Herlitz, A., Winblad, B., Fratiglioni, L. (2004) *Cognitive Functioning in Aging and Dementia: The Kungsholmen Project*. *Aging Neuropsychol. Cogn.* 2004; 11: 212–244.
- [8] Baker, P. (2009) The question is, how cruel is it? Keywords, fox hunting and the house of commons. In D. Archer (Ed.), *What's in a word-list? Investigating word frequency and keyword extraction*. Farnham: Ashgate, 125-136.
- [9] Barr, W. (2003) *Neuropsychological testing of high school athletes: preliminary norms and test-retest indices*. *Arch Clin Neuropsychol* 2003; 18: 91–101.
- [10] Barcikowska, M., Kwieciński, H. (1997) *Choroba Alzheimera Od beta-amyloidu do bezradnego człowieka*. Medipress Psychiatria i Neurologia, 1997; 1(1): 16–22.
- [11] Barcikowska, M. (2011) Rozpoznanie otępienia spowodowanego chorobą Alzheimera: zalecenia grupy roboczej National Institute on Aging and the Alzheimer's Association. *Med Prakt Neurologia* 2011; 4: 18.

- [12] Barry, D., Bates, M.E., Labouvie, E. (2008) *FAS and CFL forms of verbal fluency differ in difficulty: a meta-analytic study*. *Appl Neuropsychol* 2008; 15: 97–106.
- [13] Bartolic, E., Basso, M., Schefft, B., Glauser, T., Titanic-Schefft, M. (1999) *Effects of experimentally-induced emotional states on frontal lobe cognitive task performance*. *Neuropsychologia* 1999; 37: 677–683.
- [14] Basso, M.R., Bornstein, R.A., Lang, J.M. (1999) *Practice Effects on Commonly Used Measures of Executive Function Across Twelve Months*. *Clin. Neuropsychol.* 13: 283–292.
- [15] Bączkowska, A. (2005a) Practical applications of a cognitive corpus-based approach to English prepositions. *Acta Universitatis Nicolai Copernici*, XIV.
- [16] Bączkowska, A. (2005b) Nauczanie phrasal verbs. *JOWS*, 4.
- [17] Bączkowska, A. (2008) Contrasting Polish and English prepositions: the case of w/na deszczu vs in the rain and w/na słońcu vs. in the sun. In B. Lewandowska-Tomaszczyk (ed.) *Corpus Linguistics, Computer Tools, and Applications- State of the Art*. Frankfurt: Peter Lang.
- [18] Bączkowska, A. (2018) Functional analysis of lexical bundles in medical discourse, „*Conversatoria Linguistica*”, XII, s. 7–25. Benson, M. Benson, E. and Ilson R. (1986) *The BBI Combinatory Dictionary of English*. Amsterdam/ Philadelphia: John Benjamins Publishing Company.
- [19] Benvenuti, N., Bolioli, A., Bosca, A., Mazzei et Pietro Vigorelli, A. (2020) The “Corpus Anchise 320” and the analysis of conversations between healthcare workers and people with dementia. In: *Proceedings of the Seventh Italian Conference on Computational Linguistics CLiC-it 2020*. 51-57.
- [20] Berberian, A.A., Moraes, G.V., Gadelha, A., Brietzke, E., Fonseca, A.O., Scarpato, B.S., Vicente, M.O., Seabra, A.G., Bressan, R.A., Lacerda, A.L. (2016) Is semantic verbal fluency impairment explained by executive function deficits in schizophrenia? *Braz J Psychiatry*. 2016;38(2):121-6.
- [21] Bhatia, P., Lin, S., Gangadharaiah, R., Wallace, B., Shafran, I., Shivade, C., Du, N., and Diab, M. (2020) editors. *Proceedings of the First Workshop on Natural Language Processing for Medical Conversations*, Online, July 2020. *Association for Computational Linguistics*.
- [22] Bestegen, S. and Granger, Y. (2014) The use of collocations by intermediate vs. advanced non-native writers: A bigram-based study. *IRAL- International Review of Applied Linguistics in Language Teaching* 52(3):229-252.
- [23] Biber, D. (1988) *Variation across Speech and Writing*. New York, NY: Cambridge University Press.

- [24] Biber, D. (1990) Methodological issues regarding corpus-based analyses of linguistic variation. *Literary and Linguistic Computing* 5.257-269.
- [25] Biber, D. Finegan, E. Oostdijk, N. de Haan, P. (1994) *Intratextual variation within medical research articles Corpus-based research into language*. Amsterdam: Rodopi, 201- 222.
- [26] Biber, D., Conrad, S. and Reppen, R. (1998) *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge University Press.
- [27] Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. (1999) *The Longman Grammar of Spoken and Written English*. London: Longman.
- [28] Biber, D., Reppen, R. and Conrad. S. (2002) Developing linguistic literacy: Perspectives from corpus linguistics and multi-dimensional analysis. *Journal of Child Language* 29.449-488.
- [29] Biber, D., Conrad, S., and Cortes, V. (2003) *Lexical bundles in speech and writing: An initial taxonomy*. U A. Wilson, P. Rayson oraz T. McEnery (red.) *Corpus Linguistics by the Lune: A Festschrift for Geoffrey Leech*. Frankfurt am Main: Peter Lang, 71-92.
- [30] Biber, D., Conrad, S., and Cortes, V. (2004) *If you look at : Lexical bundles in university teaching and textbooks*. *Applied Linguistics* 25. 371-405.
- [31] Biber, D. and Conrad, S. (2005) The Frequency and Use of Lexical Bundles in Conversation and Academic Prose. In W. Teubert and M. Mahlberg (eds.), *The corpus approach to lexicography*, Thematischer Teil von *Lexicographica*. Internationales Jahrbuch für Lexicographie, 20, 2004, 56-71.
- [32] Biber, D. (2006) *University Language. A Corpus-Based Study of Spoken and Written Registers*. Amsterdam/ Philadelphia: John Benjamins Publishing Company.
- [33] Biber, D. and Gray. B. (2010) Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes* 9.2-20.
- [34] Biber, D. and Reppen, R. (2015) *The Cambridge handbook of English corpus linguistics*. Cambridge: Cambridge University Press.
- [35] Biel, Ł. (2008) Legal terminology in translation practice: dictionaries, googling or discussion forums? *SKASE Journal of Translation and Interpretation*.
- [36] Biel, Ł. (2009) Organization of background knowledge structures in legal language and related translation problems. In *Comparative Legilinguistics. International Journal for Legal Communication* 1.
- [37] Biel, Ł. (2014) *Lost in the Eurofog: The Textual Fit of Translated Law*. Frankfurt: Peter Lang.
- [38] Bolinger, D. (1976) Meaning and memory. *Forum Linguisticum* 1. 1-14.

- [39] Borkowska, A., Sobów, T. (2005) Ocena neuropsychologiczna w diagnostyce i różnicowaniu otępienia czołowo-skroniowego. *Neurologia i Neurochirurgia Polska* 2005; 39, 6: 466–475.
- [40] Borkowska, A. (2006) *Neuropsychiatria i Neuropsychologia/Neuropsychiatry and Neuropsychology*. *Neuropsychiatria i Neuropsychologia* 2006; 1, 1: 31–42.
- [41] Borkowska, A., Tomaszewska, M., Wilkość, M. (2006) *Korzystne efekty dodania walproinianu do leczenia donepezilem w chorobie Alzheimera – doniesienie wstępne*. *Farmakoter Psychiatry Neurol* 22: 19–25.
- [42] Borkowska, A., Drózdź, W., Jurkowski, P., Rybakowski, J.K. (2007) *The Wisconsin Card Sorting Test and the N-back test in mild cognitive impairment and elderly depression*. *World J. Biol. Psychiatry*. 2007; 12: 1.7.
- [43] Borkowska, A., Daniluk, J. (2008) Zaburzenia funkcji poznawczych i depresja w chorobie Alzheimera, zagadnienia neurobiologiczne. *Psychiatria* 2008;5(2):43-50.
- [44] Borkowska, A. (2011) *Pamięć i jej zaburzenia w MCI*. Medical Education. Warszawa.
- [45] Borzym, A. (2010) *Oddziaływania nefarmakologiczne w otępieniu*. W: Choroba Alzheimera. Parnowski T. (red.). Wydawnictwo Lekarskie Warszawa: PZWL; 2010: 49–51.
- [46] Bousfield, W.A., Sedgewick C.H.W. (1994) *An analysis of sequences of restricted associative responses*. *J. Gen. Psychol.* 1944; 30: 149–165.
- [47] Bower, L., Pearson, J. (2002) *Working with Specialized Language: A Guide to Using Corpora*. London. Routledge.
- [48] Bowker, L. and Barlow, M. (2008) 'A comparative evaluation of bilingual concordances and translation memory system', in E. Yuste Rodrigo (ed.) *Topics in Language Resources for Translation and Localisation*. Amsterdam/ Philadelphia: John Benjamins Publishing Company.
- [49] Brezina, V., McEnery, T., Wattam, S. (2015) *Collocations in context: A new perspective on collocational networks*. *International Journal of Corpus Linguistics* 20:2, 139-173.
- [50] Brickman, A.M., Paul, R.H., Cohen, R.A., Williams, L.M., MacGregor, K.L., Jefferson, A.L., Tate, D.F., Gunstad, J., Gordon, E. (2005) *Category and letter verbal fluency across the adult lifespan: relationship to EEG theta power*. *Arch. Clin. Neuropsychol.* 2005; 20: 561–573.
- [51] Broda, B., Piasecki, M. (2011) Parallel, Massive Processing in SuperMatrix -- a General Tool for Distributional Semantic Analysis of Corpora. *International Journal of Data Mining, Modelling and Management*.

- [52] Budvytyte, A. (2003) Der axiologische Aspekt der deutschen und litauischen Somatismen. In Burger, H., A. Häcki Buhofer & G. Gréciano (eds.) *Flut von Texten – Vielfalt der Kulturen. Ascona 2001 zu Methodologie und Kulturspezifität der Phraseologie*, 255–266. Baltmannsweiler: Schneider Verlag Hohengehren.
- [53] Bujnowska-Fedak, M.M., Grata-Borkowska, U., Sapilak, B.J. (2012) *Otępienie i depresja u pacjentów w podeszłym wieku w Praktyce Lekarza Rodzinnego*. Family Medicine & Primary Care Review 2012, 14, 3: 349–353
- [54] Burger, H., Buhofer, A., Sialm, A. (eds.) (1982) *Handbuch der Phraseologie*. Berlin: de Gruyter.
- [55] Cardebat, D., Demonet, J.F., Viillard, G., Faure, S., Puel, H., Celsis, P. (1996) *Brain functional profiles in formal and semantic fluency tasks: a SPECT study in normal*. Brain Lang. 1996; 52: 305–313.
- [56] Chan, A.S., Poon, M.W. (1999) *Performance of 7- to 95-year-old individuals in a Chinese version of the category fluency test*. J Int Neuropsychol Soc 1999; 5: 525–533.
- [57] Chan, R.C.K., Chen, E.Y.H. (2004) *Development of a Chinese verbal fluency test for the Hong Kong psychiatric setting*. Hong Kong J Psychiatry 2004; 14: 8–11.
- [58] Chertkow, H., Bub, D. (1990) Semantic memory loss in dementia of Alzheimer's type. What do various measures measure? *Brain* 1990; 113: 397–417.
- [59] Chlebda, W. (1991) *Elementy frazematyki. Wprowadzenie do frazeologii nadawcy*, Opole.
- [60] Conklin, K. and Schmitt, N. (2008) Formulaic Sequences: Are They Processed More Quickly than Nonformulaic Language by Native and Nonnative Speakers? *Applied Linguistics* 29/1:72–89.
- [61] Corpas Pastor, G. (2016) *Computerised and corpus-based approaches to phraseology: Monolingual and multilingual perspectives*. Geneva, Switzerland: Tradulex.
- [62] Cortes, V. (2004) Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes* 23, 397-423.
- [63] Cortes, V. (2006) Teaching lexical bundles in the disciplines: An example from a writing intensive history class. *Linguistics and Education*, 17, 391-406.
- [64] Coulmas, F. (1979) On the sociolinguistic relevance of routine formulae. *Journal of Pragmatics* 3 (3-4):239-266.
- [65] Cowie, A. P., Mackin, R. and McCaig, I. R. (1983) General introduction. In *The Oxford Dictionary of Current Idiomatic English*, ed. A. P. Cowie, Mackin, R. and McCaig, R., x-xvii. Oxford: Oxford University Press.

- [66] Cowie, A. P. (1994) *Phraseology*. In RE Asher (ed.) *The Encyclopedia of Language and Linguistics* Volume 6. Oxford and New York: Pergamon.
- [67] Cowie, A. P., Howarth, P. (1996) *Phraseological competence and written proficiency*. In *Language and Education*, ed. G.M. Blue and R. Mitchell, 80-93.
- [68] Cowie, A. P. (ed.) (1998) *Phraseology. Theory, analysis, and applications*. Oxford Studies in Lexicography and Lexicology. Oxford: Oxford University Press.
- [69] Cowie, A. P. (2004) *Phraseology in formal academic prose*. In Jan Aarts, Inge de Moonink and Herman Wekker (eds.), *Studies in English Language Teaching in Honour of Flor Aarts*, Amsterdam: Rodopi, 43-56.
- [70] Coxhead, A. J. (2008) *Phraseology and English for academic purposes. Challenges and opportunities*. In Meunier, F. & Granger, S. (eds.) *Phraseology in Foreign Language Learning and Teaching*. Amsterdam/ Philadelphia: John Benjamins Publishing Company, 149-161.
- [71] Crossley, M., D'arcy, C., Rawson, N.S.B. (1997) *Letter and category fluency in community-dwelling Canadian seniors: A comparison of normal participants to those with dementia of the Alzheimer or vascular type*. *J. Clin. Exp. Neuropsychol.* 1997; 19: 52–62.
- [72] Cuprych, R. (2020) Nieopublikowana praca doktorska pod tytułem: *Charakterystyczne cechy biolektalne osób starszych*. Uniwersytet Jana Kochanowskiego w Kielcach.
- [73] Daniluk, B., Szepietowska, E.M. (2009a) *Płynność semantyczna i literowa osób w różnych fazach dorosłości – część I*. *Ann UMCS Sect J* 2009a; 22: 97–110.
- [74] Daniluk, B., Szepietowska, E.M. (2009b) *Płynność semantyczna i literowa osób w różnych fazach dorosłości – czynniki modyfikujące wykonanie zadań fluencji słownej – część II*. *Ann UMCS Sect J* 2009b; 22: 111–123.
- [75] De Cock, S. (2000) *Repetitive phrasal chunkiness and advanced EFL speech and writing*. In Mair C. and Hundt M. (eds.) *Corpus Linguistics and Linguistic Theory*. Amsterdam: Rodopi, 51-68.
- [76] De Cock, S. (2004) *Preferred sequences of words in NS and NNS speech*. *Belgian Journal of English Language and Literature* 2, 225-246.
- [77] Dierckx, E., Elgelborghs, S., De Raedt, R., De Deyn, P.P., Ponjaert- Kristoffersen, J. (2007) *Mild cognitive impairment: what's in a name?* *Gerontology* 2007, 53: 28.35.
- [78] Dikmen, S.S., Heaton, R.K., Grant, I., Temkin, N.R. (1999) *Test-retest reliability and practice effects of expanded Halstead-Reitan Neuropsychological Test Battery*. *J. Int. Neuropsychol. Soc.* 1999; 5: 346–356.

- [79] Dobrovol'skij, D. (1997) Kontrastive Idiomatik Deutsch-Russisch. Zur lexicographischen Konzeption. In Gréciano, G. and A. Rothkegel (eds.) *Phraseeme in Kontext und Kontrast*, 45–59. Bochum: Brockmeyer.
- [80] Domagała, A., 2007. Zachowania językowe w demencji – struktura wypowiedzi w chorobie Alzheimerera, Lublin.
- [81] Doroszewski, W. (1969 – 1978) (red.) *Słownik języka polskiego*, t. I-XI, Warszawa.
- [82] Durco, P. (1994) *Probleme der allgemeinen und kontrastiven Phraseologie: Am Beispiel Deutsch und Slowakisch*. Heidelberg: Groos.
- [83] Ebrahimipour, M., Shahbeigi, S., Jenabi, M. (2008) *Verbal fluency performance in patients with multiple sclerosis*. *Iran J Neurol* 2008; 7: 138–142.
- [84] Ellis, N. (2002) Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24(2). 143-188.
- [85] Ellis, N. (2008) Usage-based and form-focused language acquisition: The associative learning of constructions, learned attention and the limited L2 endstate. In P. Robinson and N. Ellis (eds.) *Handbook of Cognitive Linguistics and Second Language Acquisition*. New York. Routledge.
- [86] Evert, S. (2013) Tools for the acquisition of lexical combinatorics. In R. H. Gouws, U. Heid, W. Schweickard, and H. E. Wiegand (eds.), *Dictionaries. An International Encyclopedia of Lexicography. Supplementary volume: Recent Developments with Focus on Electronic and Computational Lexicography (HSK 5.4)*, chapter 104, pages 1415-1432. Mouton de Gruyter, Berlin, New York.
- [87] Evert, S. Heinrich, P. Henselmann, K. Rabenstein, U. Scherr, E. Schmitt, M. Schröder, L. (2019) Combining machine learning and semantic features in the classification of corporate disclosures. *Journal of Logic, Language and Information*, 28, 309–330.
- [88] Firth, J. R. (1957) *Papers in linguistics, 1934-51*. London: Oxford University Press.  
Lyons, John. 1967. Firth's theory of meaning. In memory of J. R. Firth, ed. by C. E. Bazell et al., 288-302. London: Longmans.
- [89] Firth, J. R. (1968) Linguistic analysis as a study of meaning. In *Selected papers of J. R. Firth 1952-1959*, ed. F. R. Palmer. London: Longman, 12-26.
- [90] Flowerdew, L. (2012) *Corpora and Language Education*. Basingstoke: Palgrave Macmillan.
- [91] Francis, W., Kučera, H. (1964) *A Standard Corpus of Present-Day Edited American English, for Use with Digital Computers (Brown)*. Providence RD: Brown University.
- [92] Francis, G., Hunston, S., Manning, E. (1996) *Grammar Patterns 1: Verbs*. London: Harper Collins.

- [93] Fraser, K. C., Rudzicz, F., Meltzer, J. A. (2015) Linguistic Features Identify Alzheimer's Disease in Narrative Speech. *Journal of Alzheimer's Disease* 49 (2016) 407–422.
- [94] Gablasova, D., Brezina, V., McEnery, T. (2017) *Collocations in Corpus-Based Language Learning Research: Identifying, Comparing, and Interpreting the Evidence*. Special Issue: Currents in Language Learning Series: Experimental, Computational, and Corpus-Based Approaches to Language Learning: Evidence and Interpretation. vol. 67.
- [95] Gladsjo, J.A., Schuman, C.C., Evans, J.D., Peavy, G.M., Miller, S.W., Heaton, R.K. (1999) *Norms for letter and category fluency: demographic corrections for age, education, and ethnicity*. *Assessment* 1999; 6: 147–178.
- [96] Gläser, R. (1999) Zur Wiedergabe von Phraseologismen in englischen und französischen Übersetzungen ausgewählter Prosawerke von Christa Wolf. In Sabban, A. (ed.) *Phraseologie und Übersetzen*, 99–118. Bielefeld: Aisthesis Verlag.
- [97] Gliwa R., (2018) Fluencja słowna w otępieniu naczyniopochodnym – ujęcie kliniczne, [w:] *Teoria i praktyka logopedyczna. Wybrane zagadnienia*, red. E. Gacka, M. Kaźmierczak, Łódź, s. 83–102.
- [98] Gonzalez Fernandez, B., Schmitt, N. (2015) How much collocation knowledge do L2 learners have?: The effects of frequency and amount of exposure. *International Journal of Applied Linguistics* 166:1.
- [99] Goulet, P., Joannette, Y., Sabourin, L., Giroux, F. (1997) *Word fluency after a right-hemisphere lesion*. *Neuropsychologia* 1997; 35; 12: 1565–1570.
- [100] Gózdź-Roszkowski, S. (2011) *Patterns of Linguistic Variation in American Legal English*. Bern: Peter Lang.
- [101] Grabowski, Ł. (2013) 'Register variation across English pharmaceutical texts: A corpus driven study of keywords, lexical bundles and phrase frames in patient information leaflets and summaries of product characteristics.' *Procedia – Social and Behavioral Sciences* 95, 391-401.
- [102] Grabowski, Ł. (2014) On Lexical Bundles in Polish Patient Information Leaflets: A Corpus- Driven Study. *Studies in Polish Linguistics* 9 (1). 21-43.
- [103] Granger, S. and Tribble, C. (1998) *Exploiting learner corpus data in the classroom: Form focused instruction and data-driven learning*. U S. Granger (red.) *Learner Language on Computer*. Harlow: Longman.
- [104] Granger, S. and Meunier F. (2008) *Phraseology in Foreign Language Learning and Teaching*. Amsterdam/ Philadelphia: John Benjamins Publishing Company.



- [105] Granger S. (2009) The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation. In Aijmer, K. (ed.) *Corpora and Language Teaching*. Amsterdam/ Philadelphia: John Benjamins Publishing Company, 13-32.
- [106] Gries, S. Th. and Mukherjee, J. (2010) Lexical gravity across varieties of English An ICE based study of n-grams in Asian Englishes. *International Journal of Corpus Linguistics* 15(4). 520–548.
- [107] Gries, S. Th. (2013) *Statistics for linguistics with R: a practical introduction*, 2nd edn. Berlin/Boston: Walter de Gruyter.
- [108] Gries, S. Th. (2016) *Corpus-based Approaches to Construction Grammar*. Amsterdam/ Philadelphia: John Benjamins Publishing Company, 2.
- [109] Gruen, A.K., Frankle, B.C., Schwartz, R. (1990) *Word fluency generation skills of head-injured patients in an acute trauma center*. *J. Commun. Disord.* 1990; 23: 163–170.
- [110] Guernet Steiner, V.A., Mansur, L., Brucki, S.M. (2008) *Phonemic verbal fluency and age. A preliminary study*. *Dement Neuropsychol* 2008; 2: 328–332.
- [111] Gustaw, K., Woźnica, I., Bylina, J. (2008) *Rozpowszechnienie zespołów otępiennych, w tym choroby Alzheimerera, w populacji mieszkańców województwa lubelskiego*. *Medycyna Ogólna*. 14 (XLIII): 4.
- [112] Hakuta, K. (1974) A Preliminary Report on the Development of Grammatical Morphemes in a Japanese Girl Learning English as a Second Language. Working Papers on Bilingualism 3. In E. Hatch (ed.) *Studies in Second Language Acquisition*. Rowley/ Mass: Newbury House Publishers, 18-43.
- [113] Halliday, M. A. K. (1994) *An Introduction to Functional Grammar (2nd edition)*. London: Edward Arnold.
- [114] Halliday, M. A. K. (2004) *An Introduction to Functional Grammar (3rd edition)*. Oxford University Press. New York.
- [115] Harciarek, M., Williamson, J.B., Biedunkiewicz, B. (2012) *Risk factors for selective cognitive decline in dialyzed patients with endstage renal disease: evidence from verbal fluency analysis*. *J Int Neuropsychol Soc* 2012; 18: 162–167.
- [116] Hausmann, F. J. (1985) Kollokationen im Deutschen Wörterbuch. Ein Beitrag zur Theorie des lexikographischen Beispiels. In: Henning Bergenholtz / Joachim Mugdan (Hrsg.): *Lexikographie und Grammatik. Akten des Essener Kolloquiums zur Grammatik im Wörterbuch vom 28. bis 30.06.1984*. Tübingen: Max Niemeyer, 118–129.

- [117] Hausmann, F.J. and Blumenthal, P. (2006) Presentation: collocations, corpus, dictionaries. In *Collocations, Corpus, Dictionaries*, P. Blumenthal and F. J. Hausmann (eds.), 3-13. *Language Francaise* 150.
- [118] Herzyk, A. (2005) *Wprowadzenie do neuropsychologii klinicznej*. Warszawa: Wydawnictwo Naukowe Scholar.
- [119] Hessky, R. (1987) *Phraseologie. Linguistische Grundlagen und kontrastives Modell deutschungarisch*. Tübingen: Max Nimeyer.
- [120] Hillis, A., Caramazza, A. (1995) *Representation of grammatical categories of words in the brain*. *Journal of Cognitive Neuroscience*, 7, 3, 396-406.
- [121] Hodges, J.R., Patterson, K., Oxbury, S. (1992) *Semantic dementia. Progressive fluent aphasia with temporal lobe atrophy*. *Brain* 1992; 115: 1783–1806.
- [122] Hoey, M. (2005) *Lexical Priming: A new theory of words and language*. London: Routledge.
- [123] Howarth, P. A. (1996) *Phraseology in English academic writing: Lexicographica Series Maior*, 75. Tübingen: Max Niemeyer.
- [124] Howarth, P. (1998) *Phraseology and second language proficiency*. *Applied Linguistics* 19(1), 24-44.
- [125] Hunston, S. (2002) *Corpora in applied linguistics*. Cambridge, UK: Cambridge University Press.
- [126] Huo, Z., Yu, L., Yang, J., Zhu, Y., Bennett, D.A., Zhao, J. (2020) *Brain and blood metabolome for Alzheimer's dementia: findings from a targeted metabolomics analysis*. *Neurobiol Aging*. 2020;86:123-133. Erratum in: *Neurobiol Aging*. 2020;91:169.
- [127] Hyland, K. (2008a) 'As can be seen': Lexical bundles and disciplinary variation. *English for specific purposes* 27, 4-21.
- [128] Hyland, K. (2008b) *Academic clusters: Text patterning in published and postgraduate writing*. *International Journal of Applied Linguistics* 18, 41-62.
- [129] Ito, E., Hatta, T., Ito, Y. (2004) *Performance of verbal fluency tasks in Japanese healthy adults. Effect of gender, age and education on the performance*. *Jpn J Neuropsychol* 2004; 20: 254–263.
- [130] Jablonkai, R. (2010) *English in the context of European integration: A corpus-driven analysis of lexical bundles in English EU documents*. *English for Specific Purposes*, 29(4), 253-267.
- [131] Jespersen, O. (1924) *The Philosophy of grammar*. London: Allen and Unwin.

- [132] Jodzio, K.(2006) Neuropoznawcze korelaty spadku fluencji słownej po udarze prawej półkuli mózgu. *Studia Psychologiczne* 2006; 44: 5–18.
- [133] Jokinen, H., Koikkalainen, J., Laakso, H.M., Melkas, S., Nieminen, T., Brander, A., Korvenoja, A., Rueckert, D., Barkhof, F., Scheltens, P., Schmidt, R., Fazekas, F., Madureira, S., Verdelho, A., Wallin, A., Wahlund, L.O., Waldemar, G., Chabriat, H., Hennerici, M., O'Brien, J., Inzitari, D., Lötjönen, J., Pantoni, L., Erkinjuntti, T. (2020) Global Burden of Small Vessel Disease-Related Brain Changes on MRI Predicts Cognitive and Functional Decline. *Stroke*. 2020;51(1):170-178.
- [134] Kałwa, A., Rzewuska, M., Piróg-Balcerzak, A. (2008) *Analiza zależności między przebiegiem schizofrenii w początkowym okresie choroby a dysfunkcjami poznawczymi po średnio 31 latach od zachorowania*. Farmakoter Psychiatr Neurol 2008; 1: 25–32.
- [135] Karbe, H., Kertesz, A., Polk, M. (1993) *Profiles of language impairment in primary progressive aphasia*. Arch Neurol 1993; 50: 193–201.
- [136] Karlekar, S., Niu, T., Bansal, M. (2018) Detecting linguistic characteristics of Alzheimer's dementia by interpreting neural models. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: *Human Language Technologies*, Volume 2 (Short Papers), pages 701–707, New Orleans, Louisiana, June 2018. Association for Computational Linguistics.
- [137] Kashiha, H. and Chan, S. H. (2015) An exploration of lexical bundles in academic lectures: examples from hard and soft sciences. *Journal of Asia TEFL* 10 (4).
- [138] Katzev, M., Tuscher, O., Henning, J., Weiller, C., Kaller, Ch. (2013) Revisiting the functional specialization of IFG in phonological and semantic fluency: the crucial role of task demands and individual ability. *The Journal of Neuroscience*, 33 (18), 7837-7845.
- [139] Kave, G. (2005) *Phonemic fluency, semantic fluency, and difference scores: Normative data for adult Hebrew speakers*. J. Clin. Exp. Neuropsychol. 2005; 27: 690–699.
- [140] Keil, K., Kasniak, A.W. (2002) *Examining executive function in individuals with brain injury: A review*. Aphasiology 2002; 16: 305–335.
- [141] Kempler, D., Teng, E.L., Dick, M., Taussig, I.M., Davis, D.S. (1998) *The effects of age, education, and ethnicity on verbal fluency*. J. Int. Neuropsychol. Soc. 4: 531–538.
- [142] Kennedy, G. (1998) *An Introduction to Corpus Linguistics: Studies in Language and Linguistics*. London: Longman.
- [143] Kiejna, A. *Epidemiologia choroby Alzheimerera*. (2011) W: *Choroby otępienie. Teoria i praktyka*. Continuo. Leszek J. (red.). Wrocław.

- [144] Kielar-Turska, M., Byczewska-Konieczna, K. (2014) *Specyficzne właściwości posługiwania się językiem przez osoby w wieku senioralnym*, [w:] *Biomedyczne podstawy logopedii*, red. S. Milewski, J. Kuczkowski, K. Kaczorowska-Bray, Gdańsk, s. 429–443.
- [145] Kivircik Akdede, B.B., Alptekin, K., Akvardar, Y. (2005) *Quality of life in patients with obsessive-compulsive disorder: relations with cognitive functions and clinical symptoms*. *Turk Psikiyatri Derg* 2005; 16: 13–19.
- [146] Keilp, J.G., Gorlyn, M., Alexander, G., Stern, Y., Prohovnik, J. (1999) *Cerebral blood flow patterns underlying the differential impairment in category and letter fluency in Alzheimer's disease*. *Neuropsychologia* 1999; 37: 1251–1261.
- [147] Kierzyńska, A., Kaźmierski, R., Kozubski, W. (2011) *Wykształcenie a zaburzenia funkcji poznawczych u osób z chorobą Parkinsona*. *Neurol Neurochir Pol* 2011; 45: 24–31.
- [148] Kim, Y-H. (2009) *An investigation into native and non-native teachers' judgments of oral English performance: A mixed methods approach*. McGill University, Montreal, Quebec, Canada.
- [149] Kilgarriff, A., Grefenstette, G. (2003) *Web as Corpus: Introduction to the Special Issue*. *Computational Linguistics* 29 (3).
- [150] Kjellmer, G. (1987) *Aspects of English Collocations*. Amsterdam. Rodopi.
- [151] Kjellmer, G. (1994) *Some thoughts on collocational distinctiveness*. In J. Aarts and W. Meijs: 163-172.
- [152] Koivisto, K., Helkala, E.L., Reinikainen, K.J. (1992) *Population-based dementia screening program in Kuopio: the effect of education, age, and sex on brief neuropsychological tests*. *J Geriatr Psychiatry Neurol* 1992; 5: 162–171.
- [153] Kopaczyk, J. (2016) *Review of: Language in Scotland. Corpus-based studies*, by W. Anderson (ed.) (2013). *Northern Scotland* 7(1): 112-117.
- [154] Korhonen, J. (1989) *Zur syntaktischen Negationskomponente in deutschen und finnischen Verbidiomen*. In Gréciano, G. (ed.) *EUROPHRAS 88: Phraséologie contrastive: Actes du Colloque International Klingenthal-Strasbourg, 12–16 mai 1988*, 253–264. Strasbourg: Université des Sciences Humaines.
- [155] Kosmidis, M.H., Vlahou, C.H., Panagiotaki, P. (2004) *The verbal fluency task in the Greek population: Normative data, and clustering and switching strategies*. *J Int Neuropsychol Soc* 2004; 10: 164–172.
- [156] Krenn, B. (2000) *Empirical implications on lexical association measures*. In Evert, S., U. Heid, E. Lehmann and C. Rohrer (eds.) *Proceedings of the Ninth EURALEX International Congress EURALEX 2000*, 359-371. Stuttgart: EURALEX.

- [157] Kwon, S.J., Kim, H.S., Han, J.H., Bae, J.B., Han, J.W., Kim, K.W. (2021) *Reliability and Validity of Alzheimer's Disease Screening With a Semi-automated Smartphone Application Using Verbal Fluency*. *Front Neurol.* 2021;12:684902. doi: 10.3389/fneur.2021.684902. PMID: 34305793.
- [158] Latkowski, J.B., Lukas, W., red. (2004) *Medycyna rodzinna*. Warszawa: Wydawnictwo Lekarskie PZWL.
- [159] Lee T., Yuen, K., Chan, C. (2002) *Normative data for neuropsychological measures of fluency, attention, and memory measures for Hong Kong Chinese*. *J Clin Exp Neuropsychol* 24: 615–632.
- [160] Lee, N.H. (2010) What are Singaporean babies actually listening to? Prosodic aspects of multilingual motherese in Singapore. In M. Cruz-Ferreira (ed.) *Multilingual Norming*. Peter Lang. 141-172.
- [161] Leech, G. (1992) Corpora and theories of linguistic performance, *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82*, (ed.) J. Svartvik. Berlin; New York, NY: Mouton de Gruyter. 105-122.
- [162] Leech, G. (1997) Introducing corpus annotation, in Garside, R., Leech, G., and McEnery, A. (eds.) *Corpus Annotation: Linguistic Information from Computer Text Corpora*. Longman, London, pp. 1-18.
- [163] Leech, G. (2000) 'Grammars of spoken English: New implications of corpus-oriented research', *Language Learning*, 50: 3, 675-724.
- [164] Lessard-Clouston, M. (2010) Theology lectures as lexical environments: A case study of technical vocabulary use. *Journal of English for Academic Purposes* 9(4):308-321.
- [165] Levine, A.J., Miller, E.N., Becker, E.T., Selnes, O.A., Cohen, B.A. (2004) *Normative Data for Determining Significance of Test-Retest Differences on Eight Common Neuropsychological Instruments*. *Clin. Neuropsychol.* 2004; 18: 373–384.
- [166] Lewandowska-Tomaszczyk, B. (2005) *Podstawy językoznawstwa korpusowego*. Łódź: Wydawnictwo UŁ.
- [167] Lewicki, A.M. (1983) Składnia związków frazeologicznych, *Biuletyn Polskiego Towarzystwa Językoznawczego*, XL.
- [168] Loonstra, A.S., Tarlow, A.R., Sellers, A.H. (2001) *COWAT metanorms across age, education, and gender*. *Appl Neuropsychol* 2001; 8: 161–166.
- [169] López, S.R., Taussig, M.I. (1991) *Cognitive-intellectual functioning of Spanish-speaking impaired and nonimpaired elderly: implications for culturally sensitive assessment*. *J Consult Clin Psych* 1991; 3: 448–454.

- [170] Lubrini, G., Periañez, J.A., Laseca-Zaballa, G., Bernabéu-Brotons, E., Ríos-Lago, M. (2021) *Verbal Fluency Tasks: Influence of Age, Gender, and Education and Normative Data for the Spanish Native Adult Population*. Arch Clin Neuropsychol. 2021 Jul 29;acab056. doi: 10.1093/arclin/acab056. Epub ahead of print. PMID: 34323264.
- [171] L'Homme, M. C. Bertrand, C. (2000) Specialized Lexical Combinations: Should They Be Described as Collocations or in Terms of Selectional Restrictions? In *Proceedings. Ninth EURALEX International Congress*, 497-506.
- [172] Li, J. and Schmitt, N. (2009) The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing* 18: 85-102.
- [173] Łojek E., Stańczak J. (2010) Podręcznik do kalifornijskiego testu uczenia się językowego CVLT D.C. Delisa, J.H. Kramera, E. Kaplan i B. Ober. Polska normalizacja, Warszawa.
- [174] Łuczywek, E., Fersten, E. (1992) Poziom fluencji słownej przy różnych uszkodzeniach mózgu, *Studia Psychologiczne* XXX, 1–2, s. 89–98.
- [175] Majewska-Tworek, A., Zaśko-Zielińska, M. (2021) Mowa najstarszego pokolenia użytkowników polszczyzny. Badanie z wykorzystaniem metody językoznawstwa korpusowego - wstępne informacje o projekcie. W: Milewski, S., Kaczorowska-Bray, K., Kamińska, B. *STAROŚĆ, JĘZYK, KOMUNIKACJA. Nowe obszary logopedii*. Wydawnictwo Uniwersytetu Gdańskiego. Gdańsk 171-187.
- [176] Martin, A. (2001) *Functional neuroimaging of semantic memory*. W: Cabeza R, Kingstone A. red. *Handbook of functional neuroimaging of cognition*. Cambridge, MA: MIT Press; 2001, 153–186.
- [177] Marques, J., Canessa, N., Siri, S., Catricala, E., cappa, S. (2008) Conceptual knowledge in the brain: fMRI evidence for a featural organization. *Brain Research*, 1194, 90-99.
- [178] Magierska, J., Magierski, R., Fendler, M. (2012) *Clinical application of the Polish adaptation of the Montreal Cognitive Assessment (MoCA) test in screening for cognitive impairment. Zastosowanie polskiej adaptacji Montrealskiego Testu do Oceny Funkcji Poznawczych (MoCA) w przesiewowej ocenie funkcji poznawczych*. *Neurol Neurochir Pol* 2012; 46: 130–139.
- [179] Mathuranath, P.S., George, A., Cherian, P.J. (2003) *Effects of age, education and gender on verbal fluency*. *J Clin Exp Neuropsychol* 2003; 25: 1057–1064.
- [180] Mayr, U., Kliegl, R. (2000) *Complex semantic processing in old age: does it stay or does it go?* *Psychol. Aging.*, 15: 29–34.

- [181] Mayr, U. (2002) On the dissociation between clustering and switching in verbal fluency: comment on Troyer, Moscovitch, Winocur, Alexander and Stuss. *Neuropsychologia* 2002; 40: 562–566.
- [182] McDonnell, M., Dill, L., Panos, S., Amano, S., Brown, W., Giurgius, S., Small, G., Miller, K. (2020) Verbal fluency as a screening tool for mild cognitive impairment. *Int Psychogeriatr.* 2020; 32(9):1055-1062.
- [183] McEnery, T., Wilson, A. (1996) *Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- [184] McEnery, T., Gabrielatos, C. (2006) *English corpus linguistics*. In Aarts, B. & McMahon, A. (Eds.), *The Handbook of English Linguistics*. Oxford: Blackwell.
- [185] McEnery, T., Wilson, A. (2001) *Corpus Linguistics* (second edition). Edinburgh: Edinburgh University Press.
- [186] McEnery, T., Richard, X. and Yukio, T. (2006) *Corpus based language studies, and advanced resource book*. Abington, Oxon: Routledge.
- [187] McEnery, T. and Hardie, A. (2012) *Corpus linguistics. Method, Theory and Practice*. Cambridge University Press.
- [188] McKhann, G.M., Knopman, D.S., Chertkow, H. (2011) The diagnosis of dementia due to Alzheimer’s disease. Recommendations from the National Institute on Aging and the Alzheimer’s Association Workgroups. *Alzheimer’s & Dementia* 2011; 7: 263–269.
- [189] Mena Martinez, F. (2003) Creative modifications of phraseological units in English and Spanish. In Burger, H., A. Häcki Buhofer and G. Gréciano (eds.) *Flut von Texten – Vielfalt der Kulturen. Ascona 2001 zu Methodologie und Kulturspezifik der Phraseologie*, 169–182. Baltmannsweiler: Schneider Verlag Hohengehren.
- [190] Mok, E.H., Lam, L.C., Chiu, H.F. (2004) *Category verbal fluency test performance in Chinese elderly with Alzheimer’s disease*. *Dement Geriatr Cogn Disord* 2004; 18: 120–124.
- [191] Monsch, A.U., Bondi MW, Butters N (1992) *Comparisons of verbal fluency tasks in the detection of dementia of the Alzheimer type*. *Arch Neurol* 1992; 49: 1253–1258.
- [192] Mudraya, O. (2006) Engineering English: a lexical frequency instructional model. *English for Specific Purposes*, 25 (2): 23–56.
- [193] Muangpaisan, W., Intalapaporn, S., Assantachai, P. (2010) *Digit span and verbal fluency tests in patients with mild cognitive impairment and normal subjects in Thai-community*. *J Med Assoc Thai* 2010; 93: 224–230.

- [194] Nattinger, J. and DeCarrico, J. (1992) *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.
- [195] Nelson, J.S., (2006) *Fishes of the World*. 4th ed. Hoboken (New Jersey, USA): John Wiley and Sons.
- [196] Nesselhauf, N. (2005) *Collocations in a Learner Corpus*. Amsterdam/ Philadelphia. John Benjamins Publishing Company, 332.
- [197] Newman, A. (1988) The contrastive analysis of Hebrew and English dress and cooking collocations: Some linguistic and pedagogic parameters. *Applied Linguistics* 9(3): 293–305.
- [198] Nęcka, E., Orzechowski, J., Szymura, B. (2008) *Psychologia poznawcza*. Warszawa: Wydawnictwo Naukowe PWN; 2008.
- [199] Oakes, M. (1998) *Statistics for Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- [200] Opala, G. *Epidemiologia otępień*. (2011) W: Leszek J. (red.). *Choroby otępienie. Teoria i praktyka*. Continuo. Wrocław.
- [201] O’Keeffe, A., McCarthy, M. and Carter, R. (2007) *From Corpus to Classroom: Language use and language teaching*. Cambridge: Cambridge University Press.
- [202] Pakhomov, S.V.S., Eberly, L.E., Knopman, D.S. (2018) Recurrent perseverations on semantic verbal fluency tasks as an early marker of cognitive impairment. *J Clin Exp Neuropsychol*. 2018 ;40(8):832-840.
- [203] Palmer, H. E. (1933) *Second Interim Report on English Collocations*. Tokyo: Kaitakusha.
- [204] Pan, J., Jain, R., Paul, S., et al. (2015) A Internet of Things Framework for smart energy in buildings: Designs, prototype and experiments. *IEEE Internet Things Journal* 2 (6), 527-537.
- [205] Parnowski, T. (2010) *Choroba Alzheimerera*, Warszawa.
- [206] Partington, A., Morley, J. and Haarman, I. (2004) *Corpora and discourse*. Bern: Peter Lang.
- [207] Paschoal, A.M., da Silva, P.H.R., Rondinoni, C., Arrigo, I.V., Paiva, F.F., Leoni, R.F. (2021) Semantic verbal fluency brain network: delineating a physiological basis for the functional hubs using dual-echo ASL and graph theory approach. *J Neural Eng*. 2021;18(4). doi: 10.1088/1741-2552/ac0864. PMID: 34087805.



- [208] Paszenda, J. (2003) A lexicographic corpus of religious phraseological internationalisms in English, German and Polish. Criteria for the selection of data. In Burger, H., A. Häcki Buhofer and G. Gréciano (eds.) *Flut von Texten – Vielfalt der Kulturen. Ascona 2001 zu Methodologie und Kulturspezifität der Phraseologie*, 239–254. Baltmannsweiler: Schneider Verlag Hohengehren.
- [209] Pawley, A., Syder, F. H. (1983) Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (eds). *Language and communication*. 191-225.
- [210] Pearson, J. (1998) *Terms in context*. Amsterdam/ Philadelphia: John Benjamins Publishing Company.
- [211] Peng, B., Yao, X., Risacher, S.L., Saykin, A.J., Shen, L., Ning, X. (2020) ADNI. Cognitive biomarker prioritization in Alzheimer's Disease using brain morphometric data. *BMC Med Inform Decis Mak*. 2020;20(1):319. doi: 10.1186/s12911-020-01339-z. PMID: 33267852; PMCID: PMC7709267.
- [212] Perez ,M., Amayra, I., Lazaro, E., García, M., Martínez, O., Caballero, P., Berrocoso, S., López-Paz, J.F., Al-Rashaida, M., Rodríguez, A.A., Luna, P., Varona, L. (2020) *Intrusion errors during verbal fluency task in amyotrophic lateral sclerosis*. PLoS One. 2020;15(5):e0233349. doi: 10.1371/journal.pone.0233349. PMID: 32469951; PMCID: PMC7259757.
- [213] Phillips, M. (1989) *Lexical Structure of Text [Discourse analysis monographs 12]*. Birmingham: University of Birmingham.
- [214] Piirainen, E. (1995) Niederländische und deutsche Phraseologie: Zu einem kontrastiven Beschreibungsansatz. *Neerlandica Wratislaviensia* 8: 193–217.
- [215] Piñel López, R. M. (2003) Der Tod und das Sterben in der deutschen und spanischen Phraseologie: ein interkultureller Vergleich. In Burger, H., A. Häcki Buhofer and G. Gréciano (eds.) *Flut von Texten – Vielfalt der Kulturen. Ascona 2001 zu Methodologie und Kulturspezifität der Phraseologie*, 229–238. Baltmannsweiler: Schneider Verlag Hohengehren.
- [216] Piskunowicz, M., Bieliński, M., Zgliński, A., Borkowska, A. (2013) Testy fluencji słownej – zastosowanie w diagnostyce neuropsychologicznej. *Psychiatria Polska* 2013, tom XLVII, numer 3.
- [217] Podemski, R., Słotwiński K. (2003) *Zaburzenia mowy w procesach otępiennych*, [w:] *Choroby otępienne. Teoria i praktyka*, red. J. Leszek, Wrocław, s. 271–283.
- [218] Ponichtera-Kasprzykowska, M., Sobów T. (2014) *Adaptacja i wykorzystanie testu fluencji słownej na świecie*. *Psychiatria Psychologia Kliniczna*. 2014, 14 (3), p. 178–187.

- [219] Pope, C., Davis, B. H. (2011) Finding a balance: The carolinas conversation collection. *Corpus Linguistics and Linguistic Theory*, 7(1), 143-161.
- [220] Puchowska-Florek, M., Książkiewicz, B., Nowaczewska, M. (2005) *Ocena przydatności wybranych skal i testów do oceny afazji u pacjentów w ostrym okresie udaru mózgu*. *Udar Mózgu* 2005; 7: 39–47.
- [221] Pulvermüller, F. (2013) How neurons make meaning: brain mechanisms for embodied and abstract- symbolic semantics. *Trends in Cognitive Science*, 17, 9, 458-470.
- [222] Pużyński, S. (2011) Depresja w praktyce ogólnolekarskiej. *Terapia* 2011; 3: 60–64.
- [223] Quirk, R., Greenbaum, S., Leech, G. and Svartvik, J. (1985) *A Comprehensive Grammar of the English Language*. London: Pearson Longman.
- [224] Ralph, M. (2014) *Neurocognitive insights on conceptual knowledge and its breakdown*. *Philosophical Transactions of Royal Society B* 369: 20120392.
- [225] Ratcliff, G., Ganguli, M., Chandra, V. (1998) *Effects of literacy on measures of word fluency*. *Brain Lang* 1998; 61: 115–122.
- [226] Rebuschat, J., Williams, J. N. (Eds.). (2012) *Statistical learning and language acquisition*. Berlin: Mouton de Gruyter.
- [227] Reppen, R., Simpson, R. (2002) Corpus linguistics. In: Schmitt N. (ed.) *An Introduction to Applied Linguistics*. London: Arnold, 92–111.
- [228] Rosińczuk-Tonderys, J., Murzyńska, D., Kazimierska-Zajac, M. (2013) Porównanie fluencji słownej u kobiet ze starzeniem fizjologicznym i chorych z zespołami otępiennymi, *Forum Logopedyczne* 21, 88–93.
- [229] Ross, T. (2003) *The reliability cluster and switch scores for the Controlled Oral Word Association Test*. *Arch. Clin. Neuropsychol.* 2003; 18: 153–164.
- [230] Ross, T.P., Weinberg, M., Furr, A.E., Carter, S.E., Evans-Blake, L., Parham, S. (2005) *The temporal stability of cluster and switch scores using a modified COWAT procedure*. *Arch. Clin. Neuropsychol.* 2005; 20: 983–996.
- [231] Rossell, S. (2006) *Category fluency performance in patients with schizophrenia and bipolar disorder: the influence of affective categories*. *Schizophr. Res.* 2006; 82: 135–138.
- [232] Römer, U. (2005) *Progressives, Patterns, Pedagogy. A corpus-driven approach to English progressive forms, functions, contexts and didactics*. Amsterdam/Philadelphia. John Benjamins Publishing Company.
- [233] Ruff, R.M., Light, R.H., Parker, S.B., Levin, H.S. (1996) *Benton controlled oral word association test: Reliability and updated norms*. *Arch. Clin. Neuropsychol.* 1996; 11: 329–338.

- [234] Ruff, R.M., Light, R.H., Parker, S.B., Levin, H.S. (1997) The Psychological Construct of Word Fluency. *Brain Lang.* 1997; 57: 394–405
- [235] Salazar, D. (2011) A social capital framework for the study of institutional agents and their role in the empowerment of low-status students and youth. *Youth and Society* 43(3). 389-396.
- [236] Salazar, D. (2014) *Lexical Bundles in Native and Non-native Scientific Writing: Applying a corpus-based study to language teaching*. Amsterdam/ Philadelphia: John Benjamins Publishing Company.
- [237] Santos Nogueira, D., Azevedo Reis, E., Vieira, A. (2016) *Verbal Fluency Tasks: Effects of Age, Gender, and Education*. *Folia Phoniatr Logop.* 2016;68(3):124-133.
- [238] Sapilak, B.J., Pokorna-Kaławak, D. (2017) *Narzędzia geriatryczne w praktyce lekarza POZ – test rysowania zegara*. *Lekarz POZ* 6/2017: 387-391.
- [239] Schmitt, N. (ed.) (2004) *Formulaic Sequences*. Amsterdam/ Philadelphia: John Benjamins Publishing Company.
- [240] Schnur, E. (2014) Phraseological signaling of discourse organization in academic lectures: A comparison of lexical bundles in authentic lectures and EAP listening materials. *Yearbook of Phraseology*. Volume 5: Issue 1.
- [241] Schwartz, S., Baldo, J. (2001) *Distinct patterns of word retrieval in right and left frontal lobe patients: a multidimensional perspective*. *Neuropsychologia* 2001; 39: 1209–1217.
- [242] Schwartz, M. A., Ridley, A.J., Burrige, K., Firtel, A.F., Ginsberg, M. H., Borisy, G., Parsons, J. T., Horwitz, A. R. (2003) Cell migration: integrating signals from front to back. *Science*. 2003 Dec 5;302(5651):1704-9.
- [243] Scott, M., Tribble, Ch. (2006) *Textual Patterns. Key words and corpus analysis in language education*. Amsterdam/ Philadelphia. John Benjamins Publishing Company.
- [244] Scott, M. (2015) *WordSmith Tools 6.0*. Liverpool: Lexical Analysis Software.
- [245] Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- [246] Shakarami, A., Caltablano, N.J., Hajhashemi, K. (2016) Digital Discourse Markers in an ESL Learning Setting: The Case of Socialisation Forums. *International Journal of Instruction*.
- [247] Simpson-Vlach, R., Ellis. N.C. (2010) An academic formulas list: New methods in phraseology research. *Applied Linguistics* 31, 487-512.
- [248] Sinclair, J.McH. (1987) *Collins Cobuild English Language Dictionary*. HyperCollins.

- [249] Sinclair, J.McH. (1991) *Corpus, Concordance, Collocation*. Oxford: OUP.
- [250] Sinclair, J.McH. (2004) *Trust the Text: Language corpus and discourse*. London: Routledge.
- [251] Sinclair, J. McH., Mauranen, A. (2005) *Degenrate data*. Workshop held at AAACL 6 and ICAME 26, University of Michigan.
- [252] Sinclair, J. McH., Mauranen, A. (2006) *Linear Unit Grammar*. Amsterdam/ Philadelphia: John Benjamins Publishing Company.
- [253] Siyanova, A., Schmitt, N. (2008) L2 learner production and processing of collocation: A multi-study perspective. *Canadian Modern Language Review*, 64, 429–458.
- [254] Skorupka, S. (1967) *Słownik frazeologiczny języka polskiego*, t. 1-2, Warszawa.
- [255] Smadja, F. McKeown, K. and Hatzivassiloglou, V. (1996) Translating collocations for bilingual lexicons: a statistical approach. *Computational Linguistics* 22. 1-38.
- [256] Sokołowski, A., Tyburski, E., Sotys, A., Karabanowicz, E. (2020) *Sex Differences in Verbal Fluency Among Young Adults*. *Adv Cogn Psychol*. 2020;16(2):92-102.
- [257] Sosa, A.V., MacFarlane, J. (2002) Evidence for frequency-based constituents in the mental lexicon: Collocations involving the word of. *Brain and Language* 83(2):227-36.
- [258] Stephan, Y., Sutin, A.R., Luchetti, M., Aschwanden, D., Terracciano, A. (2021) *Subjective age and verbal fluency among middle aged and older adults: A meta-analysis of five cohorts*. *Arch Gerontol Geriatr*. 2021 Nov-Dec;97:104527. doi: 10.1016/j.archger.2021.104527. Epub 2021 Sep 11. PMID: 34534857.
- [259] Stolarska, U., Krocza, S., Geront, A. (2008) *Test fluencji słownej – aspekty rozwojowe w normie i patologii*. *Przegląd Lekarski* 65: 764–768.
- [260] Strauss, E., Sherman, E.M.S., Spreen, O. (2006) *A Compendium of neuropsychological tests: Administration, norms, and commentary*. New York: Oxford University Press.
- [261] Stubbs, M. (1996) *Text and Corpus Analysis*. Oxford: Blackwell.
- [262] Stubbs, M. (2004) Language Corpora. In A Davies & C Elder (eds.) *Handbook of Applied Linguistics*. Oxford: Blackwell. 106-32.
- [263] Stubbs, M. (2010) *Keyness in Texts*. Amsterdam/ Philadelphia: John Benjamins Publishing Company.

- [264] Stuss, D.T., Alexander, M.P., Hamer, L., Palumbo, C., Dempster, R., Binns, M., Levine, B., Izukawa, D. (1998) *The effects of focal anterior and posterior brain lesions on verbal fluency*. J. Int. Neuropsychol. Soc. 1998; 4: 265–278.
- [265] Sumerall, S.W., Timmons, P.L., James, A.L., Ewing, M.J., Oehlert, M.E. (1997) *Expanded norms for the Controlled Oral Word Association Test*. J. Clin. Psychol., 53: 517–521.
- [266] Sutin, A.R., Stephan, Y., Terracciano, A. (2019) Verbal fluency and risk of dementia. *Int J Geriatr Psychiatry*. 2019;34(6):863-867.
- [267] Szatkowska, I., Grabowska, A., Szymańska, O. (2000) *Phonological and semantic fluencias are mediated by different regions of the prefrontal cortex*. *Acta Neurobiol Exp (Wars)* 2000; 60: 503–508.
- [268] Szepietowska, E.M. (2010) *Płynność słowna w neuropsychologii – aspekty teoretyczne i metodologiczne*. W: Jodzio K, Szepietowska EM. red. *Neuronalne ścieżki poznania i zachowania. Rozważania interdyscyplinarne*. Lublin: Wyd. UMCS; 2010. s. 15–28.
- [269] Szepietowska, E.M., Gawda, B. (2011) *Ścieżkami fluencji werbalnej*, Lublin.
- [270] Szepietowska, E.M., Lipian, J. (2012) *Fluencja słowna neutralna i afektywna u chorych z uszkodzeniem prawej, lewej lub obu półku mózgu*. *Psychiatr Pol* 2012; 46: 539–551.
- [271] Świtalska, J. (2012) *Nasilenie i rodzaj objawów a funkcjonowanie poznawcze w chorobie afektywnej dwubiegunowej*. *Postępy Psychiatr Neurol* 2012; 21: 37–42.
- [272] Tabert, M.H., Peery, S., Borod, J., Schmidt, M., Grunwald, I., Sliwinski, M. (2001) *Lexical emotional expression across the life span: quantitative and qualitative analyses of word list generation tasks*. *J. Clin. Neuropsychol.* 2001; 15; 4: 531–550.
- [273] Talarowska, M., Florkowski, A., Gałęcki, P. (2008) *Zaburzenia funkcji poznawczych wśród chorych na cukrzycę*. *Pol Merkuriusz Lek* 2008; 15: 349–355.
- [274] Talarowska, M., Zboralski, K., Gałęcki, P. (2011) *Results of the verbal fluency test among patients suffering from depressive disorders and organic depressive disorders*. *Curr Probl Psychiatry* 2011; 12: 397–403.
- [275] Thurstone, L.L. (1938) *Primary mental abilities*. Chicago: University of Chicago Press.
- [276] Tłokiński, W., Olszewski, H. (2014) Zaburzenia mowy związane z wiekiem, W: *Biomedyczne podstawy logopedii*, red. S. Milewski, J. Kuczkowski, K. Kaczorowska-Bray, Gdańsk, s. 410–428.
- [277] Tognini-Bonelli, E. (2001) *Corpus Linguistics at Work*. Amsterdam/ Philadelphia: John Benjamins Publishing Company.

- [278] Tombaugh, T.N., Kozak, J., Rees, L. (1999) *Normative data stratified by age and education for two measures of verbal fluency: FAS and Animal Naming*. Arch. Clin. Neuropsychol. 1999; 14: 167–177.
- [279] Troyer, A.K., Moscovitch, M., Winocur, G. (1997) Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology* 1997; 11: 138–146.
- [280] Troyer, A.K., Moscovitch, M., Winocur, G., Alexander, M.P., Stuss, D. (1998) *Clustering and switching on verbal fluency: the effects of focal frontal- and temporal-lobe lesions*. Neuropsychologia 1998; 36: 499–504.
- [281] Ueda, Y. (2004) Kontrastive Phraseologie – idiomatische Wendungen mit Tierbezeichnungen als Hauptkomponenten im Deutschen und Japanischen. In Földes, C. (ed.) *Res humanae proverbiorum et sententiarum, Ad honorem Wolfgangi Mieder*, Tübingen: Gunter Narr Verlag, 351–364.
- [282] Valentin, P. (1999) Les lexèmes nominaux composés: Équivalences et traductions entre allemand et français. *Nouveaux Cahiers d'Allemand* 17: 533–540.
- [283] Van der Elst, W., Van Boxtel, M.P., Van Breukelen, G.J. (2006) *Normative data for the Animal, Profession and Letter M Naming verbal fluency tests for Dutch speaking participants and the effects of age, education, and sex*. J Int Neuropsychol Soc 2006; 12: 80–89.
- [284] Veisbergs, A. (1992) English-Latvian borrowing of idioms. *Contrastive and Applied Linguistics* (University of Latvia) 1: 36–48.
- [285] Vinogradov, V. (1946) Basic concepts of phraseology as a linguistic discipline. Leningrad University.
- [286] Vliet, E.C., Manly, J., Tang, M.X. (2003) *The neuropsychological profiles of mild Alzheimer's disease and questionable dementia as compared to age-related cognitive decline*. J Int Neuropsychol Soc 2003; 9: 720–732.
- [287] Waszkiewicz, J., Wciórka, J., Anczewska, M. (2012) *Zaburzenia językowe a inne wybrane funkcje poznawcze u osób chorujących na zaburzenia schizofreniczne*. Psychiatr Pol 2012; 46: 553–570.
- [288] Weiss, E.M., Ragland, J.D., Bressinger, C.M., Bilker, W.B., Deisenhammer, E.A., Delazer, M. (2006) *Sex differences in clustering and switching in verbal fluency tasks*. J. Int. Neuropsychol. Soc. 2006; 12: 502–509.
- [289] Williams, G. (1999) *Looking in before looking out: Internal selection criteria in a corpus of plant biology*. Papers in Computational Lexicography. Complex '99. Hungary: Budapest. : 195-204.

- [290] Williams, G. (2002) Corpus-driven lexicography and the specialised dictionary: headword extraction for the Parasitic Plant Research Dictionary. In Braasch A. and C. Povlsen (eds.), 859-864.
- [291] Williams, G. (2003) From meaning to words and back: Corpus linguistics and specialized lexicography. *ASp* 31-33, 63-76.
- [292] Wood, A., Saling, M., Abbott, D.F., Jackson, G.D. (2001) *A neurocognitive account of frontal lobe involvement in orthographic lexical retrieval: an fMRI study*. *Neuroimage* 2001; 14: 162–169.
- [293] Wray, A., Perkins, M. (2000) The functions of formulaic language: An integrated model. *Language and Communication* 20(1):1-28.
- [294] Wray, A. (2002) *Formulaic Language and the Lexicon*. New York: Cambridge University Press.
- [295] Wray, A., Fitzpatrick, T. (2008) Why can't you just leave it alone? Deviations from memorized language as a gauge of nativelike competence. In F. Meunier and S. Granger (Eds.), *Phraseology in foreign language learning and teaching*. Amsterdam/ Philadelphia: John Benjamins Publishing Company, 123-147.
- [296] Wysokiński, A., Zboralski, K., Orzechowska, A. (2010) *Normalization of the Verbal Fluency Test on the basis of results for healthy subjects, patients with schizophrenia, patients with organic lesions of the chronic nervous system and patients with type 1 and 2 diabetes*. *Arch Med Sci* 2010; 6: 438–446.
- [297] Zaharia, C. (2003) Quellen der idiomatischen Ausdrücke im Deutschen und Rumänischen. W: Burger, H., A. Häcki Buhofer & G. Gréciano (eds.) *Flut von Texten – Vielfalt der Kulturen. Ascona 2001 zu Methodologie und Kulturspezifik der Phraseologie*, 267–276. Baltmannsweiler: Schneider Verlag Hohengehren.
- [298] Zawadzka, E. (2013) *Świat w obrazach u osób po udarze mózgu*, Warszawa.
- [299] Zubenko, G.S., Zubenko, W.N., McPherson, S., Spoor, E., Marin, D.B., Farlow, M.R., Smith, G.E., Geda, Y.E., Cummings, J.L., Petersen, R.C., Sunderland, T. (2003) *A collaborative study of the emergence and clinical features of the major depressive syndrome of Alzheimer's disease*. *Am. J. Psychiatry*. 2003, 160: 857-866.

## ŹRÓDŁA INTERNETOWE

Narodowy Korpus Języka Polskiego

<http://www.nkjp.pl> (dostępność: 26.09.2021r)

## ROZDZIAŁ XIII

### ZAŁĄCZNIKI

#### 13.1. Załącznik A- Spis rycin

Nr ryciny	Nazwa ryciny	Strona
1	Projekcja przestrzenna czynników głównych testu Fluencji Słownej i MMSE	92
2	Wiek i wyniki testu MMSE badanych osób. Porównanie wyników kobiet i mężczyzn. Wartości średnie, SD, min-max.	93
3	Rozkład wartości obserwowanych względem oczekiwanych wyników testu MMSE w badanej grupie	94
4	Rozkład wyników testów fluencji słownej w badanej grupie. Wartości średnie, SD, min-max	95
5	Słowa nieprawidłowe włączone w poszczególnych testach fluencji słownej	96
6	Rozkład reszt przewidywanych względem oczekiwanych w teście fluencji słownej	100
7	Rozkład wartości obserwowanych względem oczekiwanych w teście fluencji słownej w zależności od czasu trwania zaburzeń poznawczych	106
8	Trzy najczęściej występujące słowa w wykonaniu testu fluencji przez badanych w każdej kategorii	112
9	Trzy najczęściej występujące słowa w wykonaniu testu fluencji w każdej kategorii u kobiet	113
10	Trzy najczęściej występujące słowa w wykonaniu testu fluencji w każdej kategorii u mężczyzn	113



### 13.2. Załącznik A- Spis tabel

Nr tabeli	Nazwa tabeli	Strona
I	Strukturalna klasyfikacja wiązek leksykalnych w prozie akademickiej zaadaptowana przez Bibera [27]	66
II	Klasyfikacja funkcjonalna wiązek leksykalnych [30]	68
III	Klasyfikacja funkcjonalna wiązek leksykalnych w piśmiennictwie akademickim [127]	69
IV	Interkorelacje podskal testu Fluencji Słownej	91
V	Analiza czynnikowa składowych głównych testu fluencji słownej	91
VI	Wiek i wyniki testu MMSE badanych osób. Porównanie wyników kobiet i mężczyzn.	93
VII	Wyniki testów fluencji słownej. Porównanie wyników kobiet i mężczyzn	94
VIII	Frekwencja nieprawidłowo włączonych słów w poszczególnych testach fluencji słownej	96
IX	Korelacje rang Spearmana pomiędzy wiekiem i czasem trwania zaburzeń poznawczych a wykonaniem testu MMSE i testów fluencji słownej	97
X	Korelacje rang Spearmana pomiędzy wynikami testu MMSE i testów fluencji słownej	98
XI	Wyniki analizy jednoczynnikowej ANOVA – przewidywanie wykonania testu fluencji fonologicznej (test literowy) na podstawie wyniku MMSE w badanej grupie	99
XII	Wyniki analizy jednoczynnikowej ANOVA – przewidywanie wykonania testu fluencji kategorialnej (zwierzęta) na podstawie wyniku MMSE w badanej grupie	99
XIII	Wyniki analizy jednoczynnikowej ANOVA – przewidywanie wykonania testu fluencji kategorialnej (przedmioty ostre) na podstawie wyniku MMSE w badanej grupie	99
XIV	Zależność pomiędzy wynikami testu MMSE a sprawnością wykonania	102

	testu fluencji fonologicznej – test literowy w badanej grupie	
XV	Zależność pomiędzy wynikami testu MMSE a sprawnością wykonania testu fluencji kategorialnej – zwierzęta w badanej grupie	103
XVI	Zależność pomiędzy wynikami testu MMSE a sprawnością wykonania testu fluencji kategorialnej – przedmioty ostre- w badanej grupie	104
XVII	Wyniki analizy jednoczynnikowej ANOVA – przewidywanie wykonania testu fluencji fonologicznej – test literowy na podstawie czasu trwania zaburzeń poznawczych.	105
XVIII	Wyniki analizy jednoczynnikowej ANOVA – przewidywanie wykonania testu fluencji kategorialnej (zwierzęta) na podstawie czasu trwania zaburzeń poznawczych	105
XIX	Wyniki analizy jednoczynnikowej ANOVA – przewidywanie wykonania testu fluencji kategorialnej (przedmioty ostre) na podstawie czasu trwania zaburzeń poznawczych	105
XX	Czas trwania zaburzeń poznawczych a wyniki testu fluencji słownej fonologicznej w badanej grupie	108
XXI	Czas trwania zaburzeń poznawczych a wyniki testu fluencji kategorialnej- zwierzęta	109
XXII	Czas trwania zaburzeń poznawczych a wyniki testu fluencji kategorialnej- przedmioty ostre	110
XXIII	Procentowy udział poszczególnych liter w najczęściej wymienianych słowach w NKJP w podkorpusie IPI PAN [51]	115