

Paweł Rutkowski, Małgorzata Czajkowska-Kisil,  
Joanna Łacheta, Anna Kuder

## Analiza szyku przymiotnika jako przykład wykorzystania danych korpusowych w badaniach nad gramatyką PJM

### 1. Wstęp

Dane korpusowe otwierają przed lingwistami badającymi języki migowe zupełnie nowe możliwości analityczne. O ile jeszcze do niedawna większość prac poświęconych gramatyce wizualno-przestrzennej bazowała przede wszystkim na intuicjach pojedynczych głuchych użytkowników, o tyle w ostatnich latach coraz częściej generalizacje przedstawiane w literaturze przedmiotu są weryfikowalne, gdyż wywodzą się z przeglądu danych empirycznych. Przykładem zagadnienia, które analizować można w sposób o wiele pełniejszy dzięki odwołaniu do materiałów korpusowych, jest szyk wyrazów w zdaniu migowym. Niniejszy rozdział przedstawia swoiste studium przypadku w tym zakresie, a mianowicie wstępny opis szyku składników w konstrukcjach imiennych polskiego języka migowego (PJM). Na potrzeby referowanego tu badania poddaliśmy analizie kilkugodzinną próbkę materiału wideo zaczerpniętą z tworzonych przez Pracownię Lingwistyki Migowej Uniwersytetu Warszawskiego korpusu PJM. Uzyskane przez nas wyniki pokazują, jak ważną rolę w pracach naukowych poświęconych komunikacji migowej odgrywać może oparcie ich na rzeczywistych danych liczbowych dotyczących uzusu.\*

---

\* Niniejszy rozdział zdaje sprawę z prac, które prowadzone są w ramach projektu *Ikoniczność w gramatyce i leksyce polskiego języka migowego (PJM)*, sfinansowanego ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC-2011/01/M/HS2/03661. Referowane tu wyniki zostały wcześniej zaprezentowane na konferencjach *Theoretical Issues in Sign Language Research – TISLR 11* (lipiec 2013 r., University College London) oraz *44th Poznań Linguistic Meeting – PLM* (sierpień 2013 r., Uniwersytet im. Adama Mickiewicza w Poznaniu). Jesteśmy wdzięczni uczestnikom tych spotkań za pomocne uwagi i sugestie.

## 2. Szyk konstrukcji nominalnych w językach migowych

Ze względu na trójwymiarowość przestrzeni migania oraz możliwość symultanicznej artykulacji kilku sygnałów wizualnych zagadnienie szyku (rozumianego jako konkretna linearyzacja pewnego ciągu znaków) nie należy do centralnych problemów lingwistyki migowej. Wszystkie opisane w literaturze przedmiotu języki Głuchych charakteryzują się dosyć luźnym szykiem – od reguł szeregowania znaków o wiele istotniejszy wpływ na kształt komunikatu wizualnego wydają się mieć reguły dotyczące pozycjonowania znaków w przestrzeni migania (por. dwa rozdziały zamieszczone w niniejszym tomie: *Ikoniczność nieleksykalna: reprezentacja referencjalna jako składnik tekstu w językach wizualno-przestrzennych* autorstwa Marka Świdzińskiego i Pawła Rutkowskiego oraz *Czasownik i przestrzeń* autorstwa Joanny Filipczak). Choć zakres zagadnień opisywanych w literaturze dotyczącej składni języków wizualno-przestrzennych stale się zwiększa (o czym świadczyć może chociażby szeroki przegląd, który oferuje kompendium pod redakcją Pfaua, Steinbacha i Woll, 2012), struktura syntaktyczna grup imiennych pozostaje obszarem słabo rozpoznany. Większość prac dotyczących tej problematyki wpisuje się w tradycję generatywizmu (np. MacLaughlin, 1997, Zhang, 2007, Bertone, 2010, Neidle and Nash, 2012), czyli wykorzystuje w opisie konstrukcji nominalnych model tzw. frazy przedimkowej (ang. *Determiner Phrase*, DP) zaproponowany przez Abneya (1987). Charakterystyczna dla tego podejścia teoretycznego ścisła hierarchizacja fraz funkcyjnych współtworzących szkielet konstrukcyjny grupy imiennej powinna przekładać się na szyk poszczególnych przydawek dołączanych do rzeczownika głównego (por. np. Rutkowski, 2009).

Aby sprawdzić, w jakiej mierze model DP znajduje potwierdzenie w faktach z PJM, czyli języka o — wydawać by się mogło — bardzo dużej swobodzie szyku, przeanalizowaliśmy skład wszystkich wyrażen nominalnych występujących w kilkugodzinnej próbie danych wyekstrahowanych z korpusu PJM. W próbie tej wyróżniliśmy w sumie 4827 znaków. Na materiał wideo poddany przez nas oglądowi składały się wypowiedzi dziesięciu głuchych użytkowników PJM opisujące historyjki obrazkowe o Myszcze Miki oraz film *Pear story* (Chafe, 1980, por. rozdział *Korpus polskiego języka migowego — zarys projektu* autorstwa Piotra Mostowskiego w niniejszym tomie). Naszym celem było wykorzystanie szczegółowo zanalizowanych danych empirycznych jako podstawy do udzielenia odpowiedzi na następujący zestaw pytań badawczych:

1. Czy linearyzacją grupy imiennej rządzą w PJM dające się precyzyjnie sformułować reguły składniowe?

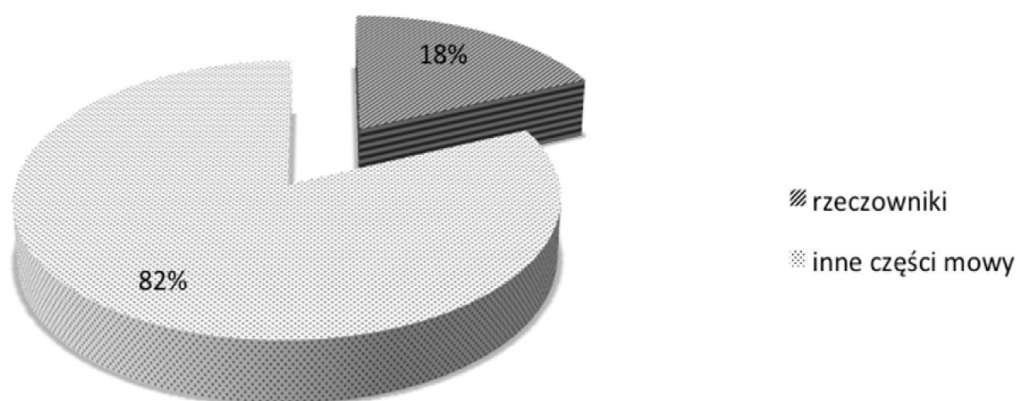
2. Jaki wpływ na składnię nominalną PJM można przypisać polszczyźnie?
3. Jaki poziom złożoności składniowej można zaobserwować w konstrukcjach nominalnych występujących w tekstach PJM?

Uzyskane przez nas wyniki zrelacjonowane są w dalszej części tego rozdziału.

### 3. Szyk przymiotnika w PJM: dane korpusowe

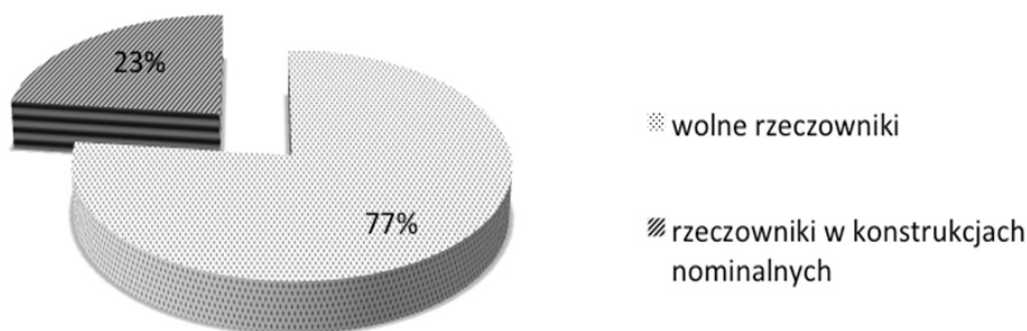
Analiza materiału empirycznego wyekstrahowanego z korpusu na potrzeby opisywanego tu projektu wymagała sporządzenia szczegółowej anotacji konstrukcji zawierających rzeczowniki. W tym celu konieczne było otagowanie wszystkich znaków składających się na badaną próbkę w zakresie ich statusu kategoryjnego (przynależności do części mowy, por. rozdział *Anotacja korpusu PJM* autorstwa Joanny Filipczak w niniejszym tomie). Rozróżniliśmy odpowiednimi znacznikami następujące składniki konstrukcji nominalnych: rzeczowniki, przymiotniki, liczebniki, elementy dzierżawcze i indeksy (elementy demonstratywne i referencjalne). Tagowanie to musiało być oczywiście przeprowadzone ręcznie, jako że obecny stopień rozwoju narzędzi informatycznych służących rozpoznawaniu obrazu nie pozwala na (pół)automatyczne segmentowanie nagrania wideo na poszczególne znaki migowe. Proces anotacji tekstu migowego różni się pod tym względem w istotny sposób od procesu anotacji tekstu pisanego (ten może być — przynajmniej w pewnym zakresie — znakowany bez zaangażowania pracy ludzkiej).

Otagowanie interesujących nas danych wykazało, iż wśród 4827 znaków składających się na analizowaną próbkę materiału korpusowego znalazło się 887 rzeczowników. Wykres 1 pokazuje odpowiednie relacje procentowe.



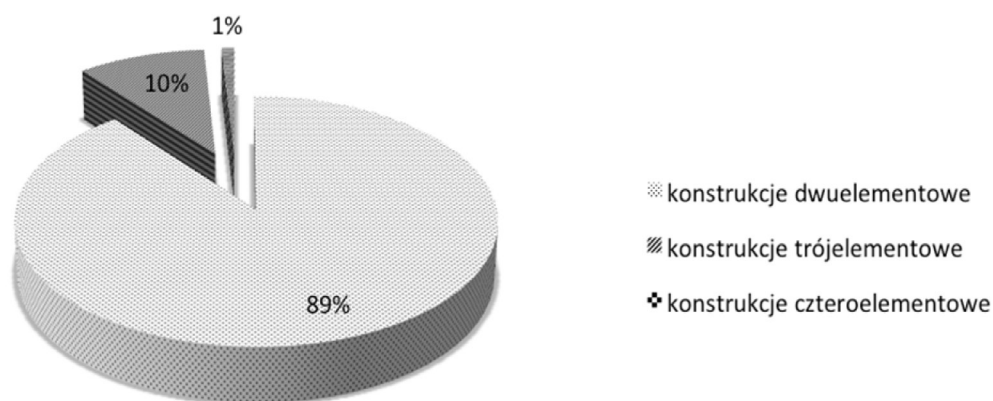
Wykres 1. Odsetek rzeczowników wśród wszystkich znaków w zanalizowanym materiale

Szczegółowe znakowanie kategoryjne pozwoliło również na określenie, jak wiele spośród 887 rzeczowników było składnikami bardziej złożonych konstrukcji imiennych. Okazało się, iż do zaledwie 205 rzeczowników w badanej próbie dołączona była jakakolwiek przydawka. Oznacza to, że 682 wyrażenia nominalne realizowane były w postaci pojedynczego rzeczownika (por. wykres 2).



Wykres 2. Odsetek wieloelementowych wyrażen nominalnych

Kolejny etap naszych prac dotyczył złożoności wyrażen imiennych w tekście PJM. Spośród 205 co najmniej dwuelementowych konstrukcji nominalnych, które odnotowaliśmy w badanej próbie, 183 składało się z dwóch elementów (np. rzeczownika i określającego go przymiotnika lub rzeczownika i kwantyfikującego go liczebnika), 20 miało trzy składniki (np. rzeczownik główny oraz przydawkę przymiotną i przydawkę rzeczowną), a jedynie dwie konstrukcje można było uznać za czteroelementowe (rzeczownik, liczebnik, dwa przymiotniki). Wykres 3 ilustruje proporcje między omawianymi tu typami konstrukcji.

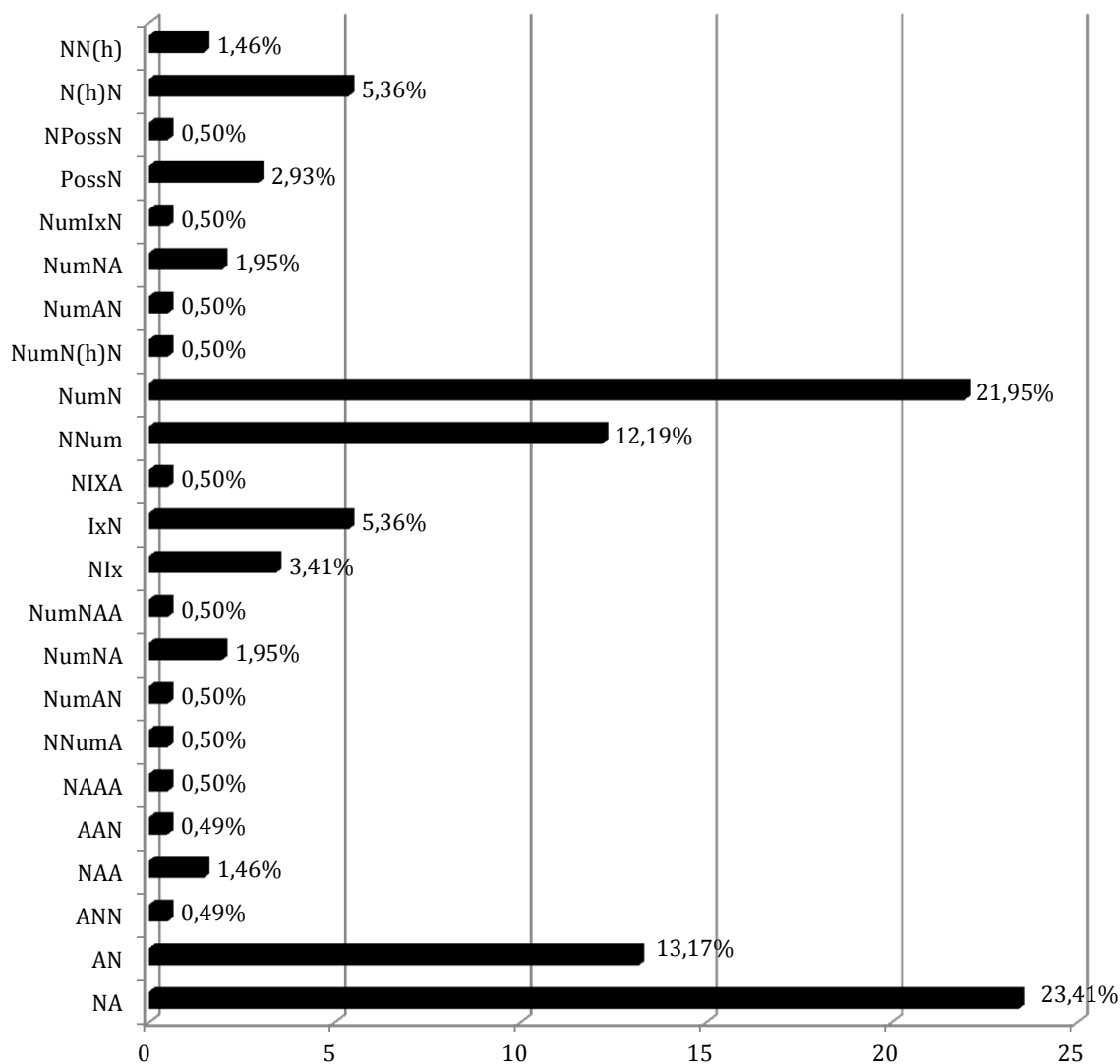


Wykres 3. Stopień złożoności konstrukcji nominalnych

Powyższe dane stanowią cenne źródło wiedzy o składni nominalnej PJM. Zwraca uwagę przede wszystkim to, iż w tekstach wizualno-przestrzennych do rzadkości zdają się należeć rozbudowane konstrukcje imienne. Być

może wynika to z samego charakteru danych migowych rejestrowanych w korpusie tworzonym przez Pracownię Lingwistyki Migowej UW. Są to w olbrzymiej większości teksty o charakterze konwersacyjnym — powstające w kontekście rozmowy przeprowadzanej twarzą w twarz między dwoma głuchymi użytkownikami PJM. Oznacza to, że ich odpowiednikiem w komunikacji fonicznej byłby język mówiony, a nie pisany. Przypuszczamy, że frekwencyjna analiza transkrypcji tekstu mówionego wykazałaby, że wieloelementowe konstrukcje nominalne występują w nim relatywnie rzadko — być może równie rzadko, co w tekście migowym. Hipotezę tę planujemy w dalszych badaniach zweryfikować za pomocą danych empirycznych dotyczących języka mówionego.

Najważniejszym aspektem relacjonowanego w tym rozdziale badania była próba opisu szyku konstrukcji nominalnych występujących w PJM. Wykres 4 pokazuje, jakie typy struktur imiennych odnotowaliśmy

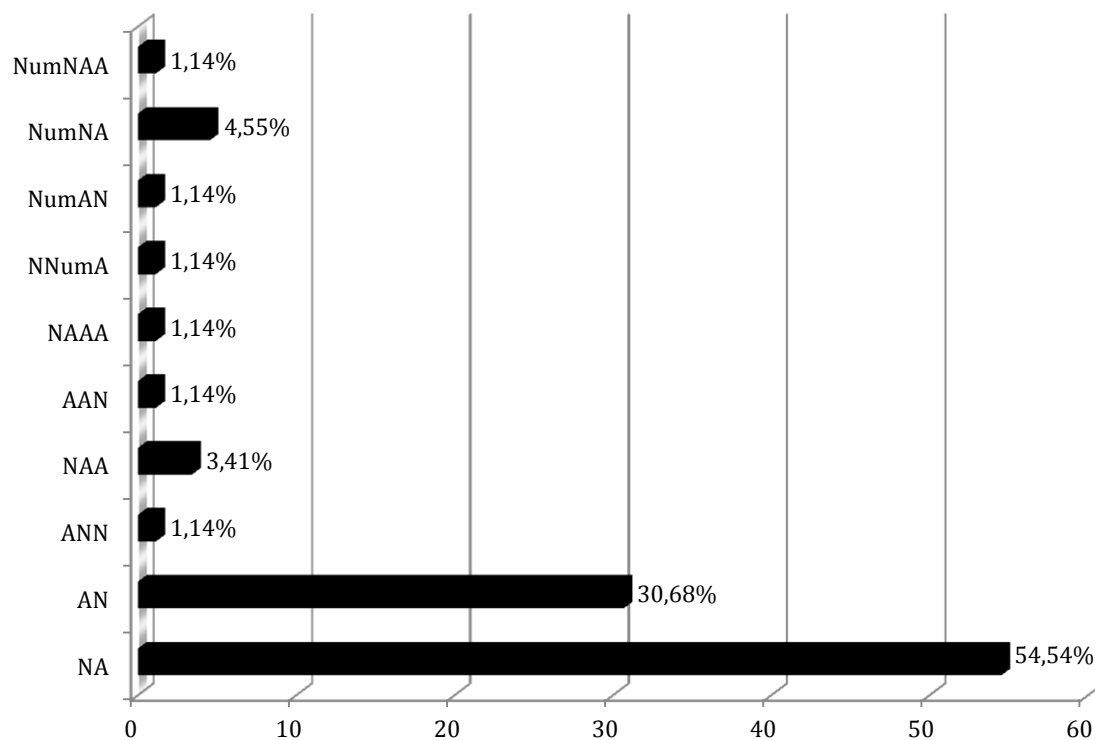


Wykres 4. Typy konstrukcji nominalnych w zanalizowanym materiale

w przebadanej próbkę. Poszczególne człony konstrukcji są na nim oznaczone jako: N = rzeczownik (ang. *noun*), A = przymiotnik (ang. *adjective*), Num = liczebnik (ang. *numeral*), Ix = indeks/element wskazujący (ang. *indexical*), Poss = element dzierżawczy (ang. *possessive*), h (po innym symbolu) = nadrzędnik (ang. *head*).

Z wykresu 4 wynika, że najczęstszym modyfikatorem rzeczownika są w tekście migowym przymiotniki. Łącznie odnotowaliśmy 87 konstrukcji przymiotnikowych o różnym stopniu złożoności. Wykres 5 przedstawia poszczególne typy wyrażeń, w których skład weszły przydawki przymiotne.

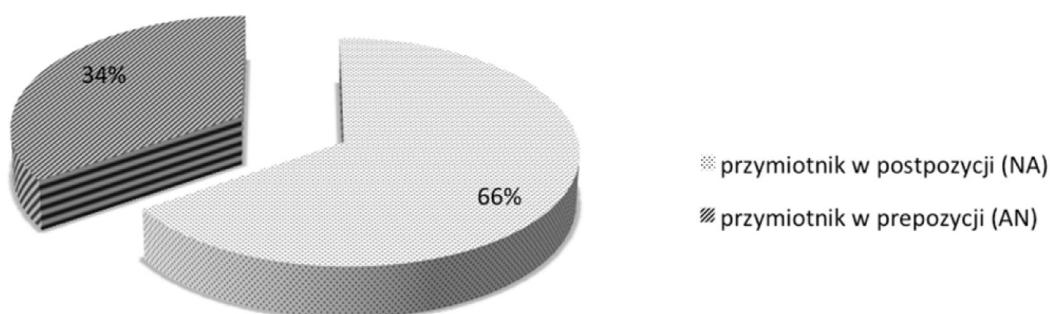
Dane, których podsumowanie stanowi wykres 5, wyraźnie wskazują na bardzo interesującą tendencję w składni PJM: w przebadanych przez nas tekstach większość przymiotników występuje w postpozycji (a nie — tak jak w polszczyźnie — w prepozycji) względem rzeczownika głównego. W 57 frazach przydawka przymiotna wyartykułowana została po swoim nominalnym nadrzędniku, a jedynie w 30 frazach szyk PJM odpowiadał szykowi polskiemu (wykres 6 oddaje te dane w postaci relacji procentowej).



Wykres 5. Konstrukcje nominalne zawierające przymiotniki

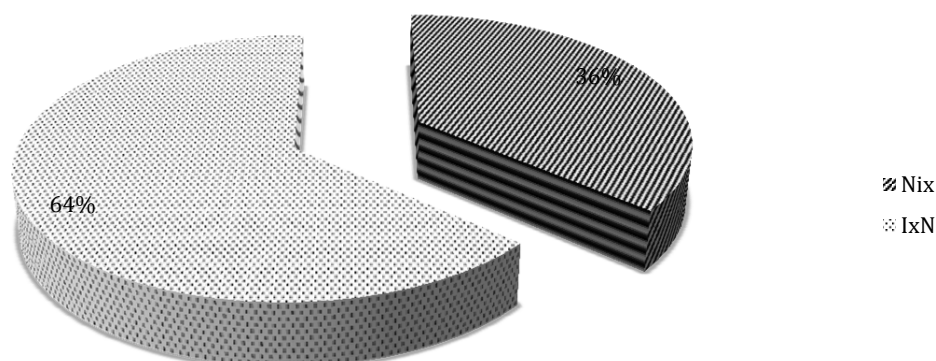
Z powyższych faktów w jasny sposób wynika, że w skład gramatyki PJM muszą (mimo jej trójwymiarowego charakteru) wchodzić reguły dotyczące szyku wyrazów. Reguły te nie mogą być przy tym uznane za zapożyczenie z linearnego języka fonicznego (jakim jest otaczająca PJM polszczyzna), gdyż

szyk NA został najwyraźniej zgramatyzalizowany na gruncie PJM niezależnie od dominującego w polszczyźnie szyku AN.

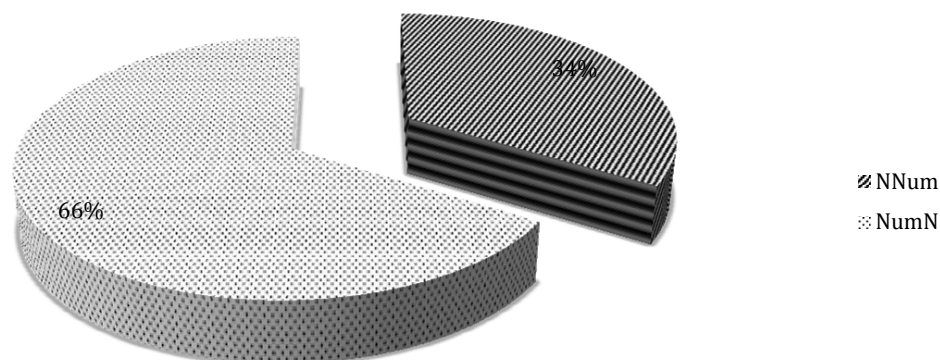


Wykres 6. Konstrukcje z przymiotnikami w prepozycji i postpozycji względem rzeczownika

Co ciekawe, postpozycyjne umiejscowienie przydawki wydaje się dotyczyć wyłącznie przymiotników, a nie innych modyfikatorów rzeczownika głównego. Elementy dzierżawcze, wskazujące i liczebne wykazują tendencję do poprzedzania rzeczownika, którego są przydawkami. Wykresy 7 i 8 przedstawiają dane procentowe dotyczące indeksów i liczebników. W wypadku składników o interpretacji dzierżawczej obraz jest jeszcze klarowniejszy: wszystkie odnotowane wystąpienia tego typu elementów wiązały się z szykiem prepozycyjnym.



Wykres 7. Konstrukcje z indeksami w prepozycji i postpozycji względem rzeczownika



Wykres 8. Konstrukcje z liczebnikami w prepozycji i postpozycji względem rzeczownika

Podstawowy szyk konstrukcji nominalnych w PJM można zatem oddać za pomocą następującego schematu (symbol „>” odpowiada stwierdzeniu „zazwyczaj występuje przed”):

INDEKS/ELEMENT DZIERŻAWCZY/LICZEBNIK > RZECZOWNIK > PRZYMIOTNIK

## 4. Podsumowanie

Przedstawiając wyłaniające się z przestudiowanych danych wstępne generalizacje dotyczące szyku wyrażeń nominalnych, niniejszy rozdział pokazać miał, jak istotną rolę w badaniach nad językami migowymi mogą odegrać dane korpusowe. Na podstawie analizy frekwencyjnej wykazaliśmy, że przymiotniki w PJM mają tendencję do występowania w postpozycji względem rzeczownika głównego, podczas gdy inne modyfikatory (takie jak liczebniki, elementy dzierżawcze i indeksy) zazwyczaj poprzedzają rzeczownik. Tak wyraźny kontrast nie jest z pewnością dziełem przypadku, a zatem nie może być zignorowany przez badacza. W odróżnieniu od intuicji gramatycznych poszczególnych użytkowników PJM (które, o czym warto wspomnieć, dalekie są od jedności w kwestii szyku NA/AN jako podstawowej i nienacechowanej linearyzacji grupy imiennej) danym korpusowym trudno zarzucić niereprezentatywność. Naszym zdaniem dostarczają one istotnych argumentów za następującymi stwierdzeniami:

1. Linearyzacją grupy imiennej rządzą w PJM zgramatyzalizowane reguły składniowe.
2. Reguły te nie są prostym zapożyczeniem linearyzacji charakterystycznych dla polszczyzny.
3. Większość konstrukcji nominalnych występujących w tekstach PJM cechuje znikoma złożoność składniowa.

## Bibliografia

- Abney, Steven P. (1987). *The English noun phrase in its sentential aspect*. Rozprawa doktorska. Cambridge, Mass.: MIT.
- Bertone, Carmela (2010). The syntax of noun modification in Italian Sign Language (LIS). W L. Brugè (red.), *University of Venice Working Papers in Linguistics*, ss. 7–28. Venezia: Dipartimento di Scienze del Linguaggio, Università Ca' Foscari.



- Chafe Wallace, red. (1980). *The Pear stories: cognitive, cultural, and linguistic aspects of narrative production*. Norwood: Ablex.
- MacLaughlin, Dawn (1997). *The structure of determiner phrases: Evidence from American Sign Language*. Rozprawa doktorska. Boston: Boston University.
- Neidle, Carol, Nash, Joan (2012). The noun phrase. W R. Pfau, M. Steinbach, B. Woll (red.), *Sign language: An international handbook*, ss. 265–292. Berlin: De Gruyter Mouton.
- Pfau, Roland, Steinbach, Markus, Woll, Bencie, red. (2012). *Sign language: An international handbook*. Berlin: De Gruyter Mouton.
- Rutkowski, Paweł (2009). *Fraza przedimkowa w polszczyźnie*. Warszawa: Wydział Polonistyki Uniwersytetu Warszawskiego.
- Zhang, Niina N. (2007). Universal 20 and Taiwan Sign Language. *Sign Language & Linguistics* 10, ss. 55–81.