

Urszula Modrzyk

Omówienie pracy magisterskiej

Można powiedzieć, że punktem wyjścia mojej pracy magisterskiej była pewna niekompatybilność.

Pracowałam w projekcie CLARIN-PL, tworząc formalny opis języka polskiego, gdy we wstępie do *Metafor w naszym życiu* Lakoffa i Johnsona natknęłam się na zdanie Tomasza Krzeszowskiego:

[...] fałszywy jest opis języka oparty na [...] tradycyjnych podziałach i rozróżnieniach, które zawsze i nieuchronnie prowadzą do zaniedbania lub wręcz pomijania w opisie przypadków pośrednich¹.

Zdanie to ilustruje problemy, z którymi zmagają się wówczas nasz zespół, a także, przypuszczalnie, każdy, kto kiedykolwiek zajmował się lingwistyką komputerową. Głównym zadaniem tej dyscypliny jest precyzyjny, formalny opis języka naturalnego w taki sposób, by otrzymana teoria znalazła swoje zastosowanie przy tworzeniu programów przetwarzających język – na przykład automatycznych tłumaczy lub wyszukiwarek². Najlepiej, by programy te były przygotowane do analizy różnorodnych wypowiedzi. Przy pomocy rygorystycznych reguł należy więc opisać jak najszerszy zakres zdań. I tu właśnie ujawnia się wspomniana niekompatybilność – jeśli nie ucieka się od „przypadków pośrednich”, pojawiają się problemy eksponujące fakt, jak bardzo matematyczny opis i binarne kategorie nie przystają do wszystkich kawałków języka. Kilka przykładów:

Sebastian to niesłychanie ciepły, pracowity i dobry człowiek.

Czy przysłówek *niesłychanie* określa tylko pierwszy z następujących przymiotników, czy całą ich grupę?

nauczyciel matematyki, gwiazda filmu, kredyt zaufania – mamy tu do czynienia z walencją rzeczownikową, posesywnością, a może jeszcze z czymś innym?

Wreszcie: czy w wyrażeniach: *komisja do spraw pracowników* czy *zbiórka na rzecz uchodźców* frazy *na rzecz* i *do spraw* są wyrażeniami przyimkowymi, czy też już przyimkami złożonymi?

W pracy magisterskiej szczegółowej refleksji poddałam tę ostatnią kwestię.

Pochylając się nad problemem pakowania żywołu języka w sztywne ramy formalnego opisu, zastanawiałam się, czy taki opis obok zastosowań praktycznych może mieć walory poznawcze. Jeśli tak – czego o języku można się z niego dowiedzieć? By odpowiedzieć na te

1 G. Lakoff, M. Johnson, *Metafory w naszym życiu*, Warszawa 2010, s. 18.

2 Por. M. Piasecki, *Cele i zadania lingwistyki informatycznej...*, op. cit., s. 285.

pytania, wpięrow musiałam opisać owe ramy, czyli analizowany przeze mnie formalizm, następnie – interesujący mnie fragment języka, by wreszcie ów fragment przebadac. I dzisiaj chciałabym pójsć tym tropem – najpierw omówić po krótkce teorię LFG, pózniej – rzec słów kilka o wyrażeniach, którymi się zajęłam, a w końcowej, najobszerniejszej części wystąpienia przedstawić założenia i podsumowanie badań.

Lexical Functional Grammar, w skrócie LFG, w zamyśle jej twórców miała być reprezentacją kompetencji składniowych, którymi dysponuje rodzimy użytkownik języka. Teoria miała dostarczyć sposobu przedstawiania uogólnień dotyczących składni, a także być narzędziem do badań lingwistycznych. Równie ważne było, by przy pomocy LFG dało się sformułować model języka, który byłby matematyczny i precyzyjny³.

U podstaw tworzenia LFG leżało przekonanie o tym, że w różnych językach istnieją te same abstrakcyjne relacje gramatyczne, natomiast są one wyrażane w odmienny sposób, na przykład poprzez szyk zdania czy budowę wyrazów⁴. W związku z tym w LFG składnię zdań języka naturalnego przedstawia się przy pomocy dwóch uzupełniających się struktur: c-struktury⁵ i f-struktury⁶. C-struktura jest drzewem składnikowym reprezentującym poziom linearnej i hierarchicznej organizacji słów we frazy, drzewko to może wyglądać znajomo dla wszystkich fascynatów gramatyki generatywnej. F-struktura zaś przedstawia abstrakcyjną organizację w zdaniu funkcji, takich jak podmiot, orzeczenie czy dopełnienie. Jest pewnym przybliżeniem semantyki zdania – nie jest równoznaczna ze strukturą semantyczną⁷, lecz stanowi bazę dla reprezentacji semantyki⁸.

Jednym z obszarów problematycznych do opisanego przy pomocy formalizmu LFG były przyimki złożone i, nazwijmy je, potencjalne przyimki złożone. Krótko mówiąc: trudno było zdecydować, gdzie postawić granicę, jak odróżnić przyimek złożony od zwykłego wyrażenia przyimkowego z podrzędnikiem w dopełniaczu. Przyimek złożony byłby wielowyrasową jednostką leksykalną pełniącą funkcję przyimka; pewną całością semantyczną i składniową, podczas gdy zwykłe wyrażenie przyimkowe interpretowalibyśmy w LFG analitycznie: na przykład w wyrażeniu

3 J. Bresnan, R.M. Kaplan, *Lexical-Functional Grammar: A Formal System for Grammatical Representation*, 1982, s. 2.

4 Por. M. Darymple, *Lexical Functional Grammar*, Academic Press 2001, s. 3.

5 ang. *constituent structure*

6 ang. *functional structure*

7 W celu przedstawienia semantyki stosuje się w LFG inny typ struktur, por. M. Darymple, *Lexical Functional Grammar*, Academic Press, Academic Press 2001, ss. 217–253.

8 Por. J. Bresnan, R.M. Kaplan, *Lexical-Functional Grammar: A Formal System for Grammatical Representation*, *op. cit.*, s. 3.

na tle flagi wyraz *tło* byłby dopełnieniem przyimka *na*, a wyraz *flaga* – podrzędnikiem posesywnym wyrazu *tło*.

W pracy magisterskiej pochyliłam się nad dwoma wyrażeniami – *na rzecz* i *do spraw*. Celem analiz było sprawdzenie, w jakim stopniu każde z nich jest zleksykalizowane w funkcji przyimka. Stopień leksykalizacji fraz zbadałam, studiując ich składnię i dystrybucję, zaś wnioski o tym, czy występują w funkcji przyimka, wyciągnęłam na podstawie analiz semantyki i pisowni.

Moje badanie postanowiłam maksymalnie oprzeć na korpusie. Badałam leksykalizację – a ważnym elementem tego procesu jest utrwalenie frazy w kolejnych użyciach. Korpus, który rejestruje użycia i pozwala je analizować w kontekście, a także zliczać, wydał się więc narzędziem idealnym. Sięgnęłam więc po Narodowy Korpus Języka Polskiego.

Posłużyłam się metodami badań znanymi z językoznawstwa korpusowego, ale również przeszczepiłam na grunt korpusowy testy na leksykalizację przyimków złożonych znane z literatury przedmiotu. Dotychczas bazowały one na kompetencji językowej badacza – u mnie podstawą analiz stały się dane korpusowe.

Analiza dystrybucji sprawdzała samodzielność rzeczownikowego członu wyrażenia. Okazało się, że, choć wyrażenia typu *rzecz gustu* czy *rzecz obyczajów* są stosowane, to jednak znacznie więcej jest wyrażen *na rzecz [czegoś]*. Inaczej ma się fraza *do spraw*: wyraz *sprawy* z dopełniaczowym podrzędnikiem prawie pięć razy częściej występuje, kiedy nie ma przed nim przyimka *do*; wyrażenie typu *do spraw czegoś* ma więc znacznie mniej wystąpień niż *sprawy czegoś*. O ile w wypadku *na rzecz* dystrybucja sugerowałaby pewne utrwalenie frazy, o tyle dystrybucja *do spraw* nie jest wystarczającym powodem do wyciągnięcia takiego wniosku.

W części składniowej przeprowadziłam szereg testów na leksykalizację fraz, zaczerpniętych z literatury przedmiotu. Testy te zaadaptowałam do badań korpusowych. Wyszukałam w korpusie wszystkie zdania, w których:

- elementy wyrażenia są od siebie oddzielone wyrazem określającym element rzeczownikowy (np. *do produkcyjnych spraw firmy*)
- podrzędnikiem badanego wyrażenia nie jest element nominalny w dopełniaczu, lecz zdanie (np. *był zaangażowany do spraw, jakie można załatwić w aptece*)

Wyszukałam również zdania podrzędne z zaimkiem *który* – tutaj sprawdziłam, czy zaimek ten został wstawiony pomiędzy elementy wyrażenia (**firma, na której rzecz przelałam pieniądze...* zamiast *firma, na rzecz której przelałam pieniądze...*).

Duża frekwencja takich zdań wskazywałaby na niski stopień leksykalizacji wyrażenia.

Korpus pokazał, że fraza *na rzecz* bardzo rzadko bywa rozdzielana, a więc jej forma jest raczej utrwalona. Wyjątek stanowią zdania z zaimkami osobowymi, w których zdecydowanie

częściej, niż ich forma poprzyimkowa, występuje zaimek dzierżawczy określający wyraz *rzecz* – na przykład *na moją rzecz* ma ponad dwudziestokrotnie więcej wystąpień niż *na rzecz mnie*.

W wypadku *do spraw* analiza składni nie dała przesłanek, by wnioskować o leksykalizacji frazy – przeprowadzając każdy z testów, obok przykładów dających podstawy do takiego osądu znalazłam równie liczne kontrprzykłady. Co ciekawe, wyróżniają się tu konstrukcje, w których nadrzędnikiem jest nazwa instytucji bądź zespołu osób (na przykład *forum do spraw bezdomnych*). W ich obrębie *do spraw* nie jest niczym rozdzielane.

Analizując semantykę, zbadalam preferencję semantyczną każdego z wyrażen oraz przeprowadziłam test substytucji semantycznej. Wyjaśnię, o co chodzi w obu testach.

Analiza preferencji semantycznej⁹ miała wskazać, ile różnych typów relacji mogą oddać badane frazy. Wychodziłam z założenia, że im bardziej dane wyrażenie zleksykalizowało się w funkcji przyimka, tym bardziej jest "uniwersalne", tym więcej różnych relacji można przy jego pomocy wyrazić. Dlatego też dla każdej z badanych fraz sprawdziłam jej sto najczęstszych podrzędników i nadrzędników. Następnie podzieliłam je na kręgi znaczeniowe. Jeśli dla danego wyrażenia tych licznie reprezentowanych kręgów było dużo, i były one różnorodne – przyjmowałam to za argument za leksykalizacją wyrażenia w funkcji przyimka.

Nadrzędniki wyrażenia *na rzecz* udało mi się podzielić na 11 typów semantycznych; najliczniejsze z nich to te odnoszące się do czynienia, obrotu pieniędzmi i do czynności społecznych oraz organizacji. Podrzędniki *na rzecz* dały się zaklasyfikować do ośmiu typów znaczeniowych, gdzie najliczniej reprezentowane były czynności, zdarzenia, instytucje oraz zbiorowości. Najbardziej typowe będą więc przykłady takie jak:

fundusz na rzecz rozwoju rolnictwa, pomoc na rzecz poprawy stanu dróg, stowarzyszenie na rzecz społeczeństwa informacyjnego czy darowizna na rzecz gminy

Dla *do spraw* wydzieliłam 9 semantycznych typów nadrzędnika i 6 typów podrzędnika. Badania preferencji semantycznej wskazały, że o tym, czy jest możliwość przyimkowej interpretacji *do spraw*, decyduje typ nadrzędnika tegoż wyrażenia. Nie sposób twierdzić, że mamy do czynienia z przyimkiem złożonym w konstrukcjach typu:

rozsądne **podejście do spraw** materialnych, w **odniesieniu do spraw** poradnictwa zawodowego, **przywiązywać wagę do spraw** produkcji

Warto jednak pamiętać, że aż 80% lewych kolokatów wyrażenia odnosi się do stanowisk, organizacji, instytucji lub zespołów osób. Konstrukcje z *do spraw* utworzone przy ich użyciu precyzują zakres obowiązków, specjalizacji, władzy bądź zainteresowań bytu wyrażonego w nadrzędniku.

9 Por. V. Kamasa, *Preferencja semantyczna jako narzędzie ujawniania konstrukcji ideologicznych w tekstach. Na przykładzie miłości w polszczyźnie ogólnej i tekstach religijnych*. [w:] *Niedosłowność w języku*, seria „Język a komunikacja”, Tertium, Kraków 2016, t.37.

Badania preferencji semantycznej *do spraw* z jednej strony potwierdziły tezę, że przyimkowa interpretacja frazy może być właściwa tylko w niektórych kontekstach, z drugiej zaś strony ukazały stosunkowo niewielkie zróżnicowanie tych kontekstów, a więc również oddawanych w nich relacji. Być może więc konstrukcje z badaną frazą stanowiące nazwy gremiów, stanowisk czy instytucji należałoby raczej poddać refleksji frazeologicznej, miast doszukiwać się w nich przyimka złożonego.

Wyniki badań preferencji wykorzystałam w kolejnej analizie – teście substytucji semantycznej. W oryginalnej jego wersji chodziło o to, by zastanowić się nad jednowyrazowymi przyimkami, którymi można zastąpić badane dwuczłonowe frazy¹⁰. Jeśli takowe istniały, świadczyłyby to o leksykalizacji wyrażenia w funkcji przyimka.

Ja nieco zmieniłam ten test, wyszukując w korpusie jednowyrazowe przyimki występujące w najbardziej typowych kontekstach semantycznych badanych wyrażen. Konteksty te znałam oczywiście z badania preferencji semantycznej. Dla każdego z dwóch wyrażen udało się znaleźć jednowyrazowe zamienniki, jednak zakres użycia żadnego z nich nie pokrywał się z zakresem użycia odpowiadającego mu wyrażenia złożonego. I tak: w konstrukcjach *z na rzecz* interesującą nas frazę najczęściej zastępuje się przyimkiem *dla* lub *na* (na przykład: *działać dla rozwoju, przekazać pieniądze na fundację*). W konstrukcjach *z do spraw* frazę zastępuje się przyimkami *dla*, *d o i o d* (na przykład: *minister właściwy dla administracji, minister właściwy do szkolnictwa wyższego czy dyrektor od produkcji*).

Na koniec analiz przebadalam pisownię nazw własnych. Zgodnie z zasadami polskiej ortografii, przyimki w nazwach własnych piszemy małą literą¹¹. Użycie w interesujących mnie wyrażeniach wielkich lub małych liter świadczyłaby więc o tym, czy autorzy tych zapisów postrzegają je jako przyimki. W wypadku *na rzecz* pisownia elementu rzeczownikowego rozkłada się mniej więcej po połowie, jeśli zaś spojrzymy na *do spraw*, rzeczownik *sprawy* dwukrotnie częściej zapisano wielką literą.

Przeprowadzone analizy ukazały dwa wyrażenia będące na różnym etapie procesu leksykalizacji.

Na leksykalizację wyrażenia *na rzecz* wskazują wyniki badań jego dystrybucji. Niejednoznaczność wyników badań składniowych sugeruje zaś, iż proces ten jeszcze się nie zakończył. Zróżnicowane pod względem semantycznym konteksty każą zaś myśleć o wyrażeniu w kategorii przyimka złożonego, mimo iż jego pisownia nie jest ustabilizowana.

10 Por. B. Milewska, *Przyimki wtórne we współczesnej polszczyźnie*, 2003, ss. 35–39.

11 E. Polański, *Zasady pisowni i interpunkcji* [w:] *Wielki słownik ortograficzny PWN z zasadami pisowni i interpunkcji*, E. Polański (red.), Warszawa 2003.

W wypadku *do spraw* na leksykalizację w pewnych kontekstach wskazują badania semantyczne i, po części, badania składniowe. Analiza składni przeprowadzona na pełnej próbie, pisownia wyrażenia w nazwach własnych, a także badania dystrybucji dostarczają jednak mocnych argumentów przeciw tezie o przyimkowym charakterze frazy.

Taki stan rzeczy może sugerować, że proces leksykalizacji wspomnianych fraz zaczyna się od zmian w ich semantyce. Element nominalny powoli traci swoje wyraziste znaczenie, wskutek czego rośnie liczba różnorodnych kontekstów, w których fraza może zostać użyta. Jednocześnie następuje swego rodzaju sprzężenie zwrotne – wraz ze wzrostem zróżnicowania kontekstów oraz frekwencji występowania badanej frazy, znaczenie elementu nominalnego słabnie. Sam rzeczownik staje się coraz silniej związany z poprzedzającym go przyimkiem, a całe wyrażenie coraz rzadziej występuje bez podrzędnego rzeczownika w dopełniaczu. Silne powiązanie elementów wyrażenia, utrwalane w coraz liczniejszych wypowiedzeniach z jego użyciem, ciągnie za sobą zmiany składniowe – fraza coraz rzadziej jest rozbijana przez wyrazy, a podrzędny wyraz w dopełniaczu staje się jej obligatoryjnym prawostronnym kontekstem.

To, że udało się uchwycić dwa wyrażenia będące na różnych etapach tego samego procesu, świadczy o ciągłym charakterze tego procesu. Moje analizy pokazują, że obszary języka ulegające aktualnie zmianie nie dają się zaś łatwo i jednoznacznie zinterpretować – co sprawia, że są problematyczne przy próbach wszelkiego rodzaju formalnej klasyfikacji, tym bardziej, gdy próbuje się opisać reprezentatywny zbiór tekstów, zawierający wypowiedzenia o różnym stopniu złożoności, przynależące do różnych rejestrów i odmian języka.

Okazuje się więc, że niekompatybilność może być owocna. Abstrahując od czysto praktycznych zastosowań teorii LFG, dużym jej walorem jest według mnie potencjał nie tyle w wyjaśnianiu, co właśnie w identyfikacji tych miejsc języka, w których pewne procesy właśnie zachodzą.

Powyższe stwierdzenie można pewnie odnieść do większości formalnych opisów języka. LFG wyróżnia jednak cecha, która czyni tworzenie reprezentacji zdań w według jej prawideł szczególnie dobrą metodą wychwytywania elementów będących na tym etapie zmiany, w którym ich interpretacja semantyczna nie zgadza się z interpretacją składniową. Cechą tą jest obecność i istotna rola f-struktury. Jak już wspomniałam, ma ona być bazą do budowy reprezentacji semantyki zdania, jednocześnie odwołując się do terminów składniowych, takich jak podmiot czy dopełnienie. W wypadku zdań, w których zmiana językowa dotknęła już semantyki wyrażenia, lecz nie utrzymała się całkowicie w składni, pojawiają się więc problemy związane z interpretacją przy użyciu f-struktury. Budowanie f-struktur może być zatem metodą, która pozwala na identyfikację zmiany językowej w trakcie jej trwania oraz ukazuje pełen jej przegląd poprzez wskazanie wszystkich

miejsz, w których zmiana zachodzi. Jak starałam się pokazać, metoda ta może doskonale korelować z innymi procedurami badawczymi i uzupełniać wnioskowanie naukowe.

Bresnan J., Kaplan R.M., *Lexical-Functional Grammar: A Formal System for Grammatical Representation*, 1982.

Darymple M., *Lexical Functional Grammar*, Academic Press, Academic Press 2001.

Grochowski Maciej, *Wyrażenia funkcyjne. Studium leksykograficzne*, Kraków 1997.

Kamasa V., *Preferencja semantyczna jako narzędzie ujawniania konstrukcji ideologicznych w tekstach. Na przykładzie miłości w polszczyźnie ogólnej i tekstach religijnych*. [w:] *Niedosłowność w języku*, seria „Język a komunikacja”, Tertium, Kraków 2016, t.37.

Krążyńska Zdzisława, *Staropolskie konstrukcje z przyimkami*, Poznań 2000, t.I.

Krążyńska Zdzisława, *Staropolskie konstrukcje z przyimkami*, Poznań 2012, t.5.

Lakoff G., Johnson M., *Metafory w naszym życiu*, Warszawa 2010.

Milewska B., *Przymyki wtórne we współczesnej polszczyźnie*, 2003.

Polański E., *Zasady pisowni i interpunkcji* [w:] *Wielki słownik ortograficzny PWN z zasadami pisowni i interpunkcji*, E. Polański (red.), Warszawa 2003.

Przybylska Renata, *Polisemia przyimków polskich w świetle semantyki kognitywnej*, Kraków 2002.