

Prof. dr hab. Wiesław Lubaszewski
Instytut Informatyki AGH
Katedra Lingwistyki Komputerowej UJ
Profesor emerytowany

Ocena rozprawy doktorskiej mgr Karoliny Jankowskiej: *Lexical convergence in Polish dialogues*, przeprowadzona na zlecenie Rady Naukowej Dyscyplin Językoznawstwa i Literaturoznawstwa Uniwersytetu Adama Mickiewicza w Poznaniu.

Praca doktorska mgr Karoliny Jankowskiej jest poświęcona badaniu zjawiska konwergencji leksykalnej w dialogach. Konwergencja leksykalna to szczególny przypadek konwergencji komunikacyjnej, czyli zjawiska dostosowywania/uzgadniania/uwspólniania sposobu mówienia przez uczestników procesu komunikacji. Konwergencja może zachodzić we wszystkich warstwach języka, a więc fonetycznej, morfologicznej, leksykalnej, syntaktycznej i semantycznej. W warstwie leksykalnej konwergencja może się przejawiać w taki np. sposób, że jeden z rozmówców zaczyna używać słownictwa drugiego rozmówcy lub obaj tworzą – różne od używanego wcześniej – słownictwo wspólne.

Pojawiająca się w trakcie dialogu konwergencja leksykalna ma zazwyczaj przyczyny pozajęzykowe, jednak Autorka nie zajmuje się badaniem przyczyn konwergencji, ograniczając się do czysto językoznawczego badania tekstu dialogu i jego związku z sytuacją komunikacyjną. Innymi słowy, Doktorantka nie zastanawia się nad tym, dlaczego jeden z przechodniów zapytanych: *Którędy na Wawel?* odpowie: *Na świątlach w lewo i cały czas prosto*, a inny powtórzy element pytania: *Na Wawel? Na świątlach w lewo, potem prosto*. Autorka bada tylko teksty dwu dialogów, by stwierdzić, że w sytuacji pytania o drogę konwergencja leksykalna może się pojawić już w pierwszej wymianie replik. Do badania konwergencji Doktorantka używa korpusu dialogów, a do analizy tekstów dialogu metod i narzędzi stosowanych w językoznawczych badaniach korpusowych. Nie ma więc najmniejszych wątpliwości, że badania opisane w pracy są badaniami językoznawczymi.

Doktorantka stawia cztery tezy:

- konwergencja leksykalna przejawia się już na poziomie części mowy,
- intensywność konwergencji oraz wybór słownictwa na poziomie części mowy są zmienne i zależą od sytuacji komunikacyjnej,
- rozmówcy dobierają formy grzecznościowe, biorąc pod uwagę własny status, status rozmówcy oraz sytuację komunikacyjną,
- dialogi na temat zagadnień wywołujących emocje są dłuższe niż dialogi, w których rozmówcy dążą do osiągnięcia wspólnego celu.

Recenzowana rozprawa składa się z wstępu, części teoretycznej, części analitycznej, dodatków, spisów rysunków i tabel oraz bibliografii. Część teoretyczna *Communication Convergence*, (s. 1-44) omawia zjawisko konwergencji (*Reasons and sources of linguistic*

alignment in conversation oraz *Levels of convergence*, s.11-22), metody badań konwergencji (*Methodology of communication convergence research*, s.23-27), dane (korpusy), dzięki którym można prowadzić badania (*Resources for language processing*, s. 28-33), polskie badania nad konwergencją fonetyczną (*Linguistic convergence in Polish*, s. 34-37) oraz przydatność badań konwergencji dla technologii dialogu człowiek – komputer (*Convergence studies in view of language technology*, s. 38-41).

Część analityczna *Lexical convergence in Polish dialogues*, s.45-112 to stanowiąca oryginalne osiągnięcie Doktorantki analiza zjawiska konwergencji leksykalnej na podstawie korpusu dialogów. Rozdział *Own research on lexical convergence*, s. 45 -77 zawiera opis budowy korpusu dialogów (*Research material*, s.45-54), opis metod analizy konwergencji (*Methodology*, s.55-62) oraz przedstawia narzędzia służące do przygotowania korpusu i analizy konwergencji (*Tools* s. 63-77). Rozdział *Lexical convergence results* (s.78-112) omawia wyniki analizy. Rozdział ostatni *Final summary and conclusion* (s.113-115) przedstawia podsumowanie i wnioski płynące z analizy. Pracę uzupełniają obszerne dodatki, z których większość przedstawia próbki danych poddawanych analizie w części drugiej, spisy tabel i wykresów oraz bibliografia.

Ze względów oczywistych w omówieniu treści rozprawy Karoliny Jankowskiej skoncentruję na tych elementach, które stanowią oryginalne osiągnięcie Doktorantki.

Podkorpus dialogów

Doktorantka wyodrębniła z korpusu z korpusu akustycznego Harmonia (zbudowanego przez zespół UAM) podkorpus dialogów, na którym prowadzi własne badania. Pozyskane eksperymentalnie dialogi reprezentują 9 sytuacji komunikacyjnych, kreowanych przez instrukcje udzielone uczestnikom eksperymentu. Są to:

- współpraca przy rozwiązywaniu problemu, zad. 5,
- pytanie o drogę, zad. 6 i 7,
- perswazja w okolicznościach pozytywnych, zad. 8 i 9,
- perswazja w okolicznościach negatywnych, zad. 10 i 11,
- uzgadnianie wspólnej oceny pozytywnej przez osoby o równym statusie, zad. 12,
- uzgadnianie wspólnej oceny negatywnej przez osoby o równym statusie, zad. 13,
- uzgadnianie wspólnej oceny pozytywnej przez osoby o różnym statusie, zad. 14,
- uzgadnianie wspólnej oceny negatywnej przez osoby o różnym statusie, zad. 15,
- uzgadnianie różnic w ocenach, zad.16.

W eksperymencie wzięło udział 32 rozmówców, którzy w pewnych sytuacjach zamieniają się rolami, tj. pytający o drogę staje się udzielającym informację. W rezultacie uzyskujemy pary dialogów na ten sam temat, inicjowanych przez każdego z rozmówców. Trzeba dodać, że reguła zamiany osób inicjujących dialog nie jest stosowana w sytuacji wspólnego rozwiązywania problemu (zad. 5) oraz w sytuacjach związanych z uzgadnianiem oceny wspólnej zad. 12, 13, 14, 15) a także ocen różnych (zad. 16).

W sumie, korpus nie jest zbyt duży. Jest jednak wystarczająco zróżnicowany i reprezentatywny, by można podjąć analizę zjawiska konwergencji i – co równie istotne – sprawdzić metody badania konwergencji leksykalnej.

Metody analizy zjawiska konwergencji leksykalnej

Doktorantka bada zjawisko konwergencji z czterech perspektyw, dokładnie omówionych w części pierwszej. Są to:

- formality level, analiza zbieżności form grzecznościowych,
- lexical choices, analiza zbieżności słownictwa z wyłączeniem słownictwa funkcyjnego,
- collaborative effort, stopień uczestnictwa w dialogu wyrażający się długością wypowiedzi, oraz liczbą jednostek leksykalnych,
- language style matching, analiza zbieżności słownictwa funkcyjnego (głównie zaimki, przymyki, partykuły).

Każda z perspektyw ma własną metodę badania.

Analiza wyników

Przedstawione w pracy wyniki analizy są uporządkowane według perspektyw, a w obrębie każdej z nich Doktorantka analizuje dialogi odpowiadające konkretnym sytuacjom komunikacyjnym. Wyniki analizy są ilustrowane wykresami i tabelami, pokazującymi analizowane zjawisko od strony statystycznej. Doktorantka ostrożnie formułuje wnioski, zdając sobie sprawę z tego, że analiza dialogów spontanicznych mogłaby skorygować niektóre wyniki.

Perspektywa formality level. Analiza pokazuje, że częstość użycia form grzecznościowych w dialogach zależy od dystansu społecznego między rozmówcami. Im mniejszy dystans, tym mniej form grzecznościowych.

Perspektywa lexical choices. Analiza obejmuje główne części mowy, tj. rzeczowniki, czasowniki, przymiotniki. Doktorantka starannie podaje częstości występowania części mowy w poszczególnych dialogach i stara się znaleźć związek pomiędzy sytuacją komunikacyjną i statystyką poszczególnych części mowy, np. w dialogu, którego celem jest ocena dzieła sztuki dominują wyrażenia wartościujące, więc najczęstsze są przymiotniki. W pytaniu o drogę najczęstsze są rzeczowniki i czasowniki itd. Jednak najciekawsza z perspektywy badania konwergencji analiza powtarzalności jednostek leksykalnych, sprowadza się tylko do bardzo starannego wskazania jednostek powtarzających się w obrębie poszczególnych dialogów – tabela 18 wynik zbiorczy oraz tabele 19 – 24 dla poszczególnych sytuacji komunikacyjnych. Doktorantka nie podejmuje dalszej analizy zjawiska powtarzalności.

Perspektywa collaborative effort. Analiza bada długość dialogu mierzoną za pomocą liczby użytych jednostek tekstu z uwzględnieniem wkładu, jaki wnosi każdy z rozmówców. Wyniki pokazują, że jest to raczej analiza sposobu mówienia poszczególnych osób.

Perspektywa language style matching. Względna analiza dystrybucji funkcyjnych części mowy jest bardzo staranna. Doktorantka oblicza wartość współczynnika LSM dla dialogu za

pomocą dwuetapowej procedury. Najpierw oblicza LSM dla poszczególnych funkcyjnych części mowy w dialogu, a następnie wartość średnią LSM dla wszystkich występujących w dialogu. Wartość średnia to LSM całego dialogu. Sposób obliczania LSM został opisany w części wstępnej na s. 26. Wyniki analizy są prezentowane osobno dla poszczególnych sytuacji komunikacyjnych. Zbiorcze zestawienie wartości LSM wszystkich dialogów (tab.25, s.104) pokazuje, że zachowanie funkcyjnych części nie jest jednoznacznie powiązane z sytuacją komunikacyjną. Widoczne w tabeli różnice mogą raczej wskazywać na różnice pomiędzy uczestnikami dialogów a także na to, czy pomiędzy rozmówcami istnieje dystans społeczny (dialogi z nauczycielką).

Dyskusja i ocena

Zjawisko konwergencji leksykalnej w dialogach polskojęzycznych jest mało zbadane, więc Doktorantka musiała się wykazać dużą samodzielnością wybierając kryteria i metody analizy zjawiska na podstawie badań anglojęzycznych. Dodatkowo, korpus którym dysponowała został zbudowany z myślą o badaniu zjawiska konwergencji fonetycznej, co znacząco wpłynęło na rozmiar korpusu.

Do analizy zjawiska konwergencji leksykalnej w korpusie eksperymentalnym Doktorantka mogła się posłużyć metodami, w większości przypadków opracowanymi do badań dialogów spontanicznych pozyskanych z Internetu. Wreszcie, część metod użytych w pracy została opracowana dla potrzeb badań psycholingwistycznych, w których analiza leksyki ma dostarczyć informacji na temat zachowań językowych uczestników dialogu.

Wszystko, co powiedziałem wyżej, pokazuje skalę trudności zadania, przed którym stanęła Doktorantka. Dlatego też będę traktować badania opisane w pracy jako studium pilotażowe, które pozwoliło zweryfikować metody badań konwergencji leksykalnej a także pozwoli zaplanować przyszłe badania dialogów spontanicznych. Nieliczne uwagi krytyczne w niewielkim stopniu wpłyną na ostateczną pozytywną ocenę pracy.

Analiza wyników za pomocą metod użytych w pracy pokazuje, że perspektywa collaborative effort, której istotą jest badanie długości replik jest mało przydatna z punktu widzenia analizy leksykalnej. Wniosek Doktorantki jest sformułowany znacznie ostrożniej. Doktorantka sygnalizuje tylko wątpliwość, mówiąc że niektórzy uczestnicy niektórych dialogów reagowali zbyt pasywnie na wypowiedź partnera.

Inaczej wygląda sprawa lexical choices. Analiza przedstawiona w pracy pokazuje, że jest to perspektywa bardzo przydatna do badania leksyki. Statystyki użycia części mowy w dialogach dobrze pokazują zależność pomiędzy leksyką a sytuacją komunikacyjną. Jednak Doktorantka nie wykorzystuje w pełni możliwości, jakie daje perspektywa lexical choices. Wykonana bardzo starannie i pracowicie ręczna analiza powtarzalności jednostek leksykalnych ogranicza się do zwykłego wskazania elementów powtarzalnych, co niestety słyca analizę. Szkoda, że Doktorantka rezygnuje z użycia opisanej w części pierwszej (s.24) statystycznej metody analizy powtarzalności leksykalnej LILLA, (Lexical Indiscriminate Local Linguistic Alignment). Przecież wystarczyło podzielić każdy dialog na pary replik i dokonać obliczeń. O tym, że Doktorantka nie miałaby z tym kłopotów pokazuje bardzo

dobrze zrobiona analiza funkcyjnych części mowy, która jednak nie wymagała segmentacji dialogu. Trzeba też dodać, że podział dialogów na pary replik pozwoliłby poszerzyć zakres analizy.

Można się wreszcie zastanowić na tym, czy zastosowanie metody LILLA nie wzbogaciłoby przedstawionej w pracy analizy form grzecznościowych.

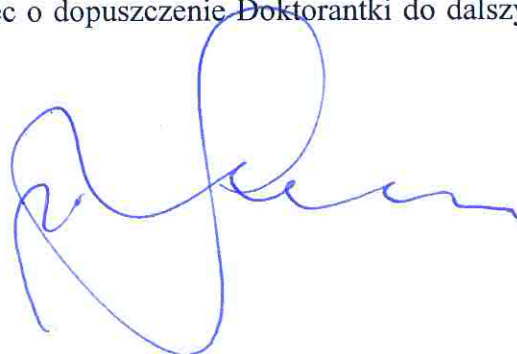
Podsumowując muszę powiedzieć, że Doktorantka dobrze sobie radzi z trudnościami, przed którymi stanęła. Praca jest oryginalna i samodzielna. Przedstawione w pracy analizy zjawiska konwergencji leksykalnej w dialogu są wystarczające, staranne i rzetelne. Doktorantka wykazała słuszność stawianych tez. Analizy wnoszą też nowe informacje na temat zachowań leksyki w dialogu. Słowem, praca mgr Karoliny Jankowskiej spełnia ustawowe i zwyczajowe wymogi stawiane pracom doktorskim z językoznawstwa. Przedstawione uwagi pokazują jedynie, że praca mogłaby być jeszcze lepsza.

Konkluzja

Przedstawiona analiza rozprawy doktorskiej mgr Karoliny Jankowskiej pozwala mi stwierdzić, że rozprawa spełnia ustawowe i zwyczajowe wymogi stawiane pracom doktorskim z językoznawstwa. Proszę więc o dopuszczenie Doktorantki do dalszych etapów przewodu doktorskiego.

W. Lubaszewski

Kraków, 2 września 2023 r.

A handwritten signature in blue ink, appearing to be 'W. Lubaszewski', written in a cursive style.