Complex Prepositions are considered to be typical examples of an ongoing process of grammaticalization. The aim of the present study is to produce an exhaustive list of Polish Complex Prepositions (henceforth CP) in a corpus-driven procedure, as well as to recognize the degree of grammaticalization of each of them.

We adopt a corpus driven approach, i.e. we treat as a candidate for a CP every single Prepositional Phrase which is (1) attested in the corpus in a sequence: 

\[ \text{[Noun or Verb]} + \text{[Preposition]} + \text{[Noun]} + \text{[Noun in genitive]} \].

In a second step, only those sequences are taken into account, which have a frequency of at least 10 hits. Subsequently we measure the (3) collocational strength between the Preposition and the Noun within the CP; (4) the number (of types) of right and left collocates of the CP; (5) the strength of the collocation between the complex preposition and its subordinate and superordinate phrases compared to the collocational strength between each of the constituents of the complex preposition separately and the adjacent words. (1) gives us a list of potential CPs, (2) and (3) show to what extent the expression is fixed, (4) informs us if the class of words opening a position for a CP is restricted (5) shows if a CP is a complement to different words than corresponding (simple) Preposition. As to (6), the quality of the intervening element may either support or weaken the complex reading of a potential CP – Adjectives and Demonstratives suggest referential interpretation of the nominal constituent of the CP, whereas the same does not necessarily hold for Possessive Pronouns. In the last step we confront our data with a dictionary of Polish Secondary Prepositions (Milewska 2003). The above procedures are mainly modifications of tests worked out in competence-based linguistics, however the important gain of approaching the question via corpus is that we can see "being prepositional" not as a dichotomic question but rather as a matter of degree.

Each test allows for establishing a hierarchy of potential CPs from "most prepositional" to "least prepositional", or PP-like for the very feature. By summing up the positions on each of the list we obtain an overall position on a scale extending from fully grammaticalized Complex Prepositions to ad hoc sequences of Prepositions and Nouns. This is in line with the view of grammaticalization considered as a gradual process (Bybee 2011, Hoffmann 2005, Poplack 2011, Torres Cacoullos & Walker 2011). In effect, a synchronic ‘gradience’ of categories (in our case the CPs) can be observed (Anderson 2001, Traugott & Trousdale 2013), what means that at any moment constructions belonging to one category can vary according to the degree of their grammaticalization.

References

