04

Palacký University of Olomouc

Dorota Jagódzka

Syntax of Complex Nominal Phrases in English and Slavic

DOROTA JAGÓDZKA

Syntax of Complex Nominal Phrases in English and Slavic[1]

The objective of this paper is a syntactic description of complex nominal expressions in English and two Slavic languages: Czech and Polish. The analysed nominal phrases are those with a head noun and multiple levels of pre-modification expressed by means of determiners, adjectives, numerals and so-called quantifiers, i.e. elements that indicate quantity, such as *many/much, plenty, little/few, all*, etc. In Czech and Polish, quantifiers can be classified into different groups according to the agreement patterns they exhibit when entering a relationship with a noun and a verb. The aim here is to compare the properties of those nominal expressions between the two Slavic languages and English and to provide some possible explanations for those special characteristics by reviewing recent research in linguistic as well as extralinguistic fields. The paper consists of data from English, Polish and Czech to show the differences between analytic and synthetic languages, which might be important and useful for interpreters, translators and students.

KEYWORDS

nominal phrase, Slavic, numerals, quantifiers

[1] This research was funded by *Ministry of Education, Youth and Sports of the Czech Republic (MŠMT in Czech)*, grant IGA_FF_2018_012 (*Internacionalizace v oblasti angloamerické literární vědy, lingvistiky a translatologie*).

1. INTRODUCTION

In many languages we can observe the classification of quantifiers into several groups based on the nature of the relations they exhibit with the head noun they modify, e.g. Veselovská (2001), Rutkowski (2002). Technically speaking, numerals constitute a subtype of quantifiers, where the first ones are numerically specific, the latter ones – existential and universal – are not. One of the most striking characteristics of numerals across languages is their division into lower numerals (1–4) and higher numerals (≥5). This is especially evident in Polish and other Slavic languages (for example Czech).

Lower numerals constitute a separate syntactic and morphological class, which is different from higher numerals. In some languages, higher numerals are not present at all (for example some of the New Guinea languages), as they are not believed to belong to a language's core vocabulary (Rutkowski 2003).

In English, the aforementioned division into two different morphosyntactic classes of numerals is not observed, as seen in the following examples:

i) *many / several / three apples were used to bake a pie*

ii) *all / a few / twenty apples were used to bake a pie*

It can be seen in (i) and (ii) above that the syntactic distribution of quantifying elements in English is not restricted according to any division between lower and higher numerals, and between universal or existential quantifiers. All of them can occur in the same place in the structure, and the forms of the nouns as well as verbs are identical with respect to case and agreement features. The next section will be concerned with the behaviour of quantifying elements in Polish and Czech and showing how different it is the English pattern in (i) and (ii) above.

2. SYNTAX OF NOMINAL PHRASES CONTAINING NUMERALS AND QUANTIFIERS IN PRESENT-DAY POLISH AND CZECH

This section will provide the basic syntactic facts about the behaviour of numerals and quantifiers in a sentence both in Polish and in Czech. First, the properties of lower numerals and universal quantifiers will be presented, and then higher numerals and existential quantifiers will be discussed. For the sake of further discussion, the quantifying elements in Polish and Czech will be divided into two groups:

Q_A - lower numerals and universal quantifiers

Q_{GEN} - higher numerals and existential quantifiers

Lower numerals and universal quantifiers (Q_A) comprise a group which consists of numerals 1-4 and universal quantifiers such as for instance *wszyscy* 'all$_{VIRAL}$' in Polish and *všechny* 'all$_{FEM}$'. Higher numerals and existential quantifiers are numerals ≥5 and universal quantifiers such as *wiele* 'many$_{FEM}$' in Polish and *mnoho* 'many' in Czech. The syntactic properties of both groups of quantifying elements in Polish and Czech will be presented in subsections 2.1 and 2.2.

2.1. LOWER NUMERALS (1–4) AND UNIVERSAL QUANTIFIERS

In Polish and Czech, lower numerals and universal quantifiers decline and agree with the noun that they precede like adjectives. The nominal expression agrees with the predicate in person, number and gender. They clearly show adjective-like declension and agreement patterns, as seen in the following data:

(1) Polish

a) *dwaj/ przystojni* *mężczyźni* *przyszli*
two/ handsome-NOM,Masc men-NOM-Masc, pl come-pastPRT,Masc,pl
'two/ handsome men came'

b) *dwie/ piękne* *kobiety* *przyszły*
two/ pretty-NOM,Fem women-NOM-Fem,pl come-pastPRT,Fem,pl
two/ pretty women came'

c) *wszyscy* *mężczyźni* *przyszli*
all-NOM,Masc men-NOM,Masc, pl come-pastPRT,Masc,pl
'all men came'

d) *wszystkie* *kobiety* *przyszły*
all-NOM,Fem women-NOM,Fem, pl come-pastPRT,Fem,pl
'all women came'

(2) Czech

dvě/všechny *dívki* *přišly*
two/all-NOM,Fem girls-NOM,Fem,pl come-pastPRT,Fem,pl
'two/all girls came'

In the above examples (1) – (2) we can see that Polish and Czech have the same pattern of agreement and distribution of lower numerals and universal quantifiers.

2.2. HIGHER NUMERALS AND EXISTENTIAL QUANTIFIERS

Higher numerals and existential quantifiers assign the genitive case to the noun they precede and always trigger the neuter agreement with the predicate. The following examples (3) – (4) illustrate this pattern:

(3) Polish

a) *pięć* *(pięknych) kobiet* *przyszło*
five-ACC[2],Fem pretty-GEN-pl, women-GEN-pl come-pastPRT,3sg,Neu
'five (pratty) women came'

b) pięciu *(przystojnych) mężczyzn* *przyszło*
five-ACC,Masc handsome-GEN-plmen-GEN,pl come-pastPRT,3sg,Neu
'five (handsome) men came'

c) *wiele* *(pięknych) kobiet* *przyszło*
many-NOM,Fem pretty-GEN-pl women-GEN,pl come-pastPRT,3sgNeu
'many (pretty) women came'

(4) Czech

mnoho / osm *chlapců* *přišlo*
many/ eight-NOM boys-GEN,pl come-pastPRT,3sgNeu
'many/eight boys came'

However, if a nominal phrase is put into the inherent case context, the Q_{GEN} does not assign the genitive case to the noun any longer. The noun and its potential premodifier(s) take the inherent case from the verb, e.g.:

(5) Polish

a) *Dałam to* *pięciu* *miłym* *koleżankom.*
five-Dat nice-Dat,pl friends-Dat,pl
'I gave it to my five nice friends'

[2] Polish higher numerals are intrinsically accusative according to the so-called Accusative Hypothesis. For further reference see for instance Miechowicz-Mathiasen (2011).

b) *Poszłam do kina* *z* *pięcioma* *koleżnakami.*
with five-Instr,pl friends-Instr,pl
'I went to the cinema with my five friends.'

c) Rozmawiałam o pięciu koleżankach z pracy.
about five-Loc,pl friends-Loc,pl
'I was talking about my five friends from work.'

An interesting question to ask is why lower and higher numerals behave differently in a morphosyntactic sense. The answer can perhaps be found in the historical development of this word class. The next section will be an attempt to reconstruct the way quantifying elements obtained their present-day status.

3. DIACHRONIC DEVELOPMENT OF NUMERALS

The following section will focus on the historical development of numeral elements, first by looking into reconstructed Proto-Indo-European data, followed by Old Polish and Old Czech.

3.1. PROTO-INDO-EUROPEAN NUMERALS

In Proto-Indo-European, there was no unified syntactic and morphological class for the numeral terms – they were very diverse in morphosyntactic terms and inflectionally belonged to all the existing declensions types, i.e. to the pronoun, noun and adjective declensions.

They did not differ from other word categories regarding internal structure. All of them consisted of three elements: the lexical part (connected with semantic meaning), the thematic vowel (the declension class classifier) and the inflectional ending (indicating syntactic relations) (Długosz-Kurbaczowa and Dubisz 2001, 250–251). The following data (6) is a reconstruction of Proto-Indo-European numerals one, two and three.

(6)

a) thematic vowel *-o-*

PIE	**oino-//eino-*	PSL **(jed)in,*	Pol. *jeden*; 'one'
PIE	**duuo-*	PSL ** dva,*	Pol. *dwa*; 'two'

b) thematic vowel *-i-*

PIE	**trei- -*	PRS **trje,*	Pol. *trzy*; 'three'

(Długosz-Kurbaczowa and Dubisz 2001, 250–251)

In (6) the numerals 1–3 have been grouped according to their thematic vowels *-o-* and *-i-*. Also, the reconstruction for Proto-Slavic (PSL) and the present-day Polish (Pol.) forms have been given for each numeral.

3.2. PROTO-SLAVIC NUMERALS

The formation of a semantic class of numeral terms began with changes within one word, namely the 'numeral name' *jeden* 'one'. In the beginning, this was a three-gender pronoun which meant 'some' or 'certain'. It declined like the demonstrative pronouns *tъ, to, ta 'this'.* The paradigm was the following:

(7) Nom sg. **jedinъ, *jedino, *jedina;*
Nom pl. **jedini, *jedina, *jediny* (and suppletive forms);
Nom dual *dъva, *dъve, **dъve.*

(Długosz-Kurbaczowa and Dubisz 2001, 250–251)

The 'numeral name' *jeden* 'one' did not have a regular dual number but only appeared in singular and plural. This initial change caused a semantic shift from **jedinъ* 'some/certain' to **jedinъ* 'one', however, the word itself remained a pronoun and not a numeral.

Other numeral nouns were created from adjectives like **thъje, *četyre*, which have an irregular declension, which enabled them to become numerals easier. In this way, the semantic class of numerals was shaped. It consisted of the following members:

(8) *jedinъ, **dъva* — originally pronouns
**trъje, *četyre* — originally adjectives
**pętъ, *sestъ, *sedmъ, *osmъ, *devętъ *desętъ *sъto* — originally nouns

It can be safe to assume that in Proto-Slavic, a new part of speech – the numeral – appeared as a semantic class that did not have its own morphological features, nor its own type of declension (consult Długosz-Kurbaczowa and Dubisz 2001).

4. NUMERALS IN OLD CZECH AND OLD POLISH

The Old Polish period is considered to be a breakthrough in the development of the Polish language in general. During that time, spoken language and its archaic Proto-Slavic constructions combined with the written language that developed under the influence of Latin. As a result,

a new written version of Old Polish emerged (Słoboda 2012)
.

Numerals in Polish started to diverge only in the Middle Ages as a separate class of words. One of the important inflectional changes was the loss of the dual number in Old Polish, which brought together the lower numerals and evened out their distribution.

The Czech language went through a similar development with a few internal differences. It kept the general tendency of Slavic languages regarding the development of numerals, which started to form as a distinct word class at the stage of transition from common Slavic into individual Slavic languages (Basaj 1974). For a long time now, scholars (Klemensiewicz, Lehr-Spławiński, and Urbańczyk 1955, Basaj 1974, Słoboda 2011, Słoboda 2012) have thought of ways to account for this special syntax of quantifying elements presented in section 2. What is the origin of such a state?

In the following sections, some of the data collected by Basaj (1974) from over 240 different historical texts written in Old Czech will be presented. The examples will illustrate the general syntactic properties that were true for constructions with numerals in that period.

4.1. FORMS OF NOUNS WITH NUMERALS 1–4 IN OLD CZECH

Numerals 1–4 agree with nouns in number, gender and case in Old Czech. Dual number started to disappear in the 15th and 16th century, and went extinct in the 17th century. So, in the Old Czech period there are two kinds of noun forms possible with numerals: plural forms and dual forms. Until the end of the 15th century we can basically notice only dual forms of nouns; plural forms became more common in the second half of the 15th century. Consider the following example for the usage of the plural form of a noun:

(9) *po* *dvú* *let*
after two years-pl
'after two years'

(Basaj 1974, 230)

Sometimes, in the same source there are both forms next to each other, dual and plural, like in: *dwie figure* DUAL - *dwie figury* PLURAL 'two figures' (Basaj 1974, 320). In the 16th century, we notice that the plural forms of nouns become more prominent in use than the dual ones. What is interesting is that dual forms occur longer when standing next to the numeral 2 than

without it. Until the end of the 16th century it was more common to use the dual form of the feminine noun for the numeral 2, than masculine or neuter.

Interestingly, in Old Czech there are examples with nouns in genitive plural forms with numerals 2–4. Consider the following examples:

(10)

a) *prosil těch trzy kraluov*
three kings-GEN,pl
'He begged those three kings.'

b) *čtyři člověků poslal*
four men-GEN,pl send-pastPRT,3sgMasc
'He sent four people.'

(Basaj 1974, 231)

The examples in (10) show the inherent case contexts, where, similarly to present-day Czech, nouns also occur in genitive plural case form.

4.2. FORMS OF NOUNS WITH NUMERALS 5–10 IN OLD CZECH

Numerals 5–10 were originally feminine nouns. When put together with a noun, it was only the numeral that got inflected; the other element (a noun) always remained in a genitive case form, like in Present-day Polish nominal constructions of the following kind: *róg* NOM *domu* GEN 'the corner of a house', *rogu* GEN *domu* GEN, *rogiem* INSTR *domu* GEN. In Old Czech, only a few examples of such a type survived, for instance:

(11)

a) *po sedmy let král jede*
seven-ACC years-GEN, pl
'for seven years the king has been going'

b) *budeš vlasti nad desety měst*
ten-ACCcities-GEN, pl
'you will be home after ten cities'

(Basaj 1974, 233)

In the above examples, it is always the numeral that governs the genitive case of the following noun. This was true for nominative, dative, genitive, and accusative. In the remaining cases, the system had changed and the noun started to get the case that was required by the context, i.e. with

dative, locative and instrumental the noun gets the inflectional endings and the numeral does not, it usually stays in nominative. Consider the following example:

(12) *kterák se jemu uvázal v šest člověků*
to six-NOM man-GEN,pl
'who tied him to six people'

(Basaj 1974, 234)

The example (12) shows the opposite situation to the present-day one, where in the inherent case contexts both the quantifying element and the noun inflect the same way.

4.3. FORMS OF THE PREDICATE WITH NOMINAL PHRASES CONTAINING NUMERALS IN OLD CZECH

Nominal phrases containing a quantifying element agree in person with the predicate of the sentence. The main verb takes the third person inflection to agree with the subject of the sentence which is exemplified in (13) below. Consider the following example:

(13) *třie* *králi* *přijeli*
threenom kingsnomPL arrive-come,pastPRT,3sg,Masc
'There arrived three kings.'

(Basaj 1974, 261)

There are no accounts for predicates in singular with noun phrases containing numerals 2–4 until the end of the 15th century (Basaj 1974).

For numerals 5–10, the main verb took either the singular or the plural form. There did not seem to be any regular restrictions as to whether the main verb should only take the plural form. From the data collected by Basaj (1974), 60% of the sentences have the main verb in the singular from. The majority of plural verb forms found in his corpus come from Bible translations. This finding can correspond to Słoboda's (2012) assumption that Latin had a significant influence on the structure of Old Polish. However, the biblical pattern did not survive because in Present-day Czech and Polish the predicate is always in the singular form when put together with a noun phrase containing a quantifying element (see section 2 of this paper).

As far as gender is concerned, the predicate takes the same gender

as the gender of the noun phrase containing numerals 1–4. Consider the example in (14) below:

(14) *pakli by sě o to ti* *třie* *páni* *dělili*
three gentlemenMasc do-pres,3pl,Masc

If the noun phrase contains numerals ≥5, then the predicate occurs in neuter in a singular or plural form. When in plural, it gets the same gender as the noun phrase, i.e. feminine, masculine or neuter. Consider the example (15):

(15) *v tu dobu* *pět* *rytieřuov* *ciesařovi* *povědĕli*
fiveNom knightsNom,pl,Masc emperor-Dat,sg tell-pastPRT,3,pl,Masc
'At that time, five knights told the Emperor'

(Basaj 1974, 266)

The pattern in (15) is quite different to what is known today about Czech and Polish. In the Present-day languages, the only option to agree in gender for the predicate and a noun phrase containing numerals ≥5 is the neuter gender (consider section 2, examples (3) and (4) in this paper).

5. WHAT CAN PSYCHOLINGUISTICS TELL US ABOUT THE NATURE OF QUANTIFYING ELEMENTS?

Language is one of the many capabilities that biology has provided humans with and it is an integral part of human cognition. It is not an isolated and independent ability; rather it is motivated by various factors of physical and mental nature.

When higher numerals appeared in a language for the first time, their status must have been different from the status of the lower numerals. They were originally nouns, as this is the category used to denote abstract concepts in a language (Rutkowski 2003).

Humans possess a sense for numbers, which can be observed in recent experiments (see Rutkowski 2003) where the findings show that people can easily capture the number of elements in a set if the set contains one to four elements. The sets that contain more elements constitute a problem for human perception. The same phenomenon is true for our memory: we can easily remember up to four elements in a row. However, the findings show that a limit for the number of elements in the surrounding neighbourhood that can be simultaneously captured by human perception also exists.

In many languages, there are morphological and syntactic differences within individual semantic classes of numerical expressions that define the boundary between lexemes denoting lower (1–4) and higher numerical values (from 5 onwards). This especially applies to cardinal numerals. The Middle Ages were a period in which languages like Old Polish and Old Czech still exhibited the internal diversity within the so-called lower numerals. Explanations of this state should be sought in Indo-European origin numerals. Perhaps originally there were only the numerals *one, two* and *three*, and the latter could then mean 'many', as Ifrah (1990, 15) states that the number *three* was already a threshold for the human numerical system which cannot be understood or defined, and "in the mind of man the discovery of numbers stopped first on the two". The proof for the existence of such an initial limitation may be the early distinction between singular, dual and plural in Indo-European languages (Słoboda 2012). Researchers dealing with this issue indicated that the differentiation of morphosyntactic numerals must be due to some universal conditions related to processes occurring in the brain. The universalism of this phenomenon manifests itself in that the division into lower and higher numerals is not a mere feature of Indo-European languages.

The meanings of numerals indicated by the lowest numbers could be perceived and presented like the meanings of basic adjectives, e.g. red, hot, round etc., i.e. as visible properties of an object or a group of objects (Hurford 1990). Lower numerals can therefore be considered perceptually as the features of objects. Rutkowski (2003) notices that perception of a set of less than 5 elements is something else (less complicated) than abstract counting – a demanding process requiring more advanced data processing in the human brain. Numerals denoting numbers above 4 may have appeared in language when people already developed their arithmetic skills. Therefore, the status of higher numerals had to be different from the status of lower numerals, which captured the number as a feature of the object (Rutkowski 2003).

6. CONCLUSION

The basis for the morphosyntactic diversity of numerals in Slavic must have been conditioned due to universal properties related to the processes taking place in the brain. The universalism of the phenomenon manifests itself in the fact that the morpho-syntactic division into lower and higher numerals is not a feature found only in Indo-European languages (Hurford 1990). However, in most European languages with the English language being a perfect example this division is not present. It is assumed that it happened because human mind could minimalize the perceptual differences between sets consisting of four and five elements.

Slavic languages, on the other hand, introduced an innovation which transformed higher numerals, which were originally adjectives, first into nouns and then into a separate numeral word category (Siuciak 2008, 16-17). This process is responsible for the special status and morpho-syntactic behaviour of numerals in Present-day Polish and Czech. The discussion concerning the historical origins of numerals in Polish and Czech seems to prove the psycholinguistic hypothesis. The status and morphosyntactic properties of these elements were indeed already different in the early stages of language development.

When attempting to translate the numeral expressions from English into Slavic, translators must remember about the special syntax they possess in the target languages. The correct use of case inflection and also the subject predicate agreement are aspects which require much attention. Not only is the knowledge of Slavic grammar an important prerequisite, but also being aware of the historical development of the numeral class may be very useful for translators and interpreters who struggle with the aforementioned issues.

REFERENCE LIST

Basaj, Mieczyslaw. 1974. *Morfologia i składnia liczebnika w języku czeskim do końca XVI wieku*. Wrocław: Zakład Norodowy im. Ossolińskich.

Długosz- Kurbaczowa Krystyna, and Stanisław Dubisz. 2001. *Gramatyka Historyczna Języka Polskiego.* Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.

Hurford, James Raymond. 1990. *Language and Numbers. The Emergence of Cognitive System*. Oxford–Nowy Jork.

Ifrah, Georges. 1990. *Dzieje liczby, czyli historia wielkiego wynalazku.* Wrocław: Zakład Narodowy im. Ossolińskich.

Klemensiewicz, Zenon, Tadeusz Lehr-Spławiński, and Stanisław Urbańczyk. 1955. *Gramatyka historyczna języka polskiego.* Warszawa: Państwowe wydawnictwo naukowe.

Miechowicz-Mathiasen, Katarzyna. 2011. "Licensing Polish Higher Numerals: an Account of the Accusative Hypothesis". In: *Current Issues in Generative Linguistics,* edited by Joanna Błaszczak, Bożena Rozwadowska, and Wojciech Witkowski. Vol. 2: 58-75.

Rutkowski, Paweł. 2002. "The Syntax of Quantifier Phrases and the Inherent vs. Structural Case Distinction." *Linguistic Research* 7(1): 43 – 74.

Rutkowski, Paweł. 2003. "Neuropsychologiczne uwarunkowania składni liczebników głównych." *Scripta Neophilologica Posnaniensia* 5: 209 – 233. Poznań: Wydział Neofilologii Uniwersytetu im. Adama Mickiewicza.

Słoboda, Agnieszka. 2011. "Pozajęzykowe uwarunkowania średniowiecznej morfologii i składni liczebników głównych." *Ling Varia* 2 (12): 91–102.

Słoboda Agnieszka. 2012. *Liczebnik w grupie nominalnej średniowiecznej polszczyzny. Semantyka i składnia.* Poznań: Wydawnictwo Rys.

Siuciak, Mirosława. 2008. *Kształtowanie się kategorii gramatycznej liczebnika*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.

Veselovská, Ludmila. 2001. "Agreement Patterns of Czech Group Nouns and Quantifiers." In *Semi-lexical Categories: The Function of -Content Words and the Content of Function Words*, edited by Norbert Corver and Henk van Riemsdijk, 273 – 320. Berlin: Mouton de Gruyter.