

The syntax of Polish cardinal numerals

ABSTRACT. This paper presents an overview of Polish cardinal numerals with reference to their history and development. It shows that higher (once nominal) numerals 5< have undergone numeralisation and now lexicalise the head of NumP, whereas the lower ones 1-4 remained adjectival and belong in its specifier. To account for the puzzling issues of the Polish numeral system, particularly for the so-called Accusative Hypothesis (a descriptive fact about numeral expressions with cardinals 5< according to which they are intrinsically accusative), an analysis is proposed in which these numeral expressions are introduced by a null light preposition (*p*). This preposition is an instance of a defective (interpretable) tense head (*iT*) and is argued to be the source of the said accusative case. An account of case changes (or lack thereof) within numeral expressions is proposed based on Case Hierarchy and case-inclusion analyses. Lastly, it is shown how the numeral expressions remain nominal (DPs) despite being introduced via *p*.

Keywords: cardinal numeral; structural/lexical case; preposition; NumP; extended projection

1 Introduction

In this paper, I present an analysis of Polish cardinal numerals and address the puzzling issues intrinsic to their internal and external syntax. I address the problem of the grammatical category of numerals, the case distribution both within and outside the numeral expression, as well as sentential agreement triggered by numeral subjects. The analysis concerns both simple and complex cardinals with a particular emphasis on their history and development. It is shown that cardinal numerals 1-4 are adjectival (and have been so since Proto-Slavic), and that higher – once nominal – numerals 5< have been reanalysed as lexicalisations of the functional head Num of NumP, i.e. have undergone numeralisation and as a result have become a separate grammatical category. I further propose how this process proceeded exemplifying it with *tysiąc* ‘thousand’ currently undergoing it. Before moving on to my proposal, I present an overview of some previous analyses of numerals and refer to Polish works discussing what is known to Slavists as the Accusative Hypothesis: a descriptive fact about numeral expressions with cardinals 5< according to which they are intrinsically accusative. To account for this vexing issue of the Polish numeral system, I propose an analysis in which accusative numeral expressions are introduced by a (null) light preposition. The analysis is couched within the framework proposed by Pesetsky & Torrego (2004) in which all case is a reflection of tense, and where also prepositions instantiate tense heads. The proposed null light preposition is an instance of a defective (interpretable) tense head (*iT*), which I argue to be the source of the accusative case. I further argue that the Polish so-called adnumeral operators (*około* ‘around/about’ and *ponad* ‘over’) instantiate such (overt) light prepositions. Further support comes from a related language, Serbo-Croatian (Franks 2002, Giusti & Leko 2004), where arguments in the form of numeral expressions with cardinals 5< in oblique case environments must be introduced by an overt light preposition (otherwise unavailable with the same verbs if the argument is not a numeral expression of the relevant type). I also propose an account of the case changes (or lack thereof) within numeral expressions via reference to Blake’s (1994) Case Hierarchy and the case-inclusion analysis proposed by Caha (2009). Lastly, I show how the numeral expressions remain nominal despite being introduced via a preposition with reference to Matushansky’s (2006) proposal of m(orphological)-merger.

2 Polish cardinal numerals: an overview

2.1 Adjectival numerals 1-4

Polish numerals 1-4 inherited their adjectival nature from Proto-Slavic and have retained this status to this day. Though their syntax was then and is to this day primarily adjectival, their Proto-Slavic declension varied between pronominal (*jedinъ* ‘one’ and *dъva* ‘two’, parallel to the demonstrative *тъ* in SG and DU respectively¹ (Comrie 1992: 805; Siuciak 2008: 17)), and nominal (*trije/tri* ‘three’ and *četyre/četyri* ‘four’). With the existing three numbers: SG, DU, and PL, *jedinъ* combined only with singular nouns and itself showed singular inflection (now we also have plural forms of *jeden*, cf. Table 1), *dъva* combined with dual forms of nouns and showed dual inflection, whereas *trije/tri* and *četyre/četyri* combined with plural nouns and showed plural inflection. Crucially, these numerals did not have their own number and gender distinctions, nor did they have a separate case paradigm, and whenever they combined with nouns they exhibited the relevant ϕ -features of the nouns in question. Upon reduction of the dual in (Middle) Polish (around mid-16th c.), just two numbers remained: SG and PL; *dъva* ‘two’, previously combining with dual nouns, began combining with plural forms (initially keeping its own dual forms, to later be influenced by the forms of the closest numerals 3-4, Siuciak 2008: 54). Further historical changes only deepened the differences, in particular, the development of new gender distinctions in PL. This caused an uneven division between the genders represented in SG (feminine, masculine, neuter) and those represented in PL (initially masculine animate, later narrowed down to masculine personal (henceforth: virile), as opposed to (literally) the rest: inanimate masculine, feminine and neuter (non-virile)).

CASE	1(SG)	1 (PL)	2 (PL)	3 (PL)	4 (PL)
NOM	<i>jeden_M, jedna_F jedno_N</i>	<i>jedne</i>	<i>dwaj/dwóch/dwu_V dwie_F, dwa_{NV}</i>	<i>trzej/trzech_V trzy_{NV}</i>	<i>czterej/czterech_V cztery_{NV}</i>
ACC	<i>jednego_M, jedną_F jedno_N</i>	<i>jedne</i>	<i>dwóch/dwu_V dwie_F, dwa_{NV}</i>	<i>trzech_V trzy_{NV}</i>	<i>czterech_V cztery_{NV}</i>
GEN	<i>jednego_{M/N}, jednej_F</i>	<i>jednych</i>	<i>dwóch/dwu</i>	<i>trzech</i>	<i>czterech</i>
DAT	<i>jednemu_{M/N}, jednej_F</i>	<i>jednym</i>	<i>dwom/dwu/dwóm</i>	<i>trzem</i>	<i>czterem</i>
INSTR	<i>jednym_{M/N}, jedną_F</i>	<i>jednymi</i>	<i>dwom/dwu_{M/N}, dwiema/dwoma_F</i>	<i>trzema</i>	<i>czterema</i>
LOC	<i>jednym_{M/N}, jednej_F</i>	<i>jednych</i>	<i>dwóch/dwu</i>	<i>trzech</i>	<i>czterech</i>

Table 1 Case paradigms 1-4 (Modern Polish, based on *Słownik Języka Polskiego*)²

Numerals 2-4 have become similar over time, however, with respect to gender distinctions, *dwa* ‘two’ still remains exceptional in that it has three gender forms in Nom/Acc, whereas *trzy* ‘three’ and *cztery* ‘four’ have only two; their declensions strongly resemble each other. *Jeden* ‘one’ shows a three-way gender distinction in SG, and inflects for case in agreement with the modified noun; these properties of *jeden* are unavailable in complex numerals, where it is always uninflected and shows no agreement whatsoever (1b,c).

¹ The following abbreviations are going to be used throughout: (i) genders: masculine animate/inanimate (M_{ANIM/INANIM}), virile (V), non-virile (NV), feminine (F), neuter (N); (ii) cases: nominative (Nom), accusative (Acc), genitive (Gen), dative (Dat), instrumental (Instr), locative (Loc), vocative (Voc); (iii) tense/mood/aspect/voice: present (PRES), past (PAST), future (FUT), perfective (PERF), imperfective (IMPERF), imperative (IMP), infinitive (INF); (iv) number: singular (SG), dual (DU), plural (PL).

² See also Comrie (1992) for a thorough presentation and discussion of all the Balto-Slavonic paradigms and their etymology.

- (1) a. Jeden chłopiec/Jedna dziewczyna/Jedno dziecko przyszedł/przyszła/przyszło.
 [one boy]_{M.NOM}/[one girl]_{F.NOM} / [one child]_{N.NOM} came_{3.SG.M/F/N.PAST}
 ‘One boy/girl/child came.’
 b. Jedne usta wystarczą.
 one_{PL.NV.NOM} lips_{NOM} suffice_{3.PL.FUT}
 ‘One set of lips is enough.’
 c. Dwudziestu (jeden) chłopców/dziewczyn/dzieci przyszło
 twenty_{NV} one_? boys_V/girls_{NV}/children_{NV} came_{3.SG.N}
 ‘Twenty (one) boys came.’

Sentential agreement subjects containing *jeden* results in full agreement with SG (1a) and PL (1b) *jeden*, but in complex numerals (1c), *jeden* appears to be ignored and the resulting agreement is default 3.SG.N, just like with 5<.³ Sentential agreement with nouns containing numerals 2-4 in virile depends on their form, of which there are two: (i) ending in *-j* (*dwaj/trzej/czterej*), with distinct Nom and Acc, and (ii) ending in *-ch* (*dwóch(dwu)/trzech/czterech*) where Nom/Acc is syncretic with Gen. The former have no influence on sentential agreement just like exemplary adjectives (2a,b)/(3a), the latter behave like numerals 5< and the numeral expression as a whole triggers 3.SG.N agreement (2c)/(3c).

- (2) a. Dwaj chłopcy przyszli.
 two_{V.NOM} boys_{V.NOM} came_{3.PL.M}
 ‘Two boys came.’
 b. Dwie dziewczyny/Dwa koty przyszły.
 two_{F.NOM} girls_{NOM}/two_{NV.NOM} cats_{NOM} came_{3.PL.NV}
 ‘Two girls/cats came.’
 c. Dwóch chłopców przyszło.
 two_{V.NOM/ACC} boys_{V.GEN} came_{3.SG.N}
 ‘Two boys came.’
- (3) a. Dwaj chłopcy zostali ukarani.
 two_{V.NOM} boys_{V.NOM} stayed_{3.PL.V} punished_{V.NOM}
 ‘Two boys were punished.’
 b. dwaj ukarani chłopcy
 two_{V.NOM} punished_{V.NOM} boys_{V.NOM}
 ‘two punished boys’
 c. Dwóch chłopców zostało ukaranych.
 two_{V.NOM/ACC} boys_{V.GEN} stayed_{3.SG.N} punished_{V.GEN}
 ‘Two boys were punished.’
 d. dwóch ukaranych chłopców
 two_{V.NOM/ACC} punished_{V.GEN} boys_{V.GEN}
 ‘two punished boys’

In the passive examples (3a,c), we also see participial agreement, which most probably should be treated in terms of concord (adjectival agreement), as shown in (3b,d) (see section 4.3.1).

Interestingly, the Nom virile forms *dwaj/trzej/czterej* cannot be used in complex numerals:

- (4) a. *Dwadzieścia/*Dwudziestu dwaj /trzej /czterej chłopcy (przyszli).

³ *Słownik Języka Polskiego* (Dictionary of Polish Language) gives a prescriptive rule according to which the last numeral imposes case on the counted noun, to the exclusion of 1, in which case it is the numeral preceding 1 that takes over this role.

- twenty_{NV} /twenty_V two_{V.NOM}/three_{V.NOM} /four_{V.NOM} boys came_{3.PL.V}
‘Twenty two/three/four boys came.’
- b. Dwudziestu dwóch /trzech /czterech chłopców przyszło.
twenty_{V.NOM/ACC} two_{V.NOM/ACC}/three_{V.NOM/ACC}/four_{V.NOM/ACC} boys came_{3.SG.N}
‘Twenty two/three/four boys came.’

2.2 Nominal numerals 5<

Polish numerals 5< inherited their nominal status from Proto-Slavic. Numerals 5-10 used to be nouns of the nominal *i*-declension (Grappin 1950: 26-27, Comrie 1992: 747, Siuciak 2008: 18). When part of the subject, it was the numeral that was syntactically more prominent and thus triggered sentential number (SG/PL) and gender (F)⁴ agreement (not the counted noun); the relation between the numeral and the counted noun resembled that of head-government in that the numeral checked lexical Gen on the noun (just like nouns do to this day), and forced plural marking on it. The lexical Gen on the noun remained unchanged throughout the case paradigm despite the varying cases of the numeral itself. The nominal status of numerals 5< has undergone change, which I present on the example of *sześć* ‘six’ in Table 2: beginning with a purely nominal inherited status, nominal *sześć* marked its complement with lexical Gen; the changes proceeded gradually, until the numerals and their complements reached congruence in all lexical cases (there was and still is no congruence in structural cases). This has been taken by historians as evidence for the influence of the syntax of numerals 2-4 on the syntax of 5<.⁵

CASE	INHERITED FORM	16 TH CENTURY	17 TH CENTURY	18 TH CENTURY
GEN	sześci _{GEN} panów _{GEN} ‘six men’s’	sześci _{GEN} panów _{GEN}	sześci _{GEN} /sześciu _{GEN} panów _{GEN}	sześciu _{GEN} panów _{GEN}
DAT	sześci _{DAT} panów _{GEN} ‘(to) six men’	sześci _{DAT} panom _{DAT}	sześciom _{DAT} panom _{DAT}	sześciu _{DAT} panom _{DAT}
INSTR	sześcią _{INSTR} panów _{GEN} ‘(with) six men’	sześcią _{INSTR} panów _{GEN}	sześcią _{INSTR} panów _{GEN}	sześciu _{INSTR} panami _{INSTR}
LOC	sześci _{LOC} panów _{GEN} ‘(about) six men’	sześci _{LOC} panów _{GEN}	sześci _{LOC} /sześciu _{LOC} panach _{LOC}	sześciu _{LOC} panach _{LOC}

Table 2 Changes affecting nominal complements of numerals 5< (Siuciak 2008: 162)

The relationship we see between the numeral and its complement in the ‘inherited form’ still exists in Modern Polish with numerals such as *tysiąc* ‘thousand’ and *million* ‘million’, both masculine (non-virile in the plural) nouns with SG and PL forms (cf. Table 3):

⁴ Both Comrie (1992: 748) and Siuciak (2008: 18) add that *desęć* ‘ten’ initially behaved slightly differently in that it triggered both masculine and feminine agreement.

⁵ Importantly, these influences were happening in both directions, i.e. we find examples of numerals 2-4 with Gen complement nouns, though these changes never became systematic (Siuciak 2008: 162): *trzy*_{NOM} *grzywnien*_{GEN} ‘three fines’ (16th c., Klemensiewicz 1930: 86), *cztery*_{NOM} *poduszek*_{GEN} ‘four pillows’ (16th c., Słoboda 2005: 18), often with a fronted noun: *miał żon*_{GEN} *dwie*_{ACC} ‘(he) had wives two’ (Paprocki 1584); where the verb marks Acc case (cf. *miał dwie*_{ACC} *żony*_{ACC} ‘(he) had two wives’); this construction can also be found today: *lat*_{GEN} *mam trzydzieści*_{ACC} *dwa*_{ACC} (cf. *mam trzydzieści*_{ACC} *dwa*_{ACC} *lata*_{GEN} ‘I am thirty-two-years old (lit. I have thirty two years)’). Such constructions are thoroughly analysed by Pereltsvaig (2008) in Russian. Pereltsvaig refers to them as ‘genitive theme’ constructions and proposes separate base-generation of the two parts (the numeral and the noun, the noun being introduced by a silent quantifying head responsible for Gen case). What supports this analysis for Polish is the fact that it is impossible to have the fronted agreeing version of *lat*_{GEN} *mam trzydzieści*_{ACC} *dwa*_{ACC}, namely **lata*_{ACC} *mam trzydzieści*_{ACC} *dwa*_{ACC} (although it is possible to say *żony miał dwie* ‘wives, he had two’, which could be simply a result of topicalization).

CASE	‘thousand zlotys’ (SG.M)	‘six thousand zlotys’ (PL.NV)
NOM	<i>tysiąc</i> _{NOM} <i>złotych</i> _{GEN}	<i>sześć tysięcy</i> _{GEN} <i>złotych</i> _{GEN}
ACC	<i>tysiąc</i> _{ACC} <i>złotych</i> _{GEN}	<i>sześć tysięcy</i> _{GEN} <i>złotych</i> _{GEN}
GEN	<i>tysiąca</i> _{GEN} <i>złotych</i> _{GEN}	<i>sześciu tysięcy</i> _{GEN} <i>złotych</i> _{GEN}
DAT	<i>tysiącowi</i> _{DAT} <i>złotych</i> _{GEN}	<i>sześciu tysiącom</i> _{DAT} <i>złotych</i> _{GEN}
INSTR	<i>tysiącem</i> _{INSTR} <i>złotych</i> _{GEN}	<i>sześcioma tysiącami</i> _{INSTR} <i>złotych</i> _{GEN}
LOC	<i>tysiącu</i> _{LOC} <i>złotych</i> _{GEN}	<i>sześciu tysiącach</i> _{LOC} <i>złotych</i> _{GEN}

Table 3 Case paradigm of *tysiąc* ‘thousand’

With 5<, the Gen case marking on the counted noun still exists in structural cases (Nom and Acc), however, this Gen has been reanalysed as structural, and like structural cases it is outranked (and trumped) by lexical ones (Babby 1987). We see this in Table 3 (*sześć tysięcy złotych* ‘six thousand zlotys’), where the (structural) Gen on *tysiąc* present in Nom and Acc (and by default also in Gen), changes accordingly into Dat, Instr, and Loc, as opposed to the (lexical) Gen on *złotych* ‘zlotys’ (in *tysiąc złotych* ‘thousand zlotys’) which remains unchanged throughout.⁶

The Modern Polish declension of 5< may not look impressive, but a lot of changes happened for this to be so (Table 4 exemplifies *pięć* ‘five’ as a representative of the group).

CASE	‘these five men’ (V)	‘these five women’ (NV)	‘these five houses’ (NV)
NOM	<i>pięciu mężczyzn</i> _{GEN}	<i>pięć kobiet</i> _{GEN}	<i>pięć domów</i> _{GEN}
ACC	<i>pięciu mężczyzn</i> _{GEN}	<i>pięć kobiet</i> _{GEN}	<i>pięć domów</i> _{GEN}
GEN	<i>pięciu mężczyzn</i>	<i>pięciu kobiet</i>	<i>pięciu domów</i>
DAT	<i>pięciu mężczyznom</i>	<i>pięciu kobietom</i>	<i>pięciu domom</i>
INSTR	<i>pięcioma mężczyznami</i>	<i>pięcioma kobietami</i>	<i>pięcioma domami</i>
LOC	<i>pięciu mężczyznom</i>	<i>pięciu kobietach</i>	<i>pięciu domach</i>

Table 4 Case paradigm of *pięć* ‘five’ with virile/non-virile nouns (Modern Polish)

⁶ For reasons of space, I will not discuss collective numerals, fractions and vague numerals, although my analysis can be applied to them, too. Collective numerals such as *dwoje/troje/czworo/pięcioro* ‘two/three/four/five’ developed slightly differently from the cardinals and have always exercised highly restricted usage. This is due to their limited distribution, which may well be caused by their narrow semantics (they used to combine with singularia tantum, e.g. *wojsko* ‘army’, *naród* ‘people’, now they combine with nouns denoting young/immature beings of unspecified sex, e.g. *dzieci* ‘children’, *kocięta* ‘kittens’, plural forms of nouns denoting persons of different genders (often, though not necessarily, generic) *czworo studentów* ‘four students’, pluralia tantum *nożyczki* ‘scissors’, *spodnie* ‘trousers’). Siuciak (2008: 43) notes that collectives were rarely used expressing a number higher than 20, and nowadays they are being ousted by alternatives: nouns *dwójka/trójka/czwórka* ‘a two/three/four’, extension of regular numerals to nouns earlier mostly combining with collectives *cztery kociaki/kocięta* ‘four kittens’ (instead of *czworo kociąt*), and formations such as *para spodni* ‘a pair of trousers’. Since collectives continue to assign lexical Gen to their counted nouns, they are treated on a par with *tysiąc* ‘thousand’. Fractions are combinations of (mostly) simple numerals and ordinal adjectives, e.g. *jedna czwarta* ‘one fourth’, *trzy czwarte* ‘three fourths/quarters’, or nouns: *pół* ‘a half’, *półtora* ‘one and a half’. Polish fractions seem to involve ellipsis of the noun *część* ‘part’ (F.SG, NV.PL), e.g. with numerals 1-2 which distinguish between F/M/N we see F agreement together with the ordinal adjective: *jedna*_{F.SG.NOM} *czwarta*_{F.SG.NOM} ‘one fourth’ no matter what the gender of the counted noun is. This suggests the presence of a (measure) noun ‘part’ within fractions, a noun which often surfaces: *czwarta*_{F.NOM} *część*_{F.NOM} *publiczności*_{NV.GEN} ‘the fourth part of the audience = one fourth of the audience’. Further support comes from the presence of lexical Gen in fractions, where neither the cardinal nor the ordinal can assign it (we know it is lexical because the counted noun remains Gen no matter the case-context: *z czwartą*_{F.INSTR} *częścią*_{F.INSTR} *publiczności*_{NV.GEN} ‘with the fourth part of the audience’). Vague numerals (*kilku* ‘several’, *wielu* ‘many’) were actually the first to take on the numeral declension (-u ending), and Siuciak (2008) argues that it was *kilku* which influenced the acquisition of this declension by numerals 5-9, as they are semantically closest to it, hence they receive treatment on a par with 5<.

In the two most common cases: structural Nom and Acc (marking subjects and objects respectively), we still see the remnants of the past paradigm. The changes were spurred by the loss of the dual number (around 16th c.) which ended with *dwa* ‘two’ entering the plural declension and thus developing the gender markings appropriate to it. The *-j* ending visible on 2-4 in Table 1 eventually became a distinctive feature of virile gender in these numerals. The same role was later taken on by the *-u* ending in numerals 5<.⁷ Importantly, only the structural cases (Nom and Acc) made the crucial gender distinctions, as shown in Table 4.

With gender changes happening so fast, it is rather obvious that nouns have not developed any special marking for the newly introduced genders, and other means were employed to express them. One of these was the Acc/Gen syncretism for virile nouns in the plural, and the continuation of the Nom/Acc syncretism for non-virile nouns.⁸ The Acc/Gen syncretism in virile plurals is responsible for the spread of the numeral *-u* ending, i.e. *-u* is the original Gen ending. Interestingly, the emergence of numerals as a separate grammatical category, the loss of their own gender and number distinctions, coincides with the gender changes and because numerals are naturally associated with plurality, these are numerals which become the primary exponents of the new virile/non-virile distinction. Indeed, Janda (1999: 217) reports that the Acc/Gen syncretism with viriles in the plural is initially used only in the presence of numerals (initially 2-4, later also 5<) and only later spreads to non-numeral expressions.

With respect to sentential agreement, these numerals always trigger 3.SG.N in all genders.

- (5) a. Pięciu chłopców przyszło
 five_{V.NOM/ACC} boys_{GEN} came_{3.SG.N}
 ‘Five boys came.’
 b. Pięć dziewczyn/kotów przyszło.
 five_{NV.NOM/ACC} girls_{GEN} /cats_{GEN} came_{3.SG.N}
 ‘Five girls/cats came.’

Numeral nouns *tysiąc* ‘thousand’ and *milion* ‘million’ trigger agreement in accordance with their own ϕ -features, although they may alternatively allow 3.SG.N agreement.

- (6) a. Tysiąc listów przyszło/przyszedł do Piotra.
 thousand_{M.NOM} letters_{NV.GEN} came_{3.SG.N/M} to Peter_{GEN}
 ‘A thousand letters came to Peter.’
 b. Ten tysięcy listów *przyszło/przyszedł do Piotra.
 this_{M.NOM} thousand_{M.NOM} letters_{NV.GEN} came_{3.SG.N/M} to Peter_{GEN}
 ‘This thousand letters came to Peter.’

This is only possible if *tysiąc* does not take its own agreeing demonstrative pronoun: the ϕ -features shared by the demonstrative are too prominent and block *tysiąc*’s numeral behaviour. I will account for this later on by proposing that it is in the process of numeralisation and that what occurs is reminiscent of *sto* ‘hundred’ (Table 5), which also started off as a noun with gender (N) and number properties (SG/PL), and lost these features in the process of numeralisation. Until mid-19th c. *sto* allowed both nominal and numeral declensions, and respectively triggered true and default agreement (Siuciak 2008: 160). Later, the numeral declension is preferred and suggested as such by grammarians (Małecki 1863, Kryński 1900).

⁷ This property also distinguishes Polish numerals 5< from the Russian ones, which do not mark animacy (Corbett 1978: 63).

⁸ See also Janda (1999: 216-219) for a discussion of the source and development of this syncretism in the plural.

CASE	NOMINAL DECLENSION (NEW POLISH)				NUMERAL DECLENSION (MODERN POLISH)	
NOM	SG:	sto	PL:	sta	SG (only):	sto
ACC		sto		sta		sto
GEN		sta//stu		set		stu
DAT		stu		stom		stu
INSTR		stem		stami		stu
LOC		stu		stach		stu

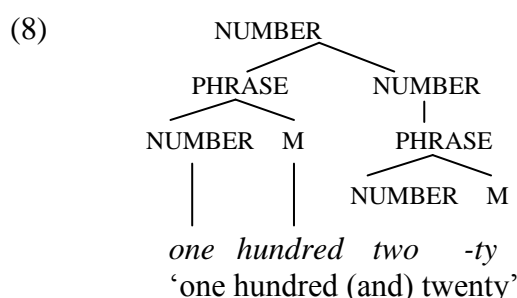
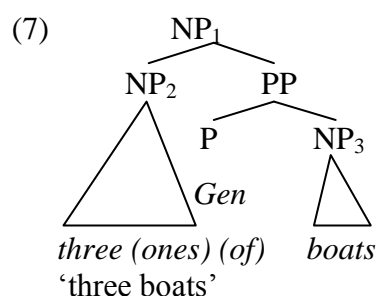
Table 5 19th c. nominal and numeral declension of *sto* ‘hundred’ (Jakubowicz 1823: 141)

3 Some previous analyses of the syntax of numerals

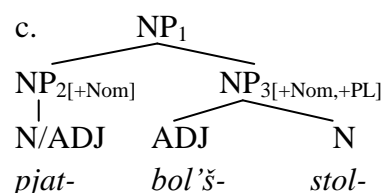
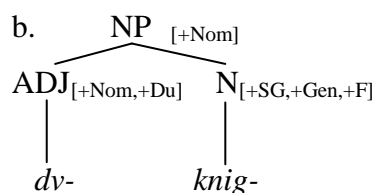
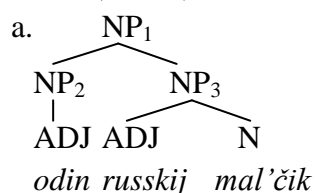
Judging by the complexity of the phenomena represented by numerals and their long history, one would expect a comparably complex and long tradition of their syntactic analyses. This, however is not the case as modern linguists were not as taken with numerals as, for instance, ancient Indian scientists responsible for the creation of numeral systems, thus combining knowledge of both grammar and arithmetics (Ifrah 1998: 845-846). Even though the numeral systems belong as much in language as in arithmetics (one intertwining with the other), they have not received as much attention in the former, as in the latter. Important works that must be mentioned are Greenberg’s (1978) and Stampe’s (1976) typological works, followed by Seiler (1990), Comrie (1992), Gvozdanović (1992, 1999) and more recently also Hurford (2003), the cognitive based approach by Heine (1997), as well as the already mentioned historical works by Ifrah. Most syntactic analyses of numerals to date seem to have concentrated upon the issue of the grammatical category of numerals, and here Slavic languages have figured quite prominently. It was (and still is) of utmost importance to establish the grammatical category of numerals, or else we cannot really analyse their syntax. Below, I review some of the most influential analyses on the subject.

3.1 Numerals: what are they?

Within the generative tradition, Perlmutter & Orešnik (1973) (7), and later Hurford (1975) (8), closely followed by Corbett (1978a, b) (9), are the first attempts to provide a syntactic analysis of numeral expressions.



(9) Corbett (1978b)



Both Perlmutter & Orešnik's (1973) and Hurford's (1975) structures are analysed by Corbett (1978a,b), who points out with respect to (7) that postulating a PP node between the two NPs, where P marks Gen, is problematic because it does not allow for distinctions between the various uses of *of*, nor does it account for why sometimes *of* is overtly realized and sometimes not (both within one language, as in English *The River Thames* vs. **The town London*, and cross-linguistically: English *a glass of water* vs. Danish *et glas vand*). Thus he changes Perlmutter & Orešnik's proposal into a binominal structure getting rid of the preposition and arriving at (12a). To account for the presence of Gen, he proposes a 'genitive insertion rule', which is dependent on the 'nouniness' of the numeral, i.e. only the more 'nouny' numerals will trigger Gen insertion. In Hurford's (1975) proposal (8) we have Number and M (M stands for 'multiplier', and may be also understood as base), where M is proposed to be represented by nouns. Corbett takes issue with the proposal that M be a substitute label for noun, and claims that in Russian the matters are not so clear-cut, and that numerals categorically-wise fall somewhere between adjectives and nouns. Moreover, he affirms that what we see in Russian is, in fact, quite general. To prove his point that Russian cardinal numerals can be graded on their level of 'nouniness', he composed the following table, which I reproduce as Table 6.

	odin	dva	tri	pjat'	sto	tysjača	million
	1	2	3	5	100	1,000	1,000,000
1. Agrees with N in syntactic number	+	—	—	—	—	—	—
2. Agrees in case (throughout)	+	—	—	—	—	—	—
3. Agrees in gender	+	(+)	—	—	—	—	—
4. Marks animacy	+	+	+	—	—	—	—
5. Has own plural	—	—	—	—	(+)	+	+
6. Takes agreeing determiner	—	—	—	—	—	+	+
7. Takes N in Gen.PL (throughout)	—	—	—	—	—	±	+

Table 6 Syntactic behaviour of Russian cardinal numerals (from Corbett 1978a: 64)

I reproduce Corbett's squish applying it to 16th c. Polish and Modern Polish (with a minor modification, i.e. I add another property: 8. Has own gender). The contents of the two tables show that Polish not only differs from Russian, but that the matters are not as simple as Corbett proposes.

	1	2	3	5	100	1,000	1,000,000
1. Agrees with N in number	+(SG)	+(DU)	+(PL)	—	—	—	—
2. Agrees in case (throughout)	+	+	+	—	—	—	—
3. Agrees in gender	+	+ Nom/Acc	+ Nom	—	—	—	—
4. Marks animacy	+	+ Acc	+ Nom	—	—	—	—
5. Has own plural	—	—	—	+	+	+	+
6. Takes agreeing determiner	—	—	—	+	+	+	+
7. Takes N in Gen.PL (throughout)	—	—	—	+	+	+	+
8. Has own gender	—	—	—	+	+	+	+

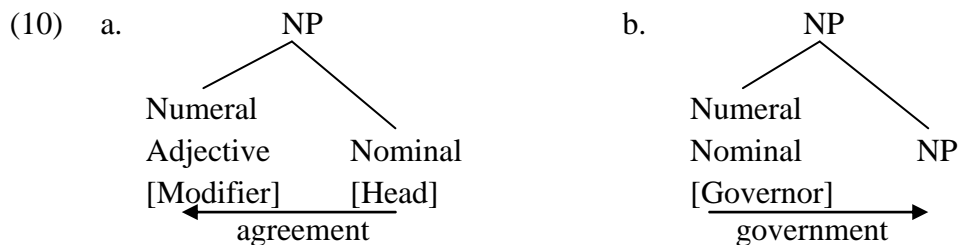
Table 7 Syntactic behaviour of 16th c. Polish cardinal numerals

	1	2	3	5	100	1,000	1,000,000
1. Agrees with N in number	+	–	–	–	–	–	–
2. Agrees in case (throughout)	+	+	+	+ except Nom/Acc	+ except Nom/Acc	–	–
3. Agrees in gender	+	+Nom/Acc/Instr	+Nom/Acc	+Nom/Acc	+Nom/Acc	–	–
4. Marks animacy	+Acc	+Nom/Acc	+Nom/Acc	+Nom/Acc	+Nom/Acc	–	–
5. Has own plural	–	–	–	–	–	+	+
6. Takes agreeing determiner	–	–	–	–	–	+	+
7. Takes N in Gen.PL (throughout)	–	–	–	+Nom/Acc	+Nom/Acc	+	+
8. Has own gender	–	–	–	–	–	+	+

Table 8 Syntactic behaviour of Modern Polish cardinal numerals

As we can see in Table 7, in 16th c. Polish there was a clear-cut division between numerals 1-4 (adjectives)⁹ and 5< (nouns) (contra Corbett 1978: 65). This division became blurred with time (cf. Table 8), in particular crucial changes affected numerals 5<, which are now seemingly closer to the adjectival paradigm, or less ‘nouny’ in Corbett’s terminology.¹⁰ The question remains: what are they? (cf. section 4).

Hurford (1987), following Corbett’s findings, elaborates on his earlier proposal and admits that whatever stands for Number and M may indeed be represented by either adjectives or nouns, but adds that the syntactic relations between the numeral and the counted noun may be of two types: the modifier-head (agreement) and head-phrase (government), where the first one is the numeral and the second the counted noun (Hurford 1987: 189).



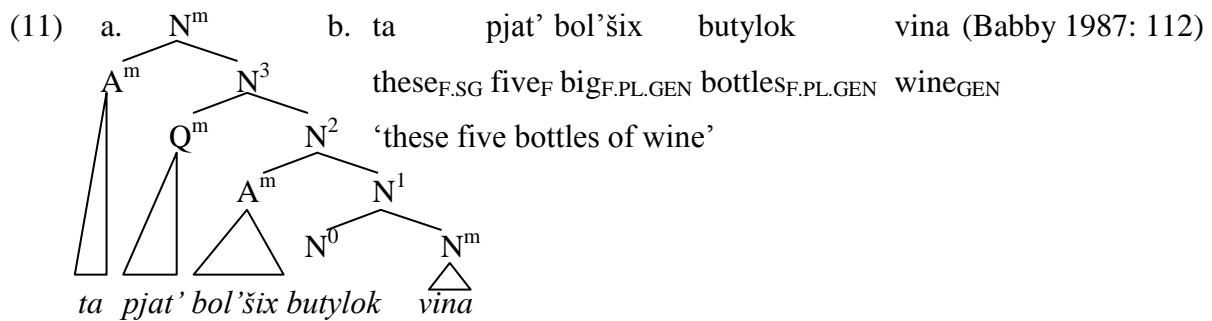
This is an important improvement in the syntactic analysis of numerals. The structures in (10) quite clearly show the head or specifier status of the given numeral (something that is missing in Corbett’s structures). The head-like status of some of the more ‘nouny’ numerals has been taken up in contemporary analyses (Nelson & Toivonen 2000, Ionin & Matushansky 2006 Bailyn & Nevins 2008), and will be elaborated on later in my own analysis. Corbett’s idea

⁹ The differences we see between 1-2 as opposed to 3 (and later 5<) follow from their separate historical development, where 2 initially represented dual of which it was the sole representative (very much like 1 in singular) and combined with dual nouns, adjectival 3-4 and nominal 5< both combined with plural nouns.

¹⁰ In Polish plural declension, animacy blends with gender in virile, and interestingly it is also neither of these in non-virile. To be more precise, the initial distinction between M_{ANIM} (as opposed to the rest), clearly singled out masculine animate beings (male persons and animals), whereas the later introduced virile/non-virile distinction narrowed it down to just male persons. So, on the one hand, we have a group of virile nouns which are both animate and have gender (M), and on the other hand, we have a huge group of non-virile nouns which includes all the rest making nothing of their animacy or gender. It is thus difficult to tell exactly what we are talking about in the plural, gender or class.

that numerals are to be found somewhere on a continuum between nouns and adjectives, although not an explanation, gives us an idea of the relevant properties. Crucially, both analyses spurred some crucial attempts at providing the answers.

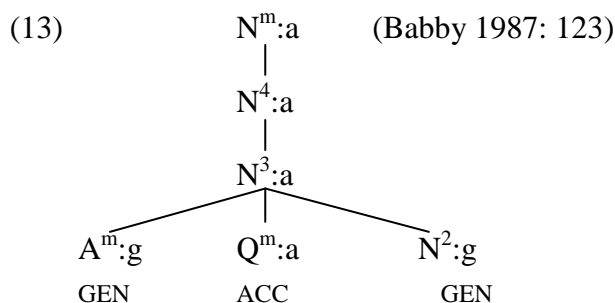
Babby (1987) analyses numerals 5< as something that one can only perceive as a hybrid category. In his structures, they are nominal heads allowed to case-mark their nominal complement (N^2), but they are merged as Q(uantifiers) (and are thus considered modifiers), which indicates that they are not head nouns of the numeral expression, although they are heads, nonetheless.



In (11), we see that whatever Q c-commands is marked with Gen of quantification (Gen on the noun *vina* is not assigned by Q, but by N^0 , and is the so-called partitive genitive, which I refer to as lexical). The goal of Babby's investigation, however, were constructions in which the regular pattern observed in (11) is disrupted by the presence of a modifier preceding Q, which he calls a prequantifier, and which does not agree with Q, but rather agrees with the Gen head noun. He refers to this phenomenon as discontinuous agreement (12):



For these constructions Babby comes up with a solution involving case-percolation in case-domains, to be precise, he proposes a ternary branching structure (13) in which Q c-commands both the prequantifier and N^2 , thus case-marking them both. The fact that Q does not bear Gen follows from the fact that being the case-assigner it cannot c-command itself, and thus case cannot percolate onto it. It bears Acc case assigned by the outside case-assigner. The reason why the other lexical items do not agree in Acc and do not exhibit the Acc case assigned by the outside case-assigner the way Q does, but instead exhibit the Gen assigned by Q results from Babby's (1987: 116) Case Hierarchy: Lexical Case > Gen(Q^m) > Nom/Acc (a descriptive statement made with respect to Russian).



Apart from the fact that ternary branching is itself a problematic issue, the discontinuous agreement should rather be analysed as an instance of movement of the modifier (prequantifier), rather than base-generation in its surface position. This solution Babby discards as untenable by saying that prequantifiers modify quantifiers not nouns. Since I cannot vouch for my Russian, I will refrain from the comments concerning the interpretation of Babby's examples, however, parallel examples are available in Polish and the interpretation of the Gen marked prequantifier can only be that of modifying the counted noun, and never that of modifying the numeral. If we want to modify the numeral, we need to employ the agreement strategy in (11). Moreover, it is possible to have the same modifier used twice with distinct cases modifying the numeral and the counted noun respectively.

- (14) a. dobre pięć butelek
 good_{NOM/ACC} five_{NOM/ACC} bottles_{GEN}
 'a good five bottles'
- b. dobrych_i pięć _{t_i} butelek
 good_{GEN} five_{NOM/ACC} bottles_{GEN}
 'five good bottles'
- c. dobre pięć dobrych butelek
 good_{NOM/ACC} five_{NOM/ACC} good_{GEN} bottles_{GEN}
 'a good five good bottles'

Similarly, Gvozdanović (1999: 190) analyses parallel Polish and Russian data with the universal quantifier *wszyscy* 'all' and numerals (examples that originally appeared in Corbett (1978b: 10)), and refers to the interaction of the scope of the numerals and quantifiers. For these, the movement analysis is the most conducive one:

- (15) a. wszystkie pięć pociągów
 all_{NOM/ACC} five_{NOM/ACC} trains_{GEN}
 'all the five trains'
- b. wszystkich_i pięć _{t_i} pociągów
 all_{GEN} five_{NOM/ACC} trains_{GEN}
 'five of all the trains'

As we can see from the translation, the moved Gen quantifier gives us a partitive reading. This is a more general pattern also observed with possessives and demonstratives:¹¹

- (16) a. te/moje pięć książek
 these/my_{NOM/ACC} five_{NOM/ACC} books_{GEN}
 'these/my five books' (the total number of books I possess is five)

¹¹ Both possessives and demonstratives have been claimed to have different base and surface positions within the nominal structure. While this is highly plausible for the possessive (see also Alexiadou et al. 2007: 567), it seems that with the demonstrative in (16b, c) we might be dealing with a difference between a mono- (16b) and biphrasal (16c) structure, i.e. between a pseudo-partitive (like English *a glass of wine*) and a true partitive (like English *a glass of the wine*; see Koptjevskaja-Tamm (2009) for a discussion of these issues). This means that whenever we have the Gen form of the demonstrative *tych* preceding a non-virile noun, it has undergone movement from below the numeral. This would entail a bi-phrasal (stacked double DP) structure. Rutkowski (2007) argues for two positions of the demonstrative within one DP, one base-position and the other derived, surface, position. His analysis, however, faces a problem of not being able to derive (16a): *te_{ACC} pięć_{ACC} książek_{GEN}*, and is only able to derive the more marked *tych_{GEN} pięć_{ACC} książek_{GEN}*. This is because the base position of the demonstrative in his analysis is always lower than the numeral and thus in the scope of the numeral's Gen of quantification.

- b. pięć tych/moich książek (partitive interpretation)
 five_{NOM/ACC} these/my_{GEN} books_{GEN}
 ‘five of these/my books’
- c. tych/moich pięć książek (partitive interpretation)
 my_{GEN} five_{NOM/ACC} books_{GEN}
 ‘five of /these my books’

In (16b,c) it does not matter whether the Gen marked possessive or demonstrative precedes or follows the numeral, the Gen marking is a way to indicate that it is in the scope of the numeral, and must thus be interpreted as such.

Some more recent proposals such as Nelson & Toivonen (2000) (Finnish, Inari Sami), Ionin & Matushansky (2006) (Russian), and Bailyn & Nevins (2008) (Russian), are all concerned with numeral systems in which the numeral marks its complement with case. Whether the numeral has the property of checking case or not sheds light on its status as either the head or the specifier. Nelson & Toivonen (2000: 184) assume head status for both numerals 2-6 checking Acc, and 7<, which check Partitive (Part) case, and also assume that the numerals head NumP (which is in accordance with my forthcoming proposal). Ionin & Matushansky (2006: 333) clearly state that the numerals must be heads, and while this is not such a problem with nominal numerals 5<, with the paucal adjectival ones 2-4 the analysis requires more work (Ionin & Matushansky 2006: 333, fn. 19). Importantly, we have a decision here that numerals represent one or the other category: they are either adjectives or nouns, they are not something in-between. A similar view is presented by Corver & Zwarts (2006), where the authors argue for a nominal status of the Dutch numerals. I will refer to these analyses in my proposals, where I will show that while Polish paucal numerals are unquestionably adjectival, the numerals 5< cannot be assigned to either adjectives or nouns, yet they check case on the counted noun and thus must be considered heads. This leads me to propose that Polish numerals 5< are lexicalisations of the functional Num head and are thus joining the ranks of functional heads such as T or v. I will also resolve the issue of paucal number and case (which is syncretic with Gen in Russian, and with Nom in Polish).

3.2 The numeral system puzzle: the Accusative Hypothesis

Polish linguists faced exactly the same problems with establishing the category of numerals as the abovementioned analyses popularized on an international basis. There exist three conflicting proposals with respect to which item constitutes the head of the numeral expression: (i) the head is the counted noun (Klemensiewicz 1937, Laskowski 1984, Mieczkowska 1994, Bobrowski 1988), (ii) the head is the numeral (Saloni & Świdziński 1987), and (iii) both the noun and the numeral are heads influencing sentential agreement in equal ways (Bogusławski 1973). There are good reasons for all of these views, and at the same time there is empirical evidence against every single one of them, i.e. with paucal numerals in all cases and with numerals 5< in non-structural cases we have congruence between the numeral and the noun, thus favouring (i); in structural cases, the congruence between the noun and the numeral is only available with paucal numerals, numerals 5< showing a case-mismatch (Nom/Acc numeral and Gen noun), thus providing evidence for (ii); with paucal numerals we have sentential agreement if in the subject position, but with numerals 5< we always have 3.SG.N, thus again pointing to (ii); combination of the two gives us (iii), but this hardly explains anything.

There is one more challenge presented by the Polish numeral system which also sheds light on the confusing approaches (i)-(iii) mentioned above. This puzzling property has been known under the name Accusative Hypothesis (Małecki 1863, Krasnowolski 1897, Szober

1928, 1953, and within the generative tradition Franks 1995, 2002, Przepiórkowski 1996, 2004, Rutkowski 2000, 2007, Rutkowski & Szczegot 2001). The Accusative Hypothesis proposes that Polish numerals 5< do not have Nom forms, i.e. they only have one (structural) case form: Acc. This Acc being structural is trumped by the more marked lexical cases and thus the differences disappear in Gen (which is the cut-off point, Franks 2002)¹²; however, the starting point for any numeral 5< (apart from those which are nouns) and its counted noun is: numeral_{ACC} noun_{GEN} (hence in all the glosses above, whenever I marked numerals Nom/Acc, they should be understood as exclusively Acc). The Acc case on the numerals 5< present also in Nom environments (e.g. the subject position) is an additional puzzling property of the Polish numeral system. This Acc case diachronically has its source in the Acc/Gen syncretism in virile nouns discussed in section 2 (where with non-virile nouns the Acc form is syncretic with Nom). Thus, what we observe in Polish should be presented as in Table 9 below (cf. Table 4); I further introduce agreeing demonstrative pronouns for the nouns to emphasize the case contrast, I also mark the relevant syncretisms in bold, and match it with the Accusative Hypothesis proposal by crossing out the Nom line (which is independently motivated by the ungrammaticality of the Nom demonstrative pronoun in V **ci*_{NOM} *pięciu*_{ACC} *mężczyzn*_{GEN} ‘these five men’, though fine in *ci*_{NOM} *mężczyźni*_{NOM} ‘these men’):

CASE	‘these five men’ (V)	‘these five women’(NV)	‘these five houses’ (NV)
NOM	*<i>ci</i> pięciu mężczyzn_{GEN}	<i>te</i> pięć kobiet_{GEN}	<i>te</i> pięć domów_{GEN}
ACC	<i>tych</i> pięciu mężczyzn_{GEN}	<i>te</i> pięć kobiet_{GEN}	<i>te</i> pięć domów_{GEN}
GEN	<i>tych</i> pięciu mężczyzn	<i>tych</i> pięciu kobiet	<i>tych</i> pięciu domów
DAT	<i>tym</i> pięciu mężczyznom	<i>tym</i> pięciu kobietom	<i>tym</i> pięciu domom
INSTR	<i>tymi</i> pięcioma mężczyznami	<i>tymi</i> pięcioma kobietami	<i>tymi</i> pięcioma domami
LOC	<i>tych</i> pięciu mężczyznach	<i>tych</i> pięciu kobietach	<i>tych</i> pięciu domach

Table 9 Case paradigm of *pięć* ‘five’ assuming the Accusative Hypothesis (Modern Polish)

Concord within the nominal projection includes all of its lexical subparts, hence demonstratives must agree with either the noun or the numeral. Judging by the paradigm in Table 9, it must be assumed that they agree with the numeral, or else the virile and non-virile paradigms would have a very different syntax (in non-virile we have *te* which can only be assumed to agree with the numeral here). The fact that the numeral does not agree in case with the counted noun is because it checks the noun’s case, and it cannot check this case on itself.

The case borne by the subject was considered the more problematic of the two issues by Polish grammarians. Schenker (1971) discusses the conflicting views of Polish grammarians on this topic. For instance, he quotes Klemensiewicz (1952) (Schenker’s translation): “Virile has no nominative whatsoever; the expressions *pięciu ludzi (przyszło)* ‘five people came, *pięciu panów (siedzi)* ‘five men are sitting’ are genitive-accusatives functioning as the subject.” In the same publication Klemensiewicz does not say what happens in non-virile, however in his later work (Klemensiewicz 1957), he identifies non-virile forms such as *pięć koszy stoi w piwnicy* ‘five baskets stand in the cellar’ as Nom (Schenker 1971: 55). Schenker sums up the varying judgments of Polish grammarians in a table similar to the one below. The contents speak for itself.

¹² The following hierarchy of case-markedness seems to be representative of the Polish case-system (I start from the least marked): Nom<Acc<Gen<Dat< Loc < Instr. This hierarchy is in agreement with that presented in Blake (1994: 89) and Caha (2009: 24) and is (mostly) supported by the existing case syncretisms. It follows then, that any case on the right of any other case will be considered here as the more marked case and will be expected to trump the structural cases (a prediction which is born out). This is also why Nom cannot trump Acc, but the opposite is possible.

	<i>pięciu studentów</i> (V) 'five (male) students'	<i>pięć studentek</i> (NV) 'five (female) students'
Grappin (1942)/Schabowska (1967)	Nom/Acc	Nom/Acc
Doroszewski (1952)	Gen	Nom
Klemensiewicz (1953)	Acc	Nom
Małecki (1863)/Szober (1928, 1953)/ Schenker (1971)	Acc	Acc

Table 10 Varying proposals for the case borne by the numeral expressions with numerals 5< (based on Schenker 1971: 56)

In what follows, I will argue with Schenker (among many others), that indeed what we see in Table 9 is the real state of the matters, i.e. that the Accusative Hypothesis is right and that it is also the only one that has any historical grounds (see the discussion on the introduced syncretisms above). The fact that Nom case does not figure on the numerals in subject positions will be explained via the case-inclusion as proposed by Caha (2009), where every more marked case includes all the less marked ones, to be more concrete: Acc includes Nom, and the case feature spelled out is always the highest, most complex one (which is why Nom cannot trump Acc). To account for the Accusative Hypothesis, and for the source of the Acc case on numerals in particular, I will argue for the existence of a null light preposition (in the sense of Franks 2002), which selects a lexicalised Num⁰, and which checks structural Acc case on the numeral expression. The analysis will be based on Pesetsky & Torrego's (2004) proposal that case is a reflection of tense, in particular that prepositions are instances of tense.

4 Analysis of the syntax of Polish numerals

4.1 Polish numerals 5< are not nouns

There are two recent proposals in the literature in which numerals are argued to be nouns: Ionin & Matushansky (2006), who propose such an analysis for the Russian numerals 5< (the paucal ones being adjectival in their analysis, too), and Corver & Zwarts (2006), who propose a similar analysis for the Dutch numerals. In what follows, I will show with reference to these proposals that Polish numerals 5< cannot be considered nouns.

Polish cardinals 5< differ from nouns in that the Gen they assign is structural not lexical (cf. section 2). Had we assumed that they are nouns nonetheless, they would immediately be exceptional in that category and would require an analysis making room for this exceptionality. The structural Gen they mark may be trumped by any other more marked case (see Jakobson 1984; Franks 1995, 2002; Caha 2009; Matushansky 2008, 2010).¹³ Since their syntactic behaviour is strictly connected to their quantifying properties, an analysis which groups them together with nouns seems rather forced. Also, if structural case is what distinguishes functional heads from lexical ones, the assumption that numerals are functional heads within the extended nominal projection goes through without further stipulations.

As discussed in Rutkowski (2007: 216-220), all Polish numerals (5< included) have nominal counterparts: *pięć* 'five' (numeral) vs. *piątka* 'a five' (noun). He provides an exhaustive discussion of why we should not group them together and consider them both nouns. Apart from the fact that it seems redundant to have two exactly synonymous nouns with exactly the same function within the same language, there is ample evidence from the syntactic behaviour of these two lexical items showing that the numeral should not be considered a noun (the data is from Rutkowski (2007) unless stated otherwise).

¹³ This applies to both Gen of negation and Gen of quantification as they both instantiate structural case.

- (17) a. Widzę pięć lingwistek.
 I-see five_{ACC} linguists_{GEN.F}
 'I can see five female linguists.'
 b. Widzę piątkę lingwistek
 I-see a-five_{ACC} linguists_{GEN.F}
 'I can see five female linguists.'
- (18) a. Pracuję z pięcioma lingwistkami/*lingwistek.
 I-work with five_{INSTR} linguists_{INSTR.F/*GEN}
 'I work with five female linguists.'
 b. Pracuję z piątką lingwistek/*lingwistkami.
 I-work with a-five_{INSTR} linguists_{GEN.F/*INSTR}
 'I work with five female linguists.'

The lexical Gen borne by the complement of the noun *piątka* remains unaffected by the lexical case of the preposition (18b), as opposed to the structural Gen checked by the numeral (18a).

Corver & Zwarts (2006: 813) show that Dutch numerals may have diminutive forms, a property typical of nominal items. This property can only be exhibited by the noun *piątka* but not by the numeral *pięć* in Polish.

- (19) a. Widzę piąteczkę lingwistek.
 I-see a-five(DIM)_{ACC} linguists_{GEN.F}
 'I can see five female linguists.'
 b. Pracuję z piąteczką lingwistek.
 I-work with a-five(DIM)_{INSTR} linguists_{GEN.F}
 'I work with five female linguists.'

Also, Polish numerals, unlike nouns, no longer have number distinctions, they have only one form. The noun *piątka*, however, comes in both singular (see above) and plural:

- (20) Demonstranci szli piątkami.
 demonstrators_{NOM} walked fives_{INSTR}
 'The demonstrators walked in groups of five.'

Numerals cannot be complements to other numerals, but nouns such as *piątka* can (my examples):

- (21) a. Dostałam pięć piątek.
 I-got five_{ACC} fives_{GEN}
 'I got five fives.' (a five = the second highest grade in the Polish school system)
 b. *Dostałam pięć pięciu.
 I-got five_{ACC} five_{GEN}

In the subject position, numeral expressions (as discussed above) trigger default 3.SG.N agreement, thus patterning with non-nominative and clausal subjects. Noun phrases headed by *piątka* can be Nom and trigger agreement which mirrors *piątka*'s ϕ -features, i.e. 3.SG/PL.F.

- (22) a. Piątka bandytów włamała się do banku.
 a-five_{NOM.F} robbers_{GEN} broke_{3.SG.F} self into bank

- ‘Five robbers broke into a bank.’
- b. Pięciu bandytów włamało się do banku.
 five_{ACC} robbers_{GEN} broke_{3.SG.N} self into bank
 ‘Five robbers broke into a bank.’

Finally, nouns like *piątka*, but crucially not cardinals like *pięć*, can be used with possessive pronouns (where the possessive is semantically connected to the numeral):

- (23) a. Nasza piątka jedzie dziś na urlop.
 our_{NOM} five_{NOM} goes today on holiday
 ‘The five of us are going on holiday today.’
- b. *Nasze pięć jedzie dziś na urlop.
 our five goes today on holiday
 ‘The five of us are going on holiday today.’

4.2 Proposal 1: Polish numerals 5< are functional Num heads, 1-4 are specifiers of NumP

I propose Polish numerals 5< to be heads of NumP and the numerals 1-4 to be specifiers of this projection. NumP is a functional projection introduced between nP/NP and DP. This kind of proposal is certainly not novel (Bernstein 1993; Przepiórkowski 1996; Giusti 1991; Zamparelli 2008; Rutkowski 2007), but not uncontroversial either (Selkirk 1977; Franks 1994; Ionin & Matushansky 2004, 2006).

There is no doubt that numerals 1-4 are adjectives: they have always been adjectival and have remained so throughout the history of Polish (cf. Comrie 1992 for an overview of the paucal adjectival numerals in Slavonic). Numerals 5<, on the other hand, have changed their status and stopped being nouns (although they are still more ‘nouny’ than adjectival (cf. Table 8)). I will thus concentrate on these numerals and propose what category they really represent, i.e. I propose that they have been reanalysed as functional heads and must be understood as lexicalisations of Num⁰. An analysis of this kind would explain their (partly) nominal behaviour in the sense that they are part of the extended projection of the noun (Grimshaw 1991, 1997, Van Riemsdijk 1998), which is also why they exhibit some agreement with the counted noun. It also accounts for why they have lost their own ϕ -features and why diachronically they were the category of choice to exhibit new gender distinctions in the plural¹⁴, and why the emergence of new genders coincided with the emergence of numerals as a separate grammatical category (Siuciak 2008). Finally, it explains their head-like behaviour with respect to case-checking and puts them as representants of Num within the ranks of other functional categories such as T or *v*.

In effect, I will distinguish between lexicalised Num⁰ checking Gen on the counted noun and forcing plural agreement, non-lexicalised paucal Num⁰ licensing paucal numerals 2-4 in its specifier (which in Polish also results in plural agreement), non-lexicalised Num⁰ with features [\pm plural].

(24) Number Phrase featural content

- | | | |
|---|---|-----------------------|
| a. Num ⁰ [+Q] [+plural] | - numerals 5< | } numeral expressions |
| b. Num ⁰ [+Q] [+paucal/plural] | - numerals 2-4 | |
| c. Num ⁰ [\pm plural] | - unquantified noun phrases (also with 1) | |

¹⁴ If we follow analyses postulating that gender is parasitic on number such as Ritter 1993, Di Domenico 1999, De Vicenzi & Di Domenico 1999, Alexiadou 2004, Müller 2000, which also follow from Greenberg’s (1963: 74) Universal No. 36

Both numeral (quantified) Num heads are assumed to license case: Gen or Paucal. In Polish, Paucal is syncretic with Nom, but in Russian it is either syncretic with Gen (hence ‘paucal genitive’, Pereltsvaig 2008) or is a case in its own right (Bailyn & Nevins 2008¹⁵). Also, Nelson & Toivonen (2000) show the different case marking strategies of lower numerals and higher numerals in Inari Sami: Acc checked by 2-6, and Part checked by 7<. These distinctions will be further supported by the selectional requirements of prepositions (e.g. distributive *po*).

The syntax of Polish 2-4 and 5< differs considerably (perhaps more so than in Russian, Ionin & Matushansky 2006), yet it coincides within the same projection, NumP. NumP may be headed by two types of (quantifying) heads: (i) a null head licensing a specifier (with adjectival paucal numerals 2-4 as specifier), checking Paucal case on the complement, (ii) a lexicalised Num head (by numerals 5<), which does not license a specifier and checks Gen case on the complement.¹⁶ Detailed analysis of the diachronic and present day material also points in that direction. Although the two groups of numerals correspond to different grammatical categories, the fact that they nowadays make use of the same projection (NumP) (although never at the same time) has been made possible by these very changes. Thanks to the numeralisation of 5<, numerals as a class group within NumP.

4.2.1 Numeralisation of numeral nouns and NumP

In this section I show how the process of numeralisation of Polish numerals proceeds and propose the general structure of the Number Phrase. I will exemplify the process of numeralisation with *tysiąc* ‘thousand’ currently undergoing it.

It follows from Tables 8 and 9 that *tysiąc* has retained its nominal nature. Yet, there are constructions in which this status may be questioned. Following the analysis presented in Przepiórkowski (2006, 2008), who shows extensive evidence to the effect that Polish distributive *po* is sensitive to the presence of numeral expressions when selecting a complement, in particular it marks numeral expressions with Acc, and nominal ones with Loc, I will show that *tysiąc* may take on either of these cases under certain conditions, i.e. when used on its own it will be treated as a noun and thus in Loc, and when used as part of a complex numeral it will be treated by *po* as a numeral and thus in Acc.

Distributive *po* is generally infelicitous with plural complements, unless they are numeral expressions (Franks 1995, Przepiórkowski 2006, 2008).¹⁷

- (25) a. Piotr dał im po jabłku.
 Peter_{NOM} gave_{3.SG.M} them_{DAT} *po* apple_{LOC}
 ‘Peter gave them an apple each.’
 b. Piotr dał im po dwa jabłka /*dwóch jabłkach.
 Peter_{NOM} gave_{3.SG.M} them_{DAT} *po* two_{ACC} apples_{ACC}/*two_{LOC} apples_{LOC}
 ‘Peter gave then two apples each.’
 c. Piotr dał im po pięć jabłek /*pięciu jabłkach.
 Peter_{NOM} gave_{3.SG.M} them_{DAT} *po* five_{ACC} apples_{ACC}/five_{LOC} apples_{ACC}

¹⁵ This type of analysis of Paucal is also hinted at by Ionin & Matushansky (2006: 333, footnote 18), though not developed in detail.

¹⁶ Both situations have their counterparts in other functional categories like *v* (assumed to check case only if it also licenses an external argument in its specifier (this is not the case in Pesetsky & Torrego’s (2004) analysis)), or *C* – the Douby Filled Comp Filter (either the specifier or the head is lexicalised), the licensing of a specifier by *T*, although the case-checking in a probe-goal system happens under Agree (hence the specifier need not be filled by the item against which the case is checked).

¹⁷ Though see Przepiórkowski 2006, 2008 for examples such as *jedne usta* ‘one (set of) lips, where *po* is fine with the plural of ‘one’ and pluralia tantum nouns.

‘Peter gave them five apples each.’

Consider the following examples with *tysiąc*:

- (26) a. Ojciec dał nam *po tysiąc /tysiącu.
 father_{NOM} gave_{3.SG.M} us_{DAT} po *thousand_{ACC}/thousand_{LOC}
 ‘Father gave us a thousand each.’
- b. Ojciec dał nam po tysiąc /tysiącu złotych.
 father_{NOM} gave_{3.SG.M} us_{DAT} po thousand_{ACC}/thousand_{LOC} zlotys_{GEN}
 ‘Father gave us a thousand zlotys each.’
- c. Ojciec dał nam po tysiąc /tysiącu dwadzieścia pięć złotych.
 father_{NOM} gave_{3.SG.M} us_{DAT} po thousand_{ACC}/thousand_{LOC} zlotys_{GEN} five_{ACC} zl._{GEN}
 ‘Father gave us a thousand and twenty-five zlotys each.’

Recall that *po* marks its nominal complements with Loc and numeral ones with Acc. In (26a) *tysiąc* must thus be analysed as nominal, judging by the ungrammaticality of its Acc form; however, in (26b, c) we have two possibilities, one in which *tysiąc* is treated as a numeral and is Acc. It still assigns lexical Gen to its complement *złoty*; however, the fact that it does not receive Loc tells us that it is not perceived as nominal by *po*. What happens during the process of numeralisation is thus a reanalysis of the lexical material (here the nominal head) as functional (Roberts & Roussou 2003: 159). In this particular case, this is achieved via merge of the noun in its own Num⁰ (if the process is complete, the noun should also lose its own number and gender distinctions, as it happened earlier with for instance *sto* ‘hundred’ whose nominal and numeral paradigms were presented in Table 5). The empty nominal projection in the complement of Num⁰ becomes available for merge of the counted noun. This process must be understood as grammaticalisation through reanalysis. In the case of *tysiąc* the process is not yet complete and the noun may still be merged either as Num, in which case it takes a nominal complement, or as N. This sheds light on how distributive *po* distinguishes between the nominal and the numeral *tysiąc*: they are different heads with different properties. With numerals for which the process is completed (5, 20, 100, etc.), there is no such choice, they are always merged in Num and only with pronouns their complement position may be unfilled (cf. (27c, d)). The resulting structure is highly reminiscent of that proposed by Panagiotidis (2002, 2003): just like there is an empty noun within each pronoun, there is also an empty noun within each numeral. Further support for such an analysis actually comes from the distribution of pronouns with numerals, which without exception follow personal pronouns.

- (27) a. nominal *tysiąc* b. numeral *tysiąc* c. numeral *tysiąc*
- DP
├── NumP
│ ├── Num
│ └── NP
│ └── \triangle
│ N
│ tysiąc

DP
├── NumP
│ ├── Num
│ │ tysiąc
│ └── NP
│ └── \triangle
│ e

DP
├── NumP
│ ├── Num
│ │ tysiąc
│ └── NP
│ └── \triangle
│ złotych
- d. nas pięciu /nas pięciu chłopców
 us_{ACC} five_{ACC}/us_{ACC} five_{ACC} boys_{GEN}
 ‘us five/us five boys’
- DP
├── D
│ nas
└── NumP
 ├── Num
 │ pięciu
 └── NP
 └── \triangle
 e/chłopców

Though the behaviour of *tysiąc* is highly reminiscent of measure nouns, measure nouns such as *godzina* ‘hour’ always receive Loc from distributive *po* (28), therefore we see that it is *tysiąc*, but not *godzina*, which is undergoing numeralisation.^{18,19}

- (28) a. Studenci mają po godzinie trzydzieści pięć minut na
 students_{PL,NOM} have_{3,PL} po hour_{LOC} thirty_{ACC} five_{ACC} minutes_{GEN} for
 każde zadanie.
 each task
 ‘Students have one hour thirty-five minutes for each task.’
 b. *Studenci mają po godzinie trzydziestu pięciu minutach na
 students_{PL,NOM} have_{3,PL} po hour_{LOC} thirty_{LOC} five_{LOC} minutes_{LOC} for
 each task
 każde zadanie.

The numeralisation of *tysiąc* also provides evidence for the quantified status of numeral expressions with paucal numerals and supports the proposals made above, i.e. even though 2-4 are adjectival, the Num head licensing them in its specifier must be [+Q] or else we would expect them to be marked Loc. So (29a,b,c) must be understood as numeral expressions in which *tysiąc* is part of a complex numeral, and must thus be analysed exclusively as numeral, and not nominal.

- (29) a. Ojciec dał nam po dwa tysiące/*dwóch tysiącach.
 father_{NOM} gave_{3,SG,M} us_{DAT} po two_{ACC} thousand_{ACC}/*two_{LOC} thousand_{LOC}
 ‘Father gave us two thousand each.’
 b. Ojciec dał nam po dwa tysiące /*dwóch tysiącach złotych.
 father_{NOM} gave_{3,SG,M} us_{DAT} po two_{ACC} thousand_{ACC}/*two_{LOC} thousand_{LOC} zlotys_{GEN}
 ‘Father gave us two thousand zlotys each.’
 c. Ojciec dał nam po dwa tysiące pięćset /*dwóch
 father_{NOM} gave_{3,SG,M} us_{DAT} po two_{ACC} thousand_{ACC} five-hundred_{ACC}/*two_{LOC}
 tysiącach pięćset złotych
 thousand_{LOC} five_{LOC} hundred_{LOC} zlotys_{GEN}
 ‘Father gave us two thousand five hundred zlotys each.’

Interestingly, in example (26c) given earlier, in which *tysiąc* itself is not multiplied, it can still be treated as a noun.

¹⁸ In (28b) we have an instance of *po* which would mark Loc on all the components, however, it is not the distributive *po*, but a preposition which means ‘after’, hence the interpretation of (28b) would be ‘after one hour and thirty five minutes, and not the intended one: ‘one hour and thirty five minutes each’.

¹⁹ See also Stavrou (2003) for a proposal of a measure phrase, a hybrid lexical-functional category in Van Riemsdijk’s (1998) terms. It is plausible that the numeralisation process could include an MP at some early stage, and that perhaps Polish measure nouns are actually occupying an M (head) position. In particular, nouns such as *masa* ‘mass’, which used in Acc *masę* (like numerals 5<) can be analysed as vague numerals (the fact that *po* will not select *masa* is probably due to the fact that it is a mass noun and thus does not fit in with the distributive meaning). If MP is in some way akin to NumP, then this would explain why approximative prepositions such as *z/około* ‘about/around’ co-occur with it (or select it) like they do with numeral expressions. This issue, however, will have to be left to future research and I do not commit here to any particular analysis of partitives or pseudo-partitives.

4.2.2 Constructing complex numerals

In this section, I propose a structure for complex numerals and thus a few words about the numeral system itself are in order. Polish uses a strictly decimal system with ‘ten’, ‘hundred’, ‘thousand’ and ‘million’ as bases. These are used in both multiplication and addition, and thus teens are formed by addition of simplex numerals 1-9 to 10, and the same is replicated in decades (combination of multiplication of 10 and addition of 1-9); decades, hundreds and thousands are formed via multiplication of bases by atomic simplex numerals 1-9 and 10.

Many complex numerals started off as separate lexical items, to be eventually spelled out as one. Conjunctions were present in teens (*na*)²⁰ and decades (*i/a*). The Middle Polish lexicalised conjunctions *i/a* ‘and’ (30) have been replaced by asyndetic conjunctions (Ionin & Matushansky 2006: 339; also referred to as juxtaposition by Hurford (2003)), as shown in the Modern Polish counterparts of (30) under (31).

- (30) a. do czterech i do dwudziestu godzin (Glaber 1535)
to four_{GEN} and to twenty_{GEN} hours_{GEN}
‘up to twenty-four hours’
b. we dwudziestu a w pięci lat (Falimirz 1534).
in twenty_{ACC} and in five_{ACC} years_{GEN}
‘in twenty-five years’
c. w stu i trzydziestu koni (Gwagnin 1611).
in hundred_{ACC} and thirty_{ACC} horses_{GEN}
‘with a hundred and thirty horses’
- (31) a. do dwudziestu czterech godzin (Modern Polish)
to two_{NV.GEN}-ten_{NV.GEN} four_{NV.GEN} hours_{NV.GEN}
‘up to twenty-four hours’
b. w dwadzieścia pięć lat
in two_{NV.NOM/ACC}-ten_{NV.NOM/ACC} five_{NV.NOM/ACC} years_{NV.GEN}
‘in twenty-five years’
c. w sto trzydzieści koni
in hundred_{NV.NOM/ACC} thirty_{NV.NOM/ACC} horses_{NV.GEN}
‘with a hundred and thirty horses’

Hundreds are good examples of multiplications. To start with, both lexical items were not only spelled out separately, but also each bore its own inflectional endings, e.g. ‘two hundred’ can be shown to have taken on the following forms before finally synthesising²¹ (based on Grappin 1950: 48, Siuciak 2008: 34, 120-121):

- (32) a. before synthesising
Nom/Acc *dwie ście*

²⁰ One very good example of this is the development of Polish teens, which I present here on the example of *jedenaście* ‘eleven’. *Jedenaście* was originally represented by three separate lexical items in Proto-Slavic: *jedinъ* ‘one’, *desęte* ‘ten’, and the connective preposition *na* ‘on’ (Greenberg’s (1978: 265) ‘superessive link’), here used as a linguistic counterpart to the arithmetic addition: *jedinъ na desęte* > *jeden na desęte* > *jeden nadzieście* > *jedennaście* > *jedenaście* (one on ten >...> eleven; based on Comrie 1992: 766).

²¹ The process began in 17th c. and finalised in 19th c., although not completely. In some cases, inflection moved onto the last conjunct, cf. *trzysta/trzystu* ‘300’ vs. *pięćset/pięciuset* ‘500’. Notice that the part –sta/-stu in ‘300’ is plural in itself and bears the relevant cases (just like with paucal numerals), whereas the part –set in ‘500’ is always Gen plural, again showing consistent behaviour with all numerals 5< (cf. Table 5). Thus ‘100’ retains its nominal (not numeral) declension when multiplied, hence we do not have forms like **trzysto* or **pięćstu*.

Gen	<i>dwu/dwoch/dwuch set</i>
Dat	<i>dwiema/dwoma stoma</i>
Instr	<i>dwiema sty/dwoma set/dwoma stoma</i>
Loc	<i>dwoch set/dwóch stach</i>

(b) synthesised forms

Nom/Acc	<i>dwieście</i>	
Gen	<i>dwóchset/dwuset</i>	} (eventually <i>dwustu</i>)
Dat	<i>dwomset/dwumstom/dwiestom/ dwiestu</i>	
Instr	<i>dwiemasty/dwustoma/dwiestoma</i>	
Loc	<i>dwuset/dwochset/dwóchset</i>	

Even though spelled together, it seems that synthesisation of these forms is not complete, as each has kept on its own inflectional suffixes.²²

Following Ionin & Matushansky's (2006) semantic proposal, I will also assume that multiplication involves selection and thus exhibits hierarchical structure, whereas addition involves conjunctions (Greenberg 1978, Hurford 1987, 2003, Von Mengden 2010). I will also incorporate Ionin & Matushansky's proposal involving Right Node Raising (RNR) (Ross 1967, Postal 1974, Bresnan 1974, Abbott 1976, Sabbagh 2007) in complex numerals and show on the basis of the visible case-marking that it is the only viable proposal for Polish. The RNR analysis also allows for an explanation of the availability of Left Branch Extractions (hence LBE) of the material preceding the noun (including the preposition) which is available in Polish:

- (33) a. Ile spędziłaś tam [ile] lat?
how-many spent_{2.SG.F} there years_{GEN}
'How many years have you spent there?'
b. Około trzydziestu pięciu spędziłam tam [~~około trzydziestu pięciu~~] lat.
about thirty_{GEN} five_{GEN} spent_{1.SG.F} there years_{GEN}
'I've spent there about thirty-five years.'

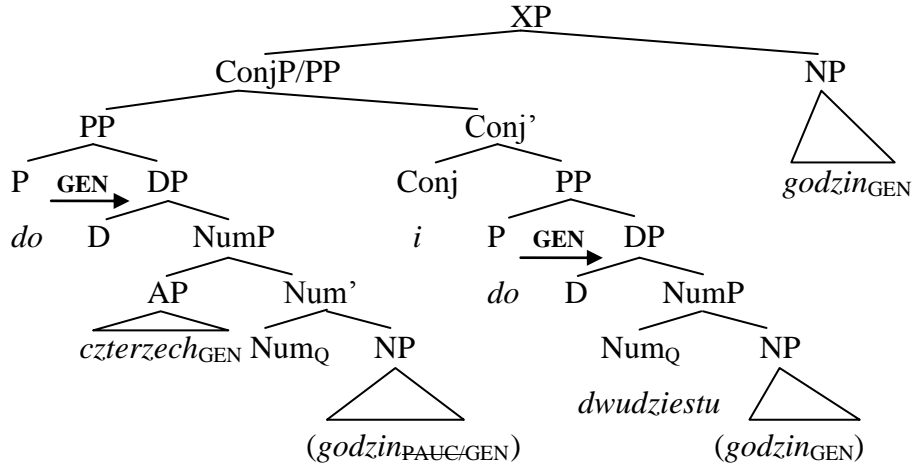
Conversely, the fact that prepositions undergo LBE together with numerals suggests that they must be merged before RNR applies rather than after, so that they can be part of the left branch (this must also be true of simple numerals). We can see in the data from 16th c. Polish that prepositions are present within every conjunct (30), and what we see as lexicalised material in 16th c. Polish must still be at least syntactically present in Modern Polish. Also, the previously given examples with *godzina* in (28), where the distributive *po* marks the two conjuncts with different cases, provide further support for the presence of separate prepositional phrases in which *po* selects and case-marks its respective nominal (Loc) and numeral (Acc) complements (hence *Studenci mieli po godzinie_{LOC} (i po) trzydzieści_{ACC} pięć_{LOC} minut_{GEN} na każde zadanie*). The proposal presented here allows for all this to happen within a single projection.

In (34a) I present what the structure of the 16th c. Polish example (30a) would look like in my proposal. In (34b), I present its Modern Polish counterpart ((31a) above) keeping the conjunction and putting the silent preposition in parentheses; in (35) I give a structure of a (made-up) complex numeral: '5028 horses'. The structures given below are highly

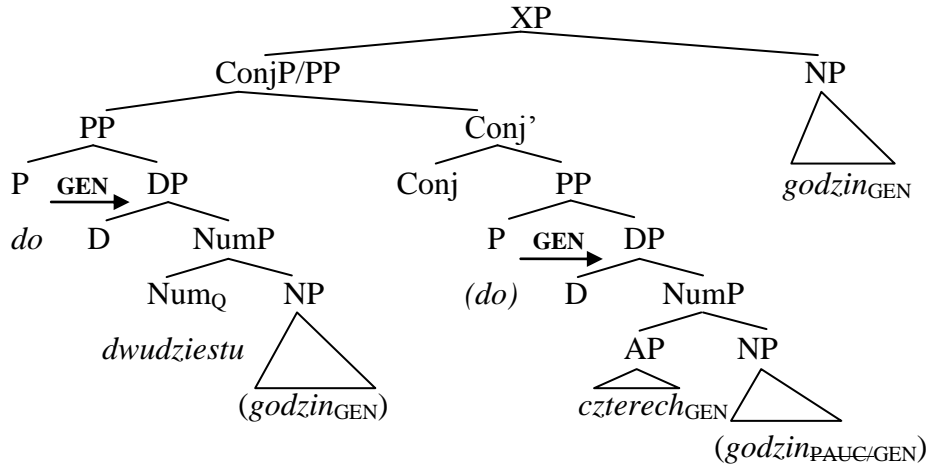
²² This is visible to this day in all combinations with *dwa* 'two' (*dwanaście_{NOM}* vs. *dwunastu_{GEN}* 'twelve'; *dwadzieścia_{NOM}* vs. *dwudziestu_{GEN}* 'twenty'; *dwieście_{NOM}* vs. *dwustu_{GEN}* 'two hundred'), as well as multiplications of hundred in 500-900 (*pięćset_{NOM}* vs. *pięciuset_{GEN}* vs. *pięciomaset_{INSTR}* 'five hundred') reminding us of their once separate past.

reminiscent of those proposed in Ionin & Matushansky (2006: 340), with the difference that I introduce the relevant labeling in accordance with my proposal.

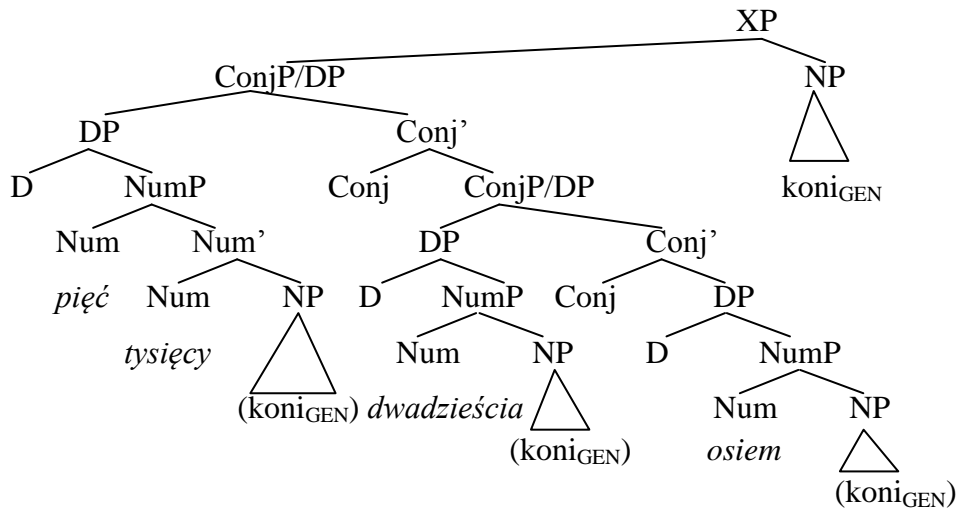
(34) a.



b.



(35)



(36) a. pięć tysięcy dwadzieścia osiem koni (5028)

five_{ACC} thousand_{GEN} twenty_{ACC} eight_{ACC} horses_{GEN}
 ‘five thousand and twenty-eight horses’

b. pięć tysięcy dwadzieścia jeden koni (5021)

five_{ACC} thousand_{GEN} twenty_{ACC} one? horses_{GEN}

- ‘five thousand and twenty-one horses’
 c. pięć tysięcy dwadzieścia dwa konie (5022)
 five_{ACC} thousand_{GEN} twenty_{ACC} two_{ACC} horses_{PAUC/NOM}
 ‘five thousand and twenty-one horses’

- (37) z pięcioma tysiącami dwudziestoma jeden/dwoma/ośmioma końmi
 with five_{INSTR} thousand_{INSTR} twenty_{INSTR} one_?/two_{INSTR}/eight_{INSTR} horses_{INSTR}
 ‘with five thousand and twenty-one/two/eight horses’

An earlier example (29c), with a structure (38) below, also supports the conjunction and RNR view of the complex numeral structure.

- (38) a. numeral *tysiąc*
-
- b. nominal *tysiąc*
-

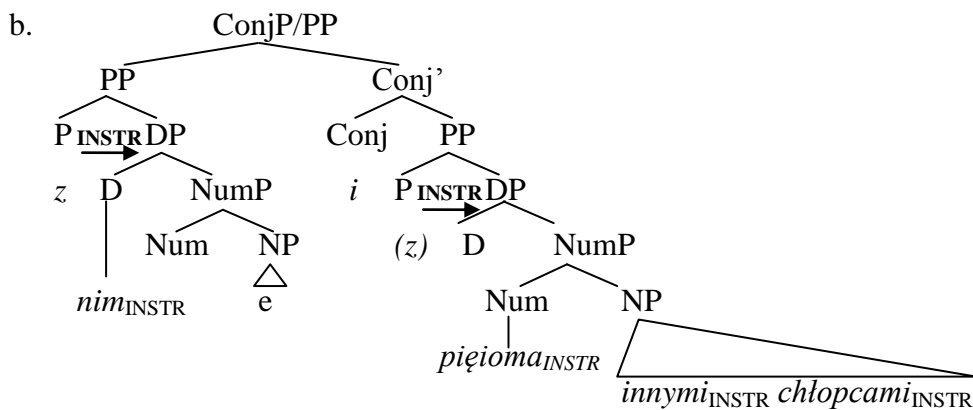
Recall that in (29c)/(38) *tysiąc* can be either nominal or numeral, which affects what counts as the complement of the numeral and when RNR applies. This kind of analysis is further supported by the constructions with LBE (which in turn support the RNR proposal):

- (39) a.* [Po tysiącu dwadzieścia pięć]_i ojciec dał nam _{t_i} złotych.
 po thousand_{LOC} twenty_{ACC} five_{ACC} ather_{NOM} gave_{3.SG.M} us_{DAT} zlotys_{GEN}
 ‘Father gave us one thousand twenty-five zlotys each.’

- b. [Po tysiąc dwadzieścia pięć]_i ojciec dał nam t_i złotych.
 po thousand_{ACC} twenty_{ACC} five_{ACC} father_{NOM} gave_{3.SG.M} us_{DAT} złoty_{SGEN}
 ‘Father gave us one thousand twenty-five zlotys each.’

Examples (40) support an analysis of pronouns proposed in (27): we have two conjuncts (in this case counting two different nouns: *e* in the case of the first one and *chłopaki* ‘boys’ in the case of the second one, so no RNR will be involved), and the grammaticality of this conjunction suggests that we are dealing with the same type of constituents (the preposition in the second conjunct may or may not be pronounced):

- (40) a. Maria spotykała się [z nim i pięcioma innymi chłopakami].
 Mary_{NOM} met_{3.SG.F.IMP} self with him_{INSTR} and five_{INSTR} other_{INSTR} boys_{INSTR}
 ‘Mary was dating him and five other boys.’



4.3 Accounting for the Accusative Hypothesis

In this section, I would like to propose an account of the Accusative Hypothesis briefly discussed earlier in section 3.2. In particular, I will identify the source of the Acc case, and I will propose that it is represented by a light preposition – a defective tense head in the sense of Pesetsky & Torrego’s (2004) proposal – one that checks case (uT on D) but cannot mark it for deletion. This is the first such attempt, as there does not exist an analysis of the Accusative Hypothesis in the literature that goes beyond the statement of the issue (Małeckie 1863, Krasnowolski 1897, Szober 1928, 1953, Franks 1995, 2002, Przepiórkowski 1996, 2004, Rutkowski 2000, 2007, Rutkowski & Szczegot 2001). It seems that linguists have always been rather thrown by the Accusative Hypothesis, which one can deduce from the following statement made by Steven Franks (2002: 142) about Czech: “My point of departure is the following outrageous claim: the subject noun phrase *těch pět hezkých dívek* ‘these five pretty girls’ in Czech [...] is not nominative, but *accusative* [...] My goal in what follows is to lay out the reasoning which led to this bizarre conclusion.” Nevertheless, Franks does not provide an explanation for where exactly this Acc case comes from, because the above claim *constitutes* his explanation of other facts he concentrates upon.

Before I move on to the analysis proper, a word or two are in order about the evidence pointing to the accusative nature of Polish numeral expressions. In section 3.2, I have already presented the relevant case paradigms (cf. Table 9). Here, I will present evidence given by Przepiórkowski (2004), where he concentrates upon numeral expressions as subjects and provides an analysis for them similar to that proposed for non-nominative subjects in

Icelandic.²³ Crucially, by proving that they behave like non-nominative subjects, Przepiórkowski proves that they cannot be Nom. As already stated earlier, I assume the Accusative Hypothesis to be right, i.e. a fact about Polish, but I also intend to go beyond the mere statement of this fact and account for the source of the case. I will show how the combination of Pesetsky & Torrego's (2004) analysis and the Accusative Hypothesis leads to the assumptions I make and in a way predicts such an analysis of Polish numeral expressions and additionally provides an analysis for Franks' (2002) Serbo-Croatian data. Below, I present Przepiórkowski's (2004) compelling evidence.

4.3.1. Accusative cardinal numerals 5<

Przepiórkowski's (2004) work constitutes a thorough analysis and discussion of the Accusative Hypothesis, which itself boasts a long linguistic tradition. Apart from presenting ample evidence in favour of the Accusative Hypothesis, Przepiórkowski (2004) also proves that when in a subject position, the numeral expressions must be considered subjects, concretely non-nominative subjects. The evidence he presents is crucial to the analysis presented here, as it hinges on the Acc case of numeral expressions.

As presented in Table 9 (section 3.2), demonstrative pronouns co-occurring with nouns are adjectival in nature and they agree with the noun they modify in ϕ -features, including case. In the case of numeral expressions, the demonstrative may agree either with the numeral or the quantified noun (see the discussion of Babby's discontinuous agreement above). With non-virile nouns we witness both possibilities as shown in (41a), i.e. we may have both a Gen form agreeing with the Gen quantified noun, and a Nom/Acc form agreeing with the numeral (*te* 'these_{NV}' is a Nom/Acc syncretic form); with virile nouns, however, only one form is available – *tych* 'these' – an Acc/Gen syncretic form. Crucially, no Nom form is available (**ci mężczyźn*). Since *te* 'these' may be either Nom or Acc and *tych* 'these' is either Acc or Gen, their common denominator is Acc. It thus seems plausible that the feminine *te* 'these' should be considered Acc, rather than Nom, or else we would have a very different syntax in the two genders in the plural. Additional evidence comes from subject-verb agreement in sentences with numeral phrase subjects: like all non-nominative subjects they trigger default 3.SG.N agreement (Przepiórkowski 2004:134).

- (41) a. Tych /te pięć kobiet /okien /kotów stało.
 these_{GEN,NV}/these_{NOM,ACC,NV} five_{NOM,ACC} women_{GEN}/windows_{GEN}/cats_{GEN} stood_{3.SG,N}
 'These five women/windows/cats stood.'
 b. Tych /*ci pięciu mężczyzn stało.
 these_{GEN,ACC,V}/these_{NOM,V} five_{ACC} men_{GEN} stood_{3.SG,N}
 'These five men stood.'

Further evidence comes from the presence of Acc/Gen adjectival predicate agreement (participial agreement), which I have already briefly mentioned in section 2.1 (examples (3)), but give here Przepiórkowski's (2004: 135) examples; since Przepiórkowski does not elaborate on how this state of the matters comes to be, I propose my own analysis.

- (42) a. Sześć samolotów zostało zakupione/zakupionych we wrześniu.
 six_{ACC} planes_{GEN} stayed_{3.SG,N} bought_{ACC/GEN} in September

²³ Przepiórkowski's (2004) publication was criticised by Saloni (2005) and later defended by Author & Co-Author (2007). The criticism concerned the proposal that Polish has non-nominative subjects. Traditional Polish grammars (and grammarians) assume sentences with Dative or Accusative subjects as subjectless impersonal constructions.

- ‘Six planes were bought in September.’
- b. Sześciu niewolników zostało *zakupieni/zakupionych w 1768r.
 six_{ACC} slaves_{GEN} stayed_{3.SG.N} bought*_{ACC/GEN} in 1768
 ‘Six slaves were bought in 1768.’

Noting that such agreement is only possible in the passive, I have proposed earlier that such agreement should be treated in terms of adjectival (participial) agreement as proposed by Chomsky (1999), i.e. during the cyclic movement of the object to the subject position, I assume that it must have moved through the relevant escape hatches so that at some point the participle finds itself in the scope of the numeral, and thus receives Gen case. This analysis, however, would leave unexplained the Acc case concord on the participle in (42a) (*sześć samolotów zostało zakupione*). There is a good reason to believe that this type of concord here is actually an error. We can see that only the form *sześć* (42a), but never the form *sześciu* (42b) can trigger such concord, even though both *sześć* and *sześciu* are Acc forms, just in different genders. Also, it appears that it is actually singular rather than plural concord, hence the gender would be neuter, rather than non-virile (although unfortunately it is not morphologically distinguishable on the participle, and we can only infer from the other lexical items which of them the participle agrees with). The fact that such concord may occur follows from the fact that *sześć*, apart from being a numeral, also functions as a neuter noun in definitions, or in sentences where we simply name or refer to the number 6. The most telling example is (43c) where I also use the passive (imagine we are talking about calligraphy). Notice the forms of the participles, which here can only be analysed as agreeing with the singular neuter noun *sześć*.

- (43) a. Sześć jest (zawsze było) liczbą parzystą.
 six_{SG.N} is (always was_{3.SG.N}) number_{F.INSTR} even_{F.INSTR}
 ‘Six is an even number.’
- b. Sześć jest/było moim szczęśliwym numerem.
 six_{SG.N} is/was_{3.SG.N} my_{M.INSTR} lucky_{M.INSTR} number_{M.INSTR}
 ‘Six is/was my lucky number.’
- c. Twoje sześć jest krzywo napisane, moje sześć zostało uznane za
 your_{SG.N} six_{SG.N} is crooked written_N, my_{SG.N} six_{SG.N} stayed_{3.SG.N} claimed_N for
 najładniej napisane.
 nicest written_N
 ‘Your six [written number] is crooked, my six was deemed as written the nicest.’

What we see in (43) – (43c) in particular – is exactly the same as what happened in (42a), but syntactically only the Gen version should be derivable, hence my claim that in (42a) the neuter Acc agreement on the participle is erroneous as it overgeneralises the properties of the noun *sześć* onto the homophonous numeral *sześć*.

Another piece of evidence in favour of the Accusative Hypothesis comes from the diachronic analysis of numeralisation. Only Acc forms of the nouns that undergo this process – *trochę*_{ACC} ‘a bit’/*masę*_{ACC} ‘mass’ – become numerals (Przepiórkowski 2004: 136):

- (44) Zostało trochę wody.
 left_{3.SG.N} bit_{ACC} water_{GEN}
 ‘There was left a bit of water.’

- (45) a. Masa ludzi przysła/*przyszło.
 mass_{NOM.F} people_{GEN} came_{3.SG.F/?*3.SG.N}

- ‘(lit.) A mass (=plenty) of people came.’
 b. Masę ludzi przyszło/ *przyszła.
 mass_{ACC} people_{GEN} came_{3.SG.N/*3.SG.F}
 ‘(lit.) A mass (=plenty) of people came.’

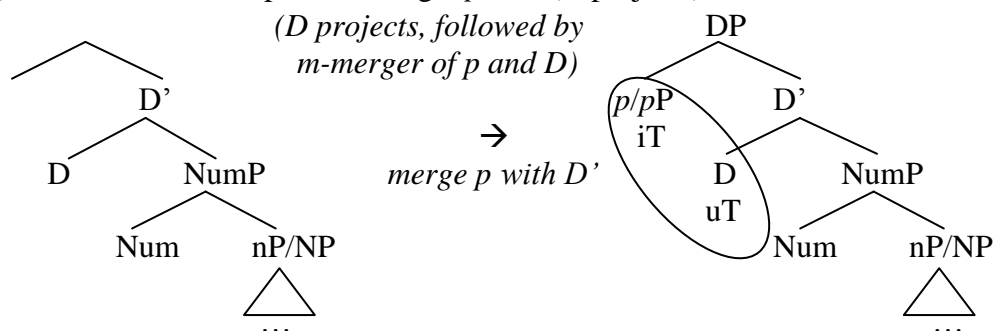
In (44) there is an Acc form of the no longer existing noun *trocha* ‘a bit’. Its Acc form survived as a numeral. In my view of numeralisation (cf. 4.2.1), this would mean that the noun *masa* was reanalysed as a numeral, and instead of being merged in NP, it is now merged in Num and is marked Acc just like all the other numerals lexicalising Num (i.e. 5<). Thus the Acc *trochę* in effect survived as the lexicalisation of the functional head Num. Collective nouns such as *grupa* ‘group’ and measure nouns such as *godzina* ‘hour’, as well as the Nom *masa* exemplified in (45a), still retain their nominal status, but I tentatively propose that they are in M(easure)P, above NP (cf. Stavrou 2003). Nevertheless, by prepositions such as distributive *po*, all of these (*grupa*, *godzina* and *masa*) are treated strictly as nominal and receive Loc accordingly.

Polish numeral expressions are thus to be analysed as Acc phrases. As we can see from the examples, the numerals have an ability to check Gen case, which is either structural (cardinals 5<) or lexical (*masę* ‘mass’). The structural Gen of quantification can be argued to be nothing more than Acc under quantification (just like it is the case with Gen of negation which also only affects structural Acc).²⁴

4.3.2 Proposal 2: a light preposition as the source of Acc case in Polish numerals

Following an analysis proposed by Pesetsky & Torrego (2004) (hence P&T 2004) in which case is a reflection of tense, and where prepositions are also assumed to be instances of tense, I would like to make the following proposal: Polish numeral expressions are introduced by a light preposition (*p*) which constitutes the source of their Acc case. It is light in the sense that it is transparent to the outside case-checkers and theta-role assigners (Franks 2002). This *p* is merged with D(P) where it is inserted at the root in accordance with the Extension Condition (Chomsky 1995). If it were to project, the phrase marker would result in a PP; instead, I propose that it does not project, but becomes the specifier of DP. This specifier being both maximal and minimal (p^0/pP), undergoes m(orphological)-merger with D (as in Matushansky’s (2006: 86) analysis of Saxon genitives). Because it is D that projects, we continue to deal with a nominal expression which may be further selected by other tense heads.

(46) Polish numeral expressions: light *p*+DP (D projects)



²⁴ A highly interesting way of dealing with such double case-marking is presented in Matushansky (2008, 2010).

The light *p* constitutes i(nterpretable)T in the spirit of P&T's analysis and its relation with u(ninterpretable)T on D results in Acc case. The transparency to the outside case-checkers can be translated here into this preposition being an instance of defective T, which cannot mark the uT on D for deletion. A similar proposal is made by P&T for unaccusative verbs, which are assumed to have a defective ϕ -incomplete T_o which "acts as a probe, just like nondefective T, triggering agreement and potentially movement, but it fails to mark uninterpretable features of its goal for deletion" (P&T 2004: 512). If all unaccusative verbs have defective T_o , then the Russian examples of Accusative Unaccusatives and Polish *-no/-to* constructions (constructions with unaccusative verbs whose accusative objects occur in the subject position) constitute evidence that a defective T head checks Acc case.

(47) Russian Finite Accusative Unaccusative [*Moskovskij komsomolec* 9/13/99]

- a. Soldata ranilo pulej. (Lavine & Freidin 2002: 258)
 soldier_{ACC} wounded_{3.SG.N} bullet_{INST}
 'A soldier was wounded by a bullet.'
- b. Podvaly zatopilo livnem.
 basements_{ACC} flooded_{3.SG.N} downpour_{INST}
 'Basements were flooded by the downpour.'

(48) Polish *-no/-to* constructions²⁵

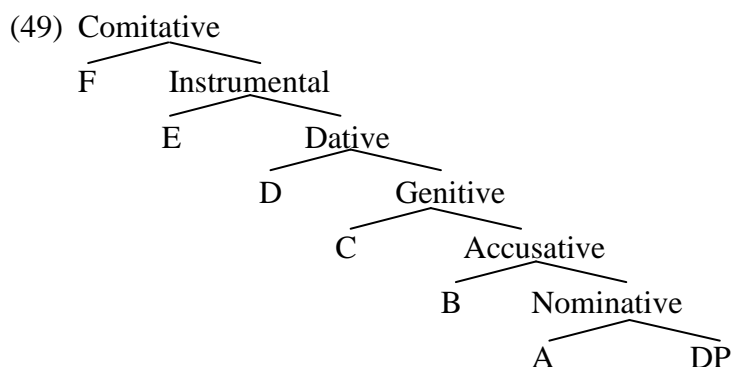
- a. Marię wezwano do sądu.
 Mary_{ACC} called_{3.SG.N} to court
 'Mary was called to court.'
- b. Marii nie wezwano do sądu. → Gen of negation
 Mary_{GEN} not called_{3.SG.N} to court
 'Mary was not called to court.'
- c. Amerykę odkryto przypadkiem.
 America_{ACC} discovered_{3.SG.N} accident_{INSTR}
 'America was discovered by accident.'
- d. Ameryki nie odkryto przypadkiem. → Gen of negation
 America_{GEN} not discovered_{3.SG.N} accident_{INSTR}
 'America was not discovered by accident.'

Since the Acc object of the unaccusative verb moves to the subject position, it must mean that its uT has not been deleted by T_o (it cannot be argued that the Acc case is lexical, since it changes into Gen under negation, like all instances of structural Acc). It does not exhibit Nom case after checking its uT against T_s because Nom is less marked than Acc. Here, I follow Caha's (2009) nanosyntactic analysis of case-inclusion as presented in (49) below²⁶:

²⁵ Compare (48) with a true passive, where the object is not marked Acc and becomes Nom when probed by T_s ; the genitive of negation cannot affect the nominative subject, even though it is the underlying object:

- (i) a. Maria została wezwana do sądu.
 Mary_{NOM} stayed_{3.SG.F/PAST} called_{3.SG.F} to court
 'Mary was called to court.'
- b. Maria nie została wezwana do sądu.
 Mary_{NOM} not stayed_{3.SG.F/PAST} called_{3.SG.F} to court
 'Mary was not called to court.'

²⁶ I agree with Caha's (2009) idea of case-inclusion in a similar sense to Jakobson's (1984) featural system, also implemented in Matushansky (2008, 2010), namely that every more marked case includes the previous one(s) and the features for which they stand (e.g. Gen of quantification is, simplifying greatly, Acc + [quantification], which fits in perfectly with what we see in Polish). I am not entirely convinced by all the remaining ideas Caha proposes, which he claims follow from the idea presented in (49), but an analysis of these issues is beyond the scope of this paper.



“The tree encodes the proposal that a nominative DP is a type of constituent, in which the DP is the complement of the feature [A]. An accusative is a similar constituent, one which is built on top of the nominative by the addition of [B], and so on.” (Caha 2009: 24). In this analysis Nom is analysed as the least marked case, Acc as the second least marked case and so on, each next case being a composite of the previous one and some feature X. It follows then, that a DP spells out the highest (most complex) case feature and even if it later on checks another case feature, this feature will only become visible if it is more complex than the one the DP already has; Polish numeral expressions epitomize this phenomenon.

Numeral expressions with their visible uT (not marked for deletion by the defective light *p* (iT)) are thus eligible for probing by both T_s and T_o . Due to this, they may still interact with the tense features of the matrix probes.²⁷ This is most clearly visible in the object position where the numeral expression, like any other DP, takes on the cases offered by the verbal T_o . In the subject position, similarly to the situation with Accusative Unaccusatives, the effects of Nom case checking will not be visible due to the Nom case being the least marked case, hence unable to trump the already present (more complex) Acc.

4.4 Combining the two proposals: light *p* selects a lexicalised Num

From a purely descriptive perspective, based on what we see in empirical evidence, the light *p* must be assumed to select a lexicalised Num. This is because we see Acc case only on these numerals which I have earlier argued to be reanalysed as a functional head Num; these include numerals 5< and such vague numerals as *masę* ‘mass’, *trochę* ‘a bit’, *kopę* ‘gross’, as well as *kilku*_{V.ACC}/*kilka*_{NV.ACC} ‘several’ and *wielu*_{V.ACC}/*wiele*_{NV.ACC} ‘many’.²⁸ This is not far from what other prepositions co-occurring with numeral expressions select, for instance *po* selects all quantified NumPs (thus including paucal numerals), and approximative *z* ‘around’/*około* ‘around’/*do* ‘up to’/*ponad* ‘over’ select not only quantified NumPs, but also M(easure)Ps.²⁹ Importantly, there is a difference between these prepositions and the null light *p*: the lexical prepositions carry particular semantics with them, whereas the null light preposition does not. Nevertheless, there are important similarities between them; for instance both *z* ‘around’ and *około* ‘around’ allow case-transmission just like the light *p* is argued to do in my analysis, and they also seem to be theta-role transmitters, as the numeral expressions in their complements remain arguments of the selecting verbs. Consider the following examples

²⁷ This also sheds light on their distribution (an issue that P&T’s analysis was most concerned with, however, from my analysis it follows that Polish numeral expressions being DPs will exhibit distribution typical of unquantified DPs, a prediction which is borne out).

²⁸ These numerals were taking on the numeral *-u* ending already in 16th c. coinciding with the numeralisation process of the first batch of numerals 5-10 (Siuciak 2008: 89). Siuciak ascribes this to the quantificational properties common to both groups of lexical items.

²⁹ I call them prepositions due to their syntactic behaviour, but in the Polish literature on the subject they are referred to as adnumeral operators (see e.g. Wajszczuk 2010).

with *około* (which selects numeral expressions and measure nouns and checks structural Gen case, hence the relevant case-changes are only visible in lexical case environments):

- (50) a. Czekałam około godziny/minuty/tygodnia. → structural Gen checked
waited_{1.SG.IMP} around hour_{GEN}/minute_{GEN}/week_{GEN} by *około*
'I've waited about an hour/minute/week.'
- b. Pomogli około pięciu tysiącom Polaków. → lexical Dat checked V
helped_{3.PL} around five_{DAT} thousands_{DAT} Poles_{GEN}
'They helped around five thousand Poles.'
- c. Nie było około pięciu tysięcy Polaków. → Gen of negation
not were_{3.SG.N} around five_{GEN} thousands_{GEN} Poles_{GEN}
'There were around five thousand Poles absent/missing.'
- d. Widziałam około pięciu tysięcy Polaków. → structural Acc checked
saw_{1.SG} around five_{GEN} thousands_{GEN} Poles_{GEN} by V (trumped by
'I saw around five thousand Poles.' structural Gen of *około*)
- e. Opiekują się około pięcioma tysiącami Polaków. → lexical Instr checked
care_{3.PL} self around five_{INSTR} thousands_{INSTR} Poles_{GEN} by V
'They are taking care of around five thousand Poles.'

I would like to argue that *około* instantiates an overt light *p* (notice, it also checks structural Gen, as opposed to the lexical cases checked by regular Ps) and should also be assumed to merge in Spec,DP like the light null *p* (cf. (46)). As we can see in the examples, the cases on the numeral expression are the ones governed by the relevant verbs, not by the light *p*. The preposition is transparent, just as expected. Also, it does not seem to project, i.e. the numeral expression remains a DP, and thus may be further selected by a regular preposition such as the distributive *po* (*około* behaves as if it were not there).

- (51) a. Dałam im po około dziesięć złotych.
gave_{1.SG.F} them_{DAT} *po* around ten_{ACC} zlotys_{GEN}
'I gave them around ten zlotys each.'

While *ponad* 'over' behaves similarly to *około* 'around' (one can substitute *ponad* for *około* in all of the above examples, although it must be noted that *ponad* may also combine with Nom), approximative *z* 'around' (hence *z_{APPROX}*), though transparent to case-marking and theta-role assignment, does not allow co-occurrence with distributive *po*. This might indicate that *z_{APPROX}* actually projects into PP above DP. There is further evidence that this indeed may be the case. In Polish we see combinations of two prepositions one below the other, and interestingly, they always seem to involve a light preposition as the lower one. Approximative *z* simply does not allow such embedding.

- (52) a. Sprzedałam po (około/ponad) pięć litrów mleka. (*około/*ponad po...)
I-sold *po*_{DISTR} (around/over) five_{ACC} litres_{GEN} milk_{GEN}
'I sold around five litres of milk each.'
- b. Czekałam do (około/ponad) pięciu godzin. (*około do...)
I-waited to (around/over) five_{GEN} hours_{GEN}
'I waited up to around five hours'
- c. w/przez (około/ponad) pięć godzin.³⁰ (*około w/przez...)

³⁰ Examples (52) are seemingly problematic in view of the hierarchy I proposed earlier where more marked cases trump the less marked ones, i.e. it is surprising that Gen marked by *około* gets trumped by Acc marked by

- in/through (around/over) five_{ACC} hours_{GEN}
 ‘In/for about/over five hours.’
- d.* Sprzedałem po z pięć litrów mleka.
 I-sold po_{DISTR} z_{APPROX} five_{ACC} litres_{GEN} milk_{GEN}
 ‘I sold around five litres of milk each.’

On the other hand, phrases with *z_{APPROX}* allow coordination with phrases introduced by light prepositions (null and overt, as exemplified in (53a)), which would suggest that they are of the same type and thus that *z_{APPROX}* is non-projecting. We can also have these coordinated phrases with light prepositions further selected by the distributive *po* (53b) (again apart from *z_{APPROX}*):

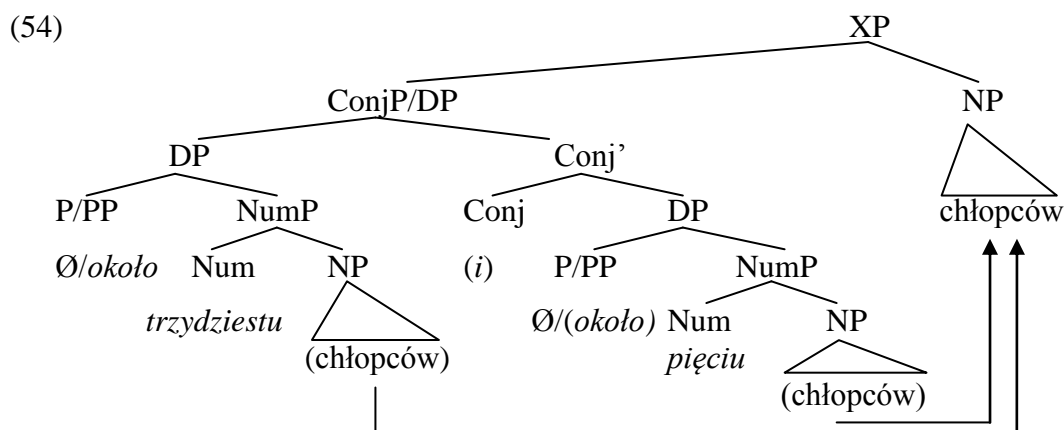
- (53) a. Zaprosiłam dwudziestu studentów, około piętnastu doktorów i z
 invited_{1.SG.F} twenty_{V.ACC} students_{GEN} around fifteen_{V.ACC} doctors_{GEN} and around
 pięciu profesorów.
 five_{V.ACC} professors_{GEN}
 ‘I invited twenty students, around fifteen PhD holders and around five professors.’
- b. Średnio sprzedałam po około pięć chlebów i ponad trzy litry
 mean sold_{1.SG} po around five_{ACC} breads_{GEN} and over three_{ACC} litres_{ACC}
 mleka na osobę.
 milk_{GEN} per person_{ACC}
 ‘On average I sold about five loaves of bread and three litres of milk per person.’

Apart from not allowing embedding by other Ps, *z_{APPROX}* has the signature properties of a light *p*, i.e. it allows case-transmission and Θ -role transmission. I will thus continue assuming that *z_{APPROX}* is a light non-projecting *p*, and that perhaps after having undergone m-merger with D it makes it impossible for the DP to be selected by other Ps.³¹ That m-merger does take place gains support from the examples with pronouns and numerals which can never co-occur with overt light prepositions, but do, as argued here, co-occur with the null one. If, as I assume, pronouns are in D, then m-merger of an overt light P and overt material in D is impossible, although a null light P will not be problematic due to its missing morphophonological realisation, hence we have *nas pięciu* ‘us five’ (with a null light *p*), but never: **około nas pięciu* (rather *około pięciu z nas* ‘around five of us’), **ponad nas pięciu* (rather *ponad pięciu z nas* ‘over five of us’) with an overt light *p*.

To give an example of a complex numeral and show how all of the made proposals work within one structure, consider *trzydziestu pięciu chłopców* ‘thirty five boys’ with a null light *p* and that of *około trzydziestu pięciu chłopców* ‘around thirty five boys’ with an overt light *p*, whose structures are predicted to be parallel:

do/około. Crucially, we are dealing here with two different types of case: lexical and structural. Following Babby’s (1987) case hierarchy in which lexical case always outranks structural case, what we see in (52) is a situation in which the structural Gen marked by *około* is outranked by the lexical Acc marked by *do*. Also, it is possible to say *po około/ponad pięciu_{LOC} dniach_{LOC}*, but the preposition *po* is not distributive here and the meaning is ‘after about five days.’

³¹ I do not have any answer as to why this should be so and will have to leave it to future research. To the best of my knowledge, this is the only preposition that behaves this way and no matter into which group of prepositions we assign it, it will be exceptional in that group: if we assume that it is a light *p*, we need to account for why it cannot be further embedded under regular (projecting) prepositions; if we assume that it is a projecting preposition, we need to account for why it cannot take DPs introduced by light *ps* as complements, which regular prepositions easily do (52). Nevertheless, approximative *z* both semantically and functionally resembles light prepositions, hence my decision to treat it as such. My conclusion is that it could itself be undergoing change.



This structure, again, explains the availability of LBE of the numeral together with the preposition, to the exclusion of the counted noun: *Okolo trzydziestu pieciu/Trzydziestu pieciu zaprosilem chłopców* ‘Around thirty-five/Thirty-five I have invited boys.’

There is more supporting evidence that the prepositional analysis is on the right track. The evidence comes from a related language, Serbo-Croatian (SC). Franks (2002) and Giusti & Leko (2004) discuss properties of the numeral expressions in SC and there is one particularly interesting property with respect to the analysis presented here: numeral expressions cannot occur in oblique case positions unless preceded by a preposition (Giusti & Leko 2004: 135).

- (53) a. Bojao sam se pet ljudi.
feared aux_{1.SG} REFL five people_{GEN}
‘I feared five people.’
b. Čuvao sam se pet ljudi.
guarded aux_{1.SG} REFL five people_{GEN}
‘I guarded myself against five people.’
c. Domogao sam se pet knjiga.
obtained aux_{1.SG} REFL five books_{GEN}
‘I obtained five books.’

According to Franks (2002: 166) the numeral expressions are fine in lexical Gen contexts because they are licensed in Gen DPs. We cannot see case on the numeral *pet* ‘five’ because 5< are indeclinable in SC (Giusti & Leko 2004: 127), the nominal complement of the numeral is Gen anyway, hence no difference will be noticeable. While in Polish numeral expressions are marked with structural Acc (second least marked case), in SC they are marked with lexical Gen. No lexical case may be trumped by any other case, even if it is a more marked one. This is true in both Polish and SC. That is why SC numeral expressions are unacceptable in Dat case contexts (54) (Franks 2002: 166) and Instr case contexts (55)/(56) (Giusti & Leko 2004: 127) (original glosses):

(54)*/?? Jovan je pomagao pet ljudi. → lexical Dat
Jovan aux_{3.SG} helped five people_{GEN}
‘Jovan helped five people.’

(55) a.* Ivan upravlja pet fabrika. → lexical Instr
Ivan manages five factories_{GEN}
‘Ivan manages five factories.’

- b.* Ivan upravlja petima fabrikama. (petima is a non-existent form)
 Ivan manages five_{INSTR} factories_{INSTR}
 ‘Ivan manages five factories.’

- (56) a.* Ivan upravlja tri fabrike. → lexical Instr
 Ivan manages three factories_{GEN}
 ‘Ivan manages three factories.’
 b. Ivan upravlja trima fabrikama.
 Ivan manages three_{INSTR} factories_{INSTR}
 ‘Ivan manages three factories.’

Judging by the grammaticality of (56b), we could say that it is not always the case that numeral expressions are completely infelicitous in oblique case contexts, only the indeclinable ones are. Then again, the numerals 1-4 are adjectival in nature also in SC (Comrie 1992), are not Num heads, and cannot assign Gen to their complement, i.e. they are not of interest here. What is of interest, however, is that the indeclinable numerals start with 5.³² This is no coincidence. Polish 5< also have a rather monotonous inflectional pattern (cf. Table 9), where the complement to the numeral is no longer Gen past the Gen. All the subsequent more marked cases affect the less marked structural Gen. In Polish, numeral expressions regularly appear in lexical case contexts, in SC they cannot do so beyond the Gen. To be more precise, they cannot do so, unless they are preceded by a preposition (57b,d).

- (57) a.* Predsjednik vlada pet zemalja.
 president rules five countries_{GEN}
 ‘The president rules five countries.’
 b. Predsjednik vlada sa pet zemalja.
 president rules with five countries_{GEN}
 ‘The president rules five countries.’
 c.* Jovan je rukovodio pet fabrika.
 Jovan aux_{3.SG} managed five factories_{GEN}
 ‘Jovan managed five factories.’
 d. Jovan je rukovodio sa pet fabrika.
 Jovan aux_{3.SG} managed with five factories_{GEN}
 ‘Jovan managed five factories.’

Here is where the matters become really interesting. This preposition is otherwise disallowed with the very same verb if the object is not a numeral expression (58b,d):

- (58) a. Predsjednik vlada zemljom.
 president rules country_{INST}
 ‘The president rules the country.’
 b.* Predsjednik vlada sa zemljom.
 president rules with country_{INST}
 ‘The president rule the country.’
 c. Jovan je rukovodio jednom fabrikom.³³

³² For a very interesting discussion on why the syntax of 1-4 differs from that of 5< see Hurford (1987), Heine (1997), Ifrah (1998), and Rutkowski (2003).

³³ As we can see the numeral *jeden* patterns with *tri* in that it is adjectival and does not show the properties exhibited by the numerals 5< (which check genitive and are heads of NumP). In (58) it is a specifier of NumP, but crucially not the head of NumP, and thus infelicitous with the preposition.

Jovan aux_{3,SG} managed one_{INST} factory_{INST}
 ‘Jovan managed one factory.’
 d.* Jovan je rukovodio **sa** jednom fabrikom.
 Jovan aux_{3,SG} managed with one factory
 ‘Jovan managed one factory.’

Franks refers to this preposition as ‘light’ and I have been making use of this term throughout the paper. He claims that it is inserted as a last resort mechanism to take on the lexical case that the numeral expression cannot take on. It is also responsible for transmitting the theta-role assigned by the verb. As we can see from the contrast between (58c) and (58d), and the absence of the preposition with *tri* ‘three’ in (56b), *jeden* ‘one’ and the paucal numerals 2-4, which are specifiers of NumP and do not co-occur with *p*. The SC preposition thus shows the selectional properties of my earlier proposed null light *p*: it selects lexicalised Num heads.

Franks (2002) suggested that SC numeral expressions are licensed in Gen contexts, but this would make SC very different from Polish or Czech. I suggest that we can have a common analysis of these languages. The fact that numeral expressions are available in the Gen contexts in (53) without the overt *p* may simply result from the Acc/Gen syncretism also present in Serbo-Croatian; just like it is impossible to tell the difference between Acc and Gen of *pięciu*_{ACC/GEN} ‘five’ in the following Polish examples: *Widziałem pięciu*_{ACC} *mężczyzn*_{GEN} ‘I saw five men’ vs. *Bąłem się pięciu*_{GEN} *mężczyzn*_{GEN} ‘I feared five men,’ so it is impossible to tell the difference between Acc and Gen of *pet ljudi* ‘five people’ in SC (53). This need not mean that all numeral expressions in SC are Gen, they may as well be Acc just like their Polish counterparts. Thus SC (53) is reminiscent of Polish where the Acc virile numeral expressions find themselves in Acc/Gen case contexts (as in the examples immediately above). The difference between Polish and SC is that Polish 5< are declinable and there is a gender distinction between virile/non-virile, which is why there is a way to distinguish Acc from Gen if we look at non-virile: *Widziałem pięć*_{ACC} *kobiet*_{GEN} ‘I saw five women’ vs. *Bąłem się pięciu*_{GEN} *kobiet*_{GEN} ‘I feared five women.’ The fact that the preposition surfaces in SC in oblique case contexts may also be connected to the indeclinability of the numerals. In Polish 5< are declinable and thus able to take on the oblique cases, in SC this job is performed by *p*, as proposed by Franks.

Based on the presented data, I find it plausible to assume the existence of a null light preposition in Polish numeral expressions. Once one becomes aware of the special connection between various prepositions and numerals, one does not need to look very far to find examples supporting its presence and find that plenty of seemingly unconnected properties could be reduced to one.

5 Conclusion

The analysis proposed in this paper has been inspired by Pesetsky & Torrego’s (2004) idea that case is a reflection of tense checking. Because Polish numeral expressions have been believed to be intrinsically accusative (the Accusative Hypothesis) for over a hundred years and no one so far has asked the question of where this accusative case originates from, it only made sense to investigate this puzzling issue. Thus I have tried to find out whether one could account for the accusative case assuming the presence of some tense-bearing head.

My investigation begins at the beginning, i.e. with an overview of Polish cardinals with emphasis on their diachronic development. Thanks to this, I have found the source of the Accusative case in numerals 5<: historically, it stems from the introduction of the Acc/Gen syncretism, which historians believe resulted from the need to make the Nom subjects more prominent as opposed to Acc objects (with the Nom/Acc syncretism so widespread it was no

easy task, hence the innovation). I have also found that paucal numerals 2-4 (and 1) inherited their adjectival nature from Proto-Slavic, whereas the (problematic accusative) numerals 5< used to be regular nouns belonging to the i-declension. After studying the way they have undergone numeralisation, a process which I proposed involves reanalysis of lexical material into functional material (here N into Num), I have proposed a structure of Number Phrase (for both simple and complex numerals) and supported my analysis with evidence based on selectional properties of various prepositions, conjunctions and Left Branch Extractions. Numerals 5< were thus proposed to be lexicalisations of Num, and thus constituting an integral part of the nominal extended projection. I have presented evidence in favour of the Accusative Hypothesis based on Przepiórkowski (2004), and included evidence from Rutkowski (2007) showing that cardinal numerals (5<) can no longer be considered nouns (contra Corver & Zwarts 2006 and Ionin & Matushansky 2004, 2006).

The only plausible candidate for a tense-head within numeral expressions seemed to be a preposition, which would then constitute the source of the accusative case of the numeral expression. Based on this assumption, I have found compelling evidence for the existence of a null light *p* which did exactly that. On the basis of the data presented in Przepiórkowski (2006, 2008), Franks (2002), and Giusti & Leko (2004), I have showed that there exist (certain) prepositions which select numerals. Moreover, I have given examples of two overt prepositions (*około* ‘around’ checking Gen, and *ponad* ‘over’ checking Nom/Acc) which show ‘light’ behaviour, i.e. they are both case-transmitters and theta-role transmitters (Franks 2002). Thus, I have argued for an analysis of numeral expressions in which they are assumed to be nominal expressions (DPs) introduced via a light preposition. This preposition, has been argued to be defective (a defective iT) in the sense that it can check uT on D but cannot mark it for deletion, hence the Acc case of the numeral expressions may be trumped by more marked cases (here I followed Caha’s (2009) nanosyntactic analysis of case-inclusion). To prove that defective tense heads (also assumed to be present in unaccusatives) can check Acc, I have presented evidence from Russian Accusative Unaccusative constructions (Lavine & Freidin 2002) as well as Polish *-no/-to* constructions and showed that this is in fact possible ; the fact that the uT on D in these constructions is not deleted can be deduced from the fact that these nominals are still visible to the T_s probe; assuming, however, that numeral expressions are also marked Acc by a defective iT, I provided evidence that such a defective tense head not only cannot delete the uT on D, but cannot even mark it for deletion, because here the uT on D is not only available for probing by $T_{s/o}$, but also enters a checking relation with $T_{s/o}$; results of this relation are only visible with cases more marked than Acc, hence not in the subject position where Nom is checked, but only in the object/complement position.

Though I have presented evidence for a selectional relationship between certain prepositions and numerals, these elements cannot and do not form a constituent to the exclusion of the lexical noun in Polish (as it is argued by Corver & Zwarts (2006) for Dutch). The behaviour of the numeral expressions is that of a nominal expression, and it does not parallel the behaviour of prepositional phrases (see also Corver & Zwarts (2006) for a similar conclusion). Making use of Matushansky’s (2006) m-merger, I have argued that the light preposition is inserted at the root to merge with D/DP and instead of projecting itself, becomes the specifier of DP, i.e. D projects. This is followed by morphological merger of the two heads. An analysis along these lines accounts for the purely nominal behaviour of the numeral expressions even when selected by an overt light preposition *około* (cf. (54)), which can be further selected by other probes.

References

- Abbott, B. 1974. Right node raising as a test for constituenthood. *Linguistic Inquiry* 7: 639–642.

- Alexiadou, A. 2004. Inflection Class, Gender and DP internal Structure. In: Müller, G., L. Gunkel & G. Zifoun (eds.) *Explorations in Nominal Inflection*. Berlin: Mouton de Gruyter.
- Alexiadou, A., M. Stavrou & L. Haegeman. 2007. *Noun Phrase in the Generative Perspective*. Berlin/New York: Mouton de Gruyter.
- Babby, L. 1987. Case, Prequantifiers, and Discontinuous Agreement In Russian. *Natural Language & Linguistic Theory* 5.1: 91-138.
- Bailyn, J. & A. Nevins. 2008. Russian Genitive Plurals are Impostors. In: Nevins, A. & A. Bachrach (eds.) *The Basis of Inflectional Identity*. Oxford University Press.
- Bernstein, J. 1993. The syntactic role of word markers in null nominal constructions. *Probus* 5: 5-38.
- Blake, B. 1994. *Case*. Cambridge University Press.
- Bobrowski, I. 1988. *Gramatyka opisowa języka polskiego. (Zarys modelu generatywno-transformacyjnego). T. 2: Od struktur wyjściowych do tekstu*. Kielce: Wyższa Szkoła Pedagogiczna im. Jana Kochanowskiego.
- Bogusławski, A. 1973. Nazwy pospolite przedmiotów konkretnych i niektóre właściwości ich form liczbowych i połączeń z liczebnikami w języku polskim. In: Topolińska, Z. & M. Grochowski (eds.) *Liczba, ilość, miara*, pp. 1-32. Wrocław: Zakład narodowy im. Ossolińskich.
- Bresnan, J. 1974. The position of certain clause particles in phrase structure. *Linguistic Inquiry*, 5: 614-619.
- Caha, P. 2009. *The nanosyntax of case*. Ph.D. dissertation, University of Tromsø.
- Chomsky, N. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Chomsky, N. 1999. Derivation by Phase. *MITWPL* 18. Cambridge, MA: MIT Press.
- Comrie, B. 1992. Balto-Slavonic. In: Gvozdanović, J. (Ed.) *Indo-European Numerals* (Trends in Linguistics/Studies and Monographs 57), pp.717-833. Berlin/New York: Mouton de Gruyter.
- Corbett, G.G. 1978a. Universals in the syntax of cardinal numerals. *Lingua* 46: 355-368.
- Corbett, G.G. 1978b. Problems in the syntax of Slavonic numerals. *The Slavonic and East European Review* 56.1: 1-12.
- Corver, N. & J. Zwarts. 2006. Prepositional numerals. *Lingua* 116: 811-835.
- De Vicenzi, M. & E. Di Domenico. 1999. A distinction between ϕ -features: the role of gender and number in the retrieval of pronoun antecedents. *Rivista di Linguistica* 11.1: 41-74.
- De Vicenzi, M. 1999. Differences between the morphology of gender and number: evidence from establishing coreference. *Journal of Psycholinguistic Research* 28.5: 537-553.
- Doroszewski, W. 1952. *Podstawy gramatyki polskiej*. Warszawa: PWN.
- Falimirz, S. 1534. *O ziołach y o mocy gich*. Kraków: F. Ungler.
- Franks, S. 1994. Parametric properties of numeral phrases in Slavic. *Natural Language & Linguistic Theory* 12: 597-674.
- Franks, S. 1995. *Parameters of Slavic Morphosyntax*. Oxford University Press. Oxford.
- Franks, S. 2002. A Jakobsonian feature based analysis of the Slavic numeric quantifier genitive. *Journal of Slavic Linguistics* 10: 141-181.
- Giusti, G. 1991. The categorial status of quantified nominals. *Linguistische Berichte* 136: 438-452.
- Giusti, G. & N. Leko. 2004. The Categorial Status of Quantity Expressions. *Lingvistički Vidici* 34/05: 121-183.
- Glaber, A. (1535). *Problemata Aristotelis. Gadki z pisma wielkiego philozofa Aristotela*. Kraków: F. Ungler.
- Grappin, H. 1942. *Grammaire de la langue polonaise*. Paris: Droz.

- Grappin, H. 1950. *Les noms de nombre en polonaise*. Kraków: Polska Akademia Umiejętności.
- Greenberg, J. 1963. Some Universals of Grammar with particular Reference to the Order of Meaningful Elements. In: Greenberg, J. (ed.) *Universals of Human Language*. Cambridge, MA: MIT Press.
- Greenberg, J. 1978. Generalizations about Numeral Systems. In: Greenberg, J. (ed.) *Universals of Human Language, Vol.3*, pp. 249-295. Stanford, CA: Stanford University Press.
- Grimshaw, J. 1991. Extended projection. (unpublished manuscript)
- Grimshaw, J. 1997. Locality and Extended Projection. In Coopmans, P., M. Everaert & J. Grimshaw (eds.) *Lexical Specification and Insertion*, pp. 115-33. Amsterdam; Philadelphia: John Benjamins.
- Gvozdanović, J. 1999. Some remarks on numeral morphosyntax in Slavic. In: Gvozdanović, J. (ed.) *Numeral Types and changes Worldwide*, pp.187-196. Berlin/New York: Mouton de Gruyter.
- Gwagnin, A. 1611. *Kronika Sarmacyey Europskiej*. Kraków: M. Loba.
- Heine, B. 1997. *Cognitive foundations of Grammar*. New York/Oxford: Oxford University Press.
- Hurford, J. 1975. *The linguistic theory of numerals*. Cambridge: Cambridge University Press.
- Hurford, J. 1987. *Language and number*. Oxford: Basil Blackwell.
- Hurford, J. 2003. The interaction between numerals and nouns. In: Plank, F. (Ed.) *Noun Phrase Structure in the Languages of Europe*, pp. 561-620. Berlin/New York: Mouton de Gruyter.
- Ifrah, G. 1998. *The Universal History of Numbers*. London: The Harvill Press.
- Ionin, T. & O. Matushansky. 2004. A singular plural. In B. Schmeiser, V. Chand, A. Kelleher & A. Rodriguez (eds), *WCCFL 23: Proceedings of the 23rd West Coast Conference on Formal Linguistics*, 101-114. Somerville, MA: Cascadilla Press.
- Ionin, T. & O. Matushansky. 2006. The Composition of Complex Cardinals. *Journal of Semantics* 23: 315–360.
- Jakobson, R. 1984. *Russian and Slavic Grammar. Studies 1931-1981*. Berlin/New York/Amsterdam: Mouton Publishers.
- Janda, L. 1999. Whence virility? The rise of a New gender distinction in the history of Slavic. In: Mills, M. (ed.) *Slavic Gender Linguistics*, pp. 201-228. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Jakubowicz, M. 1892. *Gramatyka języka polskiego*. Wilno.
- Klemensiewicz, Z. 1930. Liczebnik główny w polszczyźnie literackiej. Historia formy i składni. *Prace Filologiczne* 15: 1-130.
- Klemensiewicz, Z. 1937. *Składnia opisowa współczesnej polszczyzny kulturalnej*. Kraków: PAN.
- Klemensiewicz, Z. 1952. *Podstawowe wiadomości z gramatyki języka polskiego*. Kraków: Wyd. Studium Słowiańskiego Uniwersytetu Jagiellońskiego.
- Klemensiewicz, Z. 1957. *Zarys składni polskiej*. Warszawa: PWN.
- Koptjevskaja-Tamm, M. 2009. A lot of grammar with a good portion of lexicon: towards a typology of partitive and pseudo-partitive nominal construction. In: Helmbrecht et al. (eds.) *Form and Function in Language Research. Papers in Honour of Christian Lehmann*, pp. 329-346. Berlin/New York: Mouton de Gruyter.
- Krasnowolski, A. 1897. *Systematyczna składnia języka polskiego*. Warszawa: Drukarnia Estetyczna K. Sierpińskiego.
- Kryński, A. 1900. *Gramatyka języka polskiego*. Warszawa: M. Arct.

- Laskowski, R. 1984. Liczebnik. In: Grzegorzczkova, R., R. Laskowski & H. Wróbel (eds.) *Gramatyka współczesnego języka polskiego. Morfologia*, pp. 283-293. Warszawa: PWN.
- Lavine, J. & R. Freidin. 2002. The Subject of Defective T(ense) in Slavic. *Journal of Slavic Linguistics* 10: 253–289.
- Małecki, A. 1863. *Gramatyka języka polskiego większa*. Lwów.
- Matushansky, O. 2006. Head Movement in Linguistic Theory. *Linguistic Inquiry* 37.1: 69–109.
- Matushansky, O. 2008. A Case Study of Predication. In: Marušič, F. and R. Žaucer (eds.), *Studies in Formal Slavic Linguistics. Contributions from Formal Description of Slavic Languages* 6.5.: 213-239. Frankfurt am Main: Peter Lang.
- Matushansky, O. 2010. Some cases of Russian. In: Dudchuk, P. et al. (eds.) *Proceedings of FDSL* 7.5.
- Mieczkowska, H. 1994. *Kategoria gramatyczna liczebników w ujęciu konfrontatywnym polsko-słowackim*. Kraków: Uniwersytet Jagielloński.
- Müller, N. 2000. Gender and number in acquisition. In: Unterbeck, B. & M. Rissanen (Eds.) *Gender in Grammar and Cognition*. Berlin/New York: Mouton de Gruyter.
- Nelson, D. & I. Toivonen. 2000. Counting and the grammar: case and numerals in Inari Sami. In: Nelson, D. & P. Foulkes (Eds.) *Leeds Working Papers in Linguistics* 8, pp. 179- 192.
- Panagiotidis, P. 2002. *Pronouns, Clitics and Empty Nouns*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Panagiotidis, P. 2003. One, Empty Nouns, and θ -Assignment. *Linguistic Inquiry* 34.2: 281-292.
- Paprocki, B. 1584. *Herby Rycerstwa Polskiego*. Kraków: Biblioteka Polska.
- Pereltsvaig, A. 2008. Split Phrases in Colloquial Russian. *Studia Linguistica* 62.1: 5–38.
- Perlmutter, D. & J. Orešnik. (1973). Language-particular rules and explanation in syntax. In: Anderson, S. & P. Kiparsky (Eds.) *A Festschrift for Morris Halle*, pp. 419-459. New York: Holt Rinehart.
- Pesetsky, D. & E. Torrego. 2004. Tense, Case, and the Nature of Syntactic Categories. In: Guéron, J. and J. Lecarme (eds.) *The Syntax of Time*. Cambridge, Massachusetts: MIT Press.
- Postal, P. M. 1974. *On raising*. Cambridge, MA: MIT Press.
- Przepiórkowski, A. 1996. Case assignment in Polish: Towards an HPSG analysis. Grover, C. & E. Vallduvi (eds.) *Studies in HPSG, Edinburgh Working Papers in Cognitive Science*, 12: 191–228. Centre for Cognitive Science, University of Edinburgh.
- Przepiórkowski, A. 2004. O wartości przypadku podmiotów liczebnikowych. *Bulletin de la Société Polonaise de Linguistique* LX: 133-143.
- Przepiórkowski, A. 2006. O dystrybucyjnym PO i liczebnikach jedynkowych. *Polonica* XXVI–XXVII, 171–178.
- Przepiórkowski, A. 2008. Generalised quantifier restrictions on the arguments of the Polish distributive *po*. *Cognitive Studies/Études Cognitives* 8: 11-29.
- Ritter, E. 1993. Where's Gender? *Linguistic Inquiry* 24.4: 795-803.
- Roberts, I. & A. Roussou. 2003. *Syntactic Change. A minimalist Approach to Grammaticalization*. Cambridge University Press.
- Ross, J. R. 1967/1986. *Infinite syntax*. Norwood: Ablex Publishing
- Rutkowski, P. 2000. Składnia polskich grup liczebnikowych: próba opisu formalnego. *Poradnik Językowy* 8:10–28.
- Rutkowski, P. 2003. On the universal neuropsychological basis of the syntax of numerals. *Journal of Universal Language* 4.2: 147-182.
- Rutkowski, P. 2007. *Hipoteza frazy przedimkowej jako narzędzie opisu składniowego polskich grup imiennych*. Ph.D. dissertation, Warsaw University.

- Rutkowski, P. & K. Szczegot. 2001. On the syntax of functional elements. In: Przepiórkowski, A. & P. Bański (Eds.), *Generative linguistics in Poland: Syntax and morphosyntax*. Warszawa: Instytut Podstaw Informatyki PAN, 187–196.
- Sabbagh, J. 2007. Ordering and linearizing rightward movement. *Natural Language & Linguistic Theory* 25: 349–401.
- Saloni, Z. 2005. O przypadkach w języku polskim (na marginesie artykułu Adama Przepiórkowskiego). *Biuletyn Polskiego Towarzystwa Językoznawczego LXI*: 27–48.
- Saloni, Z. & M. Świdziński. 1987. *Składnia współczesnego języka polskiego*. Warszawa: PWN.
- Schabowska, M. 1967. *Rzeczowniki ilościowe w języku polskim*. Wrocław: PAN.
- Schenker, A. 1971. Some remarks on Polish quantifiers. *The Slavic and East European Journal* 15.1: 54–60.
- Seiler, H. 1990. A dimensional view of numeral systems. In: Croft, W., K. Denning, S. Kemmer (eds.) *Studies in Typology and Diachrony: Papers presented to Joseph H. Greenberg on his 75th birthday* (Typological Studies in Language 20), pp. 187–208. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Selkirk, E. 1977. Some remarks on noun phrase structure. In: Culicover, P., Wasow, T., A. Akmajian (eds.) *Studies in Formal Syntax*. Academic Press, New York, pp. 285–316.
- Siuciak, Mirosława. 2008. *Kształtowanie się kategorii gramatycznej liczebnika w języku polskim*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- Słoboda, Agnieszka. 2005. Składnia grupy imiennej w staropolszczyźnie. Na przykładzie wielkopolskich rot sądowych XIV i XV wieku. *Biblioteczka Poznańskich Studiów Polonistycznych. Seria Językoznawcza* 28. Poznań.
- Stampe, D. 1976. Cardinal Number Systems. In: Muwfen, S., C. Walker & S. Steever (Eds.) *Papers from the Twelfth regional Meeting, Chicago Linguistic Society, April 23–25, 1976*, pp. 594–609. Chicago: Chicago linguistic Society.
- Stavrou, M. 2003. Semi-lexical nouns, classifiers, and the interpretation(s) of the pseudopartitive construction. In: Coene, M. & Y. D'Hulst (Eds.) *From NP. to DP, Vol 1: The syntax and semantics of noun phrases*, pp. 329–353. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Szober, S. 1928. *Trzy piękne córki było nas u matki*: Formy podmiotu i orzeczenia w zdaniach z podmiotem logicznym, określonym przydawką liczebnikową. *Język Polski* XIII.4: 97–106.
- Szober, S. 1953. *Gramatyka języka polskiego*. Warszawa: Nasza Księgarnia (3rd edition).
- Szymczak, M., H. Szkiładź, S. Bik, M. Krajewska, B. Pakosz, E. Sobol, C. Szkiładź (eds.). (1999). *Słownik Języka Polskiego PWN*. Warszawa: Wydawnictwo Naukowe PWN.
- Van Riemsdijk, H. 1998. Categorical feature magnetism: The endocentricity and distribution of projections. *Journal of Comparative Germanic Linguistics* 2: 1–48.
- Von Mengden, F. 2010. *Cardinal Numerals: Old English from a Cross-Linguistic Perspective*. Berlin/New York: Mouton de Gruyter.
- Wajszczuk, J. 2010. Functional class (so-called “Part of Speech”) assignment as a kind of meaning-bound word syntactic information. *Cognitive Studies/Études Cognitives* 10: 15–33. Warsaw: SOW Publishing House.
- Zamparelli, R. 2008. “Dei ex machina: A note on plural/mass indefinite determiners.” *Studia Linguistica* 62.3: 301–327.