
Case Syncretism, Features, and the Morphosyntax of Polish Numeral Phrases

Gilbert C. Rappaport

University of Texas at Austin

Contents

1	Introduction	123
2	The paradox	123
3	Higher numerals	125
4	Lower numerals	132
5	Conclusion	135

1 Introduction

Numeral phrases in Polish, as in other West and East Slavic languages, exhibit puzzling behavior. In particular, to judge by the morphosyntax of the constituents of the numeral phrase, the internal structure of the phrase appears to depend upon the external context of that phrase. The problem is a well-known paradox of Slavic syntax, and a considerable literature has been devoted to it.¹ The purpose of this paper, which complements other work by the author (Rappaport 2002, 2003), is to develop a Minimalist theory of formal features, case syncretism, and lexical representation which effectively removes the paradox by maintaining (despite appearances) the usual independence of internal phrase structure and external context, while placing the morphosyntactic variation where such anomalous facts belong: in the lexicon. After defining the apparent paradox more carefully in section 2, we develop our analysis for two different categories of numeral in sections 3 and 4. This paper will focus on Polish material, with a few limited comparative observations involving Russian. The approach can, and, we believe, should, be extended to encompass the grammatical patterns observed in the other pertinent Slavic languages, as well as additional details of Polish itself.

2 The paradox

Consider first a simple Noun Phrase such as *te miłe dziewczynki* ‘these nice girls’ in an oblique case context. As an example of such a position we take the complement

¹See, for example, Babby 1987, Corbett 1993, Franks 1995, Halle 1990; 1995, Mel’čuk 1985, and Przepiórkowski 1999, 2003, along with references cited there.

of the verb *zachwycić się* ‘to be captivated by’ governing the instrumental case. The morphosyntax of this phrase in (1) makes it clear that *dziewczynki* ‘girls’ is the head, and the demonstrative *te* ‘these’ and the adjective *miłe* ‘nice’ agree with that head:

- (1) Zachwyciłem się [tymi miłymi dziewczynkami]_{INST.}
 ‘I was captivated by [these_{INST.PL} nice_{INST.PL} girls_{INST.PL}].’

If we substitute a numeral, for example, *sześć* ‘six’, for the demonstrative pronoun, the numeral agrees with the head noun just as the demonstrative pronoun does:

- (2) Zachwyciłem się [sześcioma miłymi dziewczynkami]_{INST.}
 ‘I was captivated by [six_{INST} nice_{INST.PL} girls_{INST.PL}].’

In contrast, if we perform this substitution of a numeral for a demonstrative in a direct case position (assigned the nominative or accusative case), the agreement pattern within the numeral phrase takes a different form:

- (3) a. [Te miłe dziewczyny]_{NOM} bawią się na podwórzu.
 ‘[These_{NOM.PL} nice_{NOM.PL} girls_{NOM.PL}] are playing in the courtyard.’
 b. [Sześć miłych dziewczynek]_{NOM} bawi się na podwórzu.
 ‘[Six_{NOM} nice_{GEN.PL} girls_{GEN.PL}] are playing in the courtyard.’

While in (3a) it is clear that the noun in the subject phrase is the head of that phrase, in (3b) it appears that the numeral is the head, taking the phrase *miłe dziewczyny* as a genitive complement. The morphosyntactic agreement pattern of (3b) thus resembles not that of the numeral phrase in (2), but that of the non-numeral subject phrase in (4):

- (4) [Grupa miłych dziewczynek]_{NOM} bawi się na podwórzu.
 ‘[A group_{NOM} (of) nice_{GEN.PL} girls_{GEN.PL}] are playing in the courtyard.’

Following Babby (1987), we will call the agreement pattern of the numeral phrase in (2) *homogeneous*, and that in (3b) — *heterogeneous*. The paradox of Polish numeral phrases (and those of many other Slavic languages) is the correlation of homogeneous morphosyntax with oblique case contexts, and of heterogeneous morphosyntax — with direct case contexts. That the *internal* structure of a phrase depends upon its *external* context in this way is a conclusion difficult to accept, and arguments against it will be presented here. For now we will adopt the following working definition: a *numeral* represents a distinct part of speech which participates in such a contextually-determined alternation between homogeneous and heterogeneous morphosyntax in its phrase. The goal of this paper is to provide an explanatory account of this behavior.

Before proceeding, some terminological distinctions are in order. First, there are ‘numbers’ which are not ‘numerals’. On the one hand, the number *jeden* ‘one’ always exhibits homogeneous morphosyntax, so it is simply an adjective in category. On the other hand, the numbers *tysiąc* ‘thousand’, *milion* ‘million’, and *miliard* ‘billion’ always exhibit heterogeneous morphosyntax, and correspondingly are simply nouns. Conversely, some ‘numerals’ are not ‘numbers’: quantitative expressions such as *wiele* ‘many’ and *ile* ‘how many’, among others, exhibit the alternation in question, despite

the fact they do not denote cardinal quantities. Nevertheless, such words qualify by definition as numerals.

As for numerals proper, we distinguish two categories in Polish. Section 3 investigates the default situation, which we will associate with *higher numerals*: the genitive case is regularly assigned to the quantified noun and, correspondingly, to its agreeing modifiers. *Lower numerals* (*dwa* ‘two’, *trzy* ‘three’, *cztery* ‘four’, and *oba*, *obydwa* ‘both’), which do not assign the genitive case, will be discussed in section 4. We will develop in this paper a more precise definition of numeral, which will reveal greater similarity between higher and lower numerals than may be apparent at first glance.

3 Higher numerals

We will assume, with some simplification, that the standard distinction of three genders among singular Polish nouns (masculine, feminine, neuter) is reduced in the plural to a two-way distinction: masculine personal (typically, at least one member of the denoted group is a human being of masculine gender) and non-masculine personal (everything else). The morphosyntax of higher numeral phrases headed by a non-masculine personal noun was illustrated above ((2)–(3)), and displays the grammatical paradox discussed previously, which we state as follows:

- (5) a. Direct case positions (nominative and accusative) → heterogeneous morphosyntax
 b. Oblique case positions (all others) → homogeneous morphosyntax

But a higher numeral phrase headed by a masculine personal noun in an accusative position is problematic:

- (6) Widzę [sześciu miłych chłopców]_{ACC}.
 ‘I see [six_{GEN} nice_{GEN.PL} boys_{GEN.PL}].’

Sense could be made of (6) in either of two ways. It could be assimilated to (7) by seeing here the heterogeneous morphosyntax expected in this direct case position.

- (7) Widzę [bandę miłych chłopców]_{ACC}.
 ‘I see [a band_{ACC.SG} (of) nice_{GEN.PL} boys_{GEN.PL}].’

That is, the numeral and the phrase following it are in the genitive for different reasons. But there is no obvious reason for the genitive marking of the numeral. It appears to exhibit the genitive accusative, a syncretism of the accusative with the genitive characteristic of masculine animate nouns in the singular and masculine personal nouns in the plural. However, a numeral (unlike *chłopiec* ‘boy’) is neither, so that it appears to be agreeing in these features with the following noun, unlike in (7).

Alternatively, (6) could be assimilated to the homogenous pattern of (8):

- (8) Widzę [tych miłych chłopców]_{ACC}.
 ‘I see [these_{ACC=GEN.PL} nice_{ACC=GEN.PL} boys_{ACC=GEN.PL}].’

But the morphological substitution of the genitive for the accusative hardly seems adequate to qualify the direct object of (8) as a syntactically oblique position, the en-

vironment in which numeral phrases exhibit homogeneous morphosyntax (see (5)).

Ironically, the solution is suggested by an even more counter-intuitive fact of Polish grammar: masculine personal higher numeral phrases take the same form in the nominative as in the accusative:

- (9) [Sześciu miłych chłopców]_{NOM} bawi się na podwórzu.
'[Six_{GEN} nice_{GEN.PL} boys_{GEN.PL}] are playing in the courtyard.'

In this nominative case position there can be no appeal to a syncretism with the genitive case, so that the genitive accusative does not appear to be part of the explanation for (6); in the absence of the numeral (cf. (10a)), the quantified noun stands in the nominative, as expected, so that (9) does not reflect homogeneous morphosyntax. But to view (9) as exhibiting heterogeneous morphosyntax, and analogous to (10b), provides no account for the genitive on the numeral itself.

- (10) a. [Mili chłopcy]_{NOM} bawią się na podwórzu.
'[Nice_{NOM.PL} boys_{NOM.PL}] are playing in the courtyard.'
b. [Banda miłych chłopców]_{NOM} bawi się na podwórzu.
'[A band_{NOM} (of) nice_{GEN.PL} boys_{GEN.PL}] is playing in the courtyard.'

In other words, neither of our analogous patterns (10a)–(10b) can account for the case forms exhibited by (9), just as neither (7) nor (8) shed light on (6).

Let us assume without argument here (following, eg., Babby 1987 and others) that the syntactic structure of numeral phrases is the same, regardless of whether they exhibit homogeneous or heterogeneous morphosyntax. In both cases, the quantified noun is the head. To assume otherwise seems to present an unworkable syntactic mechanism. The observed variation is in the principles of morphological marking. Our approach will be to address the following two questions:

- (11) a. How can we best account for the two kinds of morphosyntax exhibited by numeral phrases (their *internal* grammar)?
b. How can this contrast in internal grammar be correlated with the associated contrast in syntactic context between direct and oblique case contexts (their *external* grammar), as defined in (5)?

We propose a *Minimalist* approach (Chomsky 2000, 2001a, 2001b), in which the fundamental properties of the case-agreement system are regulated by the operation *Agree*. We take this mechanism ‘off the shelf’, as it were, and place the distinctive aspects of Polish morphology where they belong: in the lexicon. We begin by defining this basic operation and reviewing how it functions in assigning structural nominative and accusative cases in a basic transitive clause structure. With this as background it will be possible to proceed to our analysis of Polish numeral phrases.

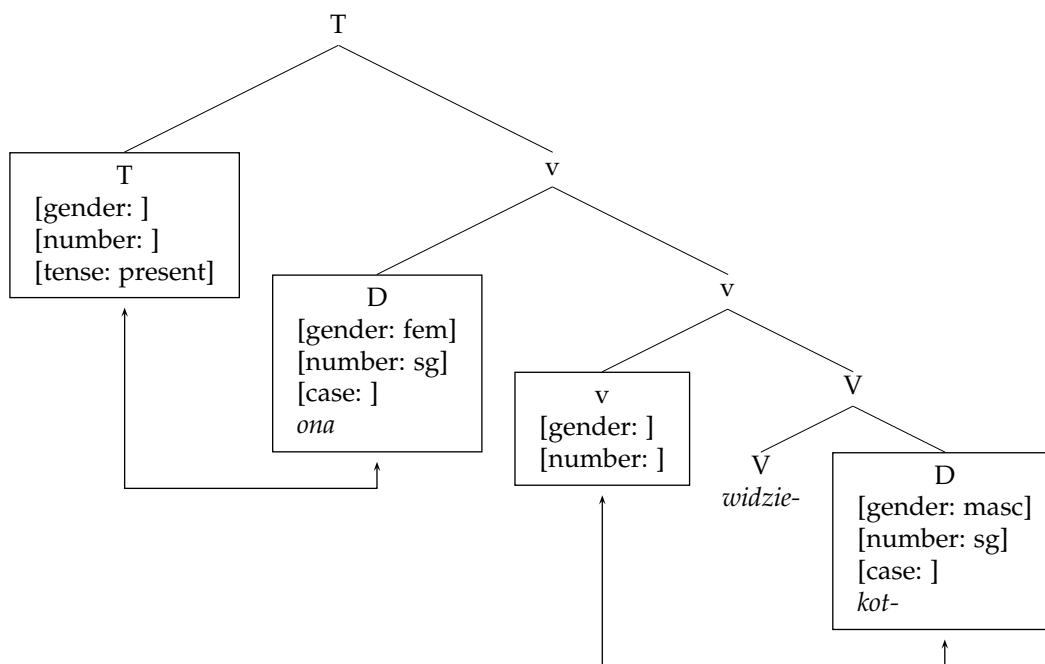
Formal features (i.e., information) consist of a *type* and, optionally, a *value*. For example, the feature type for case can be *valued* (taking the form, for example, [case: nominative]) or *unvalued* (taking the form [case:]). Two features on different syntactic objects *match* if they contain the same type, regardless of their value. A category is *active* if it contains an unvalued feature.

Two syntactic objects A and B undergo the operation *Agree* if and only if:

- (12)
- B is in the *search space* of A (is c-commanded by A and is appropriately local)
 - A and B contain at least one pair of matching features; and
 - both A and B are active

On the application of *Agree*, any unvalued matching features on either category are assigned the value of its counterpart. It is further a property of the functional categories T and *v* they they can assign a (structural) case (nominative and accusative, respectively) on the application of *Agree*. We illustrate *Agree* with the following example:²

- (13) Ona widzi kota.
‘She sees a cat.’



The functional category *v*, associated with transitive constructions, Agrees with the direct object *D* (Determiner Phrase), because each is active (with unvalued features) and the two categories contain matching features (for gender and number). On the application of *Agree*, the values masculine and singular are assigned to the corresponding unvalued features for gender and number on *v*, and the value accusative is assigned to the unvalued case feature on the direct object. *Agree* applies in similar fashion between the T (Tense), responsible for predicate agreement, and the subject *D* (merged as Specifier-of-*v*, to raise to Spec-of-T under conditions we need not discuss

²The feature [person] is ignored here throughout.

here). The result is that the case feature on D is assigned the value nominative, and values for gender and number are copied from this subject D to T.

We now apply this mechanism in straightforward fashion to account for the morphosyntax of Polish numeral phrases. We begin with the heterogeneous morphosty-nax of higher numeral phrases in a direct case position, as in:

- (14) Widzimy osiem dziewczynek.
- | | |
|---------------------------------------|---|
| [case: ?]
[gender:]
[number:] | [case:]
[gender: fem]
[number: pl] |
| ↑—————↑ | |
- ‘We see [eight_{ACC=NOM} girls_{GEN.PL}]_{ACC}.’

The operation Agree can apply between the numeral phrase and its complement: each is active (containing unvalued features) and the two categories contain matching features. Numerals do not morphologically distinguish number (because they are always plural), nor does the gender value [feminine] have any morphological effect, so that the application of Agree in this example results in no overt agreement of the numeral with its head noun in this case.

How can we account for the genitive case on *dziewczynek* ‘girls’? One possibility would be to assume that higher numerals are associated with the genitive case and this feature is copied onto the noun it modifies. But this cannot be correct: if the numeral is marked with the genitive, it should take the genitive form itself. But that is ungrammatical:

- (15) *Widzimy [ośmiu dziewczynek]_{ACC}.
‘We see [five_{GEN} girls_{GEN.PL}].’

In order to obviate this problem, we assume an ‘abstract’ *quantitative* case, spelled out via syncretism with the genitive. That is, the numeral is associated in the lexicon with the valued case feature [case: quant(itative)]; this value is copied under Agree onto the modified noun. There is a simple morphological syncretism rule for Nouns and Adjectives replacing the quantitative case by the genitive case ([case: quant] → [case: gen]); the latter case is then spelled out as usual. As argued in more detail in Rappaport 2003, this syncretism should not be shocking: it has its analogy in the expression of animacy in Russian, exclusively in the form of a syncretism of the accusative with the genitive. The result should be no more mysterious than the appearance of the ‘genitive’ modifier in *Znam twojego kolegę* ‘I know your_{GEN} colleague_{ACC}’. We will demonstrate below that higher numerals in Polish actually have no nominative case form. The numeral form seen in (14) is the result of a direct spell-out rule of the quantitative case, adding a null desinence to the stem of the nominal: [case: quant] → ∅.

The homogenous morphosyntax of higher numeral phrases in oblique case positions is more straightforward. The numeral acts like a modifier because it is one: the case feature is unvalued, along with its gender and number features, as is the case for attributive adjectives. Thus, the availability of the two morphosyntactic patterns

reduces to a single, simple lexical fact: numerals have an optional value for case. The lexically-valued case feature [case: quant] gives heterogeneous morphosyntax; lexically unvalued case feature [case:] gives homogeneous morphosyntax. This is our answer to question (11a).

We now turn to a second question: how to correlate each of these patterns with the external context. The goal is to derive the following descriptive generalization:

- (16) a. A valued case feature (giving heterogeneous morphosyntax) is found in direct case positions;
b. An unvalued case feature (giving homogeneous morphosyntax) is found in oblique case positions.

In order to derive these generalizations, we utilize the independently-motivated distinction of two mechanisms of case assignment:

- (17) a. *Inherent* (‘lexical’) *case assignment* is selection, a condition on Merge (the insertion of lexical items into syntactic structure). A lexical item is visible to Merge if associated in the lexicon with the case feature required by a governing lexical item, in association with its semantic role assigning properties. The operation Agree is not involved.
b. *Structural* (‘syntactic’) *case assignment* applies to a lexical item which has been Merged with an unvalued case feature; a value is assigned in the syntax by Agree, independently of any lexical or semantic properties of either category.

It follows that while a noun has inherent values for number and gender, it may or may not have an inherent value for case. It must have one, however, in order to be visible to selection. In similar fashion, numerals optionally have a case value on Merge, in contrast, they can have no inherent values for the other features.

We begin with deriving the homogenous morphosyntax of oblique case positions, as in (2), in which the verb requires that its complement be in the instrumental case. A properly cased numeral phrase (with head noun lexically specified as [case: Inst]) satisfies the selectional requirement of the verb and is visible to Merge. Merging instead an uncased noun phrase in this position would violate the selection requirement of the verb. This scenario is ruled out by the assumed mechanism of inherent case assignment: a phrase headed by such a noun could not be selected because it would be invisible.

Now consider how heterogeneous morphosyntax results in a direct case position. In such a context, an uncased numeral would be active and would therefore Agree with an object in the accusative case (=nominative in this example, by syncretism for non-masculine personal plural nouns):

- (18) *Widzę [pięć dziewczyny]_{ACC}.
‘I see [five_{ACC=NOM} girls_{ACC=NOM.PL}].’

The result is ungrammatical. We rule it out morphologically: higher numerals in Polish have no nominative case form. In contrast, a numeral whose case feature is

valued as quantitative would invoke heterogeneous morphosyntax within the numeral phrase (cf. (14)); what looks like the nominative in such constructions is really the quantitative. Such a phrase is analogous to such quantified nominal phrases as *po dziewczynie* ‘a girl_{PREP} each’, *koło dwustu dziewczyn* ‘about two hundred_{GEN} girls_{GEN.PL}’, etc., in that the quantified noun is assigned case by a phrase-level mechanism, not by the clause-level mechanisms associated with T and v (cf. also the genitive of negation in the direct object position). It is a grammatical property of structural case that it need not be realized if a nominal phrase has an internal source of case, such as the quantifiers in the examples cited. Conversely, such quantified phrases are limited to direct positions, because they cannot satisfy the selectional requirements associated with oblique positions. This distinction in the two mechanisms of case assignment provides our answer to question (11b).

An interesting detail of Polish is that there are lexical items such as *dużo* ‘many’, *mało* ‘few’, *trochę* ‘little’, and *moc* ‘a lot of’ which behave like numerals in exhibiting heterogeneous syntax in direct case positions, but simply cannot appear in oblique cases. The explanation is a simple one in lexical terms: if the alternation of the two patterns is the result of the optionality of valuing the case feature, then the limited distribution of such items indicates that they are obligatorily specified for the feature [case:quant]. By not permitting an unvalued case feature, they rule themselves out of oblique positions.

Polish contrasts in an interesting way with Russian, which otherwise displays analogous behavior among its numeral phrases. Our discussion of the object position so far has been based on a non-masculine personal noun. In the case of a masculine personal noun, Russian and Polish differ:

- (19) a. Polish: *Widzę [pięciu chłopców]_{ACC}.*
 ‘(I) see [five_{GEN} boys_{GEN.PL}].’
 b. Russian: *Vižu [pjat’ mal’čikov]_{ACC}.*
 ‘(I) see [five_{NOM} boys_{GEN.PL}].’

At first glance, it appears that the contrast reduces to homogeneous morphosyntax in Polish versus heterogeneous morphosyntax in Russian. That is, in Polish the genitive form of the numeral represents the genitive accusative, the result of agreeing with the following noun, also in the genitive accusative. In Russian, since the numeral is not in the genitive, it cannot be agreeing with the genitive noun which follows; quite the contrary, the numeral is the source of that genitive.

But there are two indications that there is more going on here than this. First, if the Polish numeral phrase in (19a) is homogeneous, then why is its non-masculine personal analogue (18) (repeated here) ungrammatical?

- (18) **Widzę [pięć dziewczyny]_{ACC}.*
 ‘I see [five_{ACC=NOM} girls_{ACC=NOM.PL}].’

That is, if homogeneous morphosyntax is the result of an unvalued case feature, then why is an unvalued case feature possible on a masculine personal noun, but not on others? Second, why is it that in Polish we appear to see the same contrast in subject

position?

- (20) a. [Pięć dziewczyn] śpi na podwórzu.
 ‘[Five_{NOM} girls_{GEN.PL}] are sleeping in the courtyard.’
 b. [Pięciu chłopców] śpi na podwórzu.
 ‘[Five_{GEN} boys_{GEN.PL}] are sleeping in the courtyard.’

To invoke the genitive syncretism of the accusative for masculine personals to distinguish (19) from (18) would leave (20a)–(20a) completely unexplained. We are left with the following mysterious situation: in Polish (but not in Russian), an apparent alternation of homogeneous:heterogeneous morphosyntax is observed among direct positions, in both nominative and accusative positions, apparently correlated with gender.

Our approach here, as at several other points in the discussion, is to ensure that anomalous facts are placed in the lexicon and to minimize unmotivated syntactic alternations. Consequently, we will attempt to assimilate these Polish facts to the generalizations given in (16), for which we have already provided an account. That is, we claim that the Polish masculine personal numeral phrases in subject (20b) and object (19a) positions display heterogeneous morphosyntax (despite appearances) as much as do their non-masculine personal Polish counterparts (20a) and (14), as well as do their Russian analogues (e.g., (19b)). In all cases, the Polish numeral is associated with the valued case feature [case: quant]. The Russian rule spelling out quantitative case for higher numerals is maximally simple: a null desinence is appended to the stem of the numeral. In Polish, gender is a factor in the spell-out of the quantitative case in precisely the way it is for the accusative: in the presence of the feature [gender: masculine personal], the feature [case: quant] is replaced in the morphology as [case: gen], and spelled out as such. That is:

- (21) a. Russian: [case: quant] → ∅
 b. Polish:
 (i) [case: quant] → [case: gen] / __ [gender: masc. pers.]
 (ii) Elsewhere, [case: quant] → ∅

The Polish rule (21bii) mirrors the syncretism of the accusative plural with either the nominative or genitive. That Polish higher numeral phrases in nominative positions take the same form that they would in an accusative position is one of the arguments presented by Przepiórkowski (1999) in support what he calls the *accusative-subject hypothesis* of Polish numeral phrases. But our approach is more fine-grained: we are attempting to identify the source of the case marking and explain the phrase-internal facts. Within our analysis, we cannot associate higher numerals in Polish with the accusative, because that would not derive the morphological genitive for non-masculine personal nominals (as in (20a) and (14)). As we will see in section 4, this *is* the best analysis of lower numeral phrases, but the facts of higher numerals are different. Alternatively, it is tempting to state (21bii) as the replacement of the quantitative by the accusative, with the accusative on the numeral then subject to its general replacement by either the nominative or genitive in the plural. This option, alas, is not open to

us, as we have concluded that higher numerals have no nominative case form, so that the null desinence has to be associated directly with the quantitative case, as we have done in (21bii).

An interesting detail emerges here for free. On the one hand, we are treating (20b) and (19a) as exhibiting heterogeneous morphosyntax, analogous to (4) and (7), in which the quantified noun is the complement of a lexical noun. But that analogy is imprecise, indeed, at this point, seen to be inaccurate, because the heterogeneous morphosyntax of numeral phrases results from the application of Agree between the (cased) numeral and the following noun (the head of the numeral phrase), while in (4) and (7) the collective nouns *grupa* ‘group’ and *banda* ‘band’ truly are the heads of their respective phrases. So the analogy would not permit the numerals in (19a) and (20b) to have gender values of masculine personal, since that is the gender of the quantified noun, not the numeral. But on our analysis, under Agree, even as the numeral assigns case to the quantified noun, the numeral itself takes on the matching features of that noun. Later, in the morphology those features (including [gender: masculine personal]) are available to condition spell-out, as in (21bii).

4 Lower numerals

Lower numerals in oblique positions exhibit the homogeneous morphosyntax of the higher numerals:

- (22) Zachwyciłem się [dwoma miłymi dziewczynkami]_{INST}
 ‘I was captivated by [two_{INST} nice_{INST.PL} girls_{INST.PL}].’

The situation in direct case positions presents a more complex picture. Consider the following:

- (23) a. [Dwie miłe dziewczynki]_{NOM} śpią.
 ‘[Two_{NOM} nice_{NOM.PL} girls_{NOM.PL}] are sleeping.’
 b. [Dwaj mili chłopcy]_{NOM} śpią.
 ‘[Two_{NOM} nice_{NOM.PL} boys_{NOM.PL}] are sleeping.’
 c. [Dwóch miłych chłopców] śpi.
 ‘[Two_{GEN} nice_{GEN.PL} boys_{GEN.PL}] are sleeping.’

Examples in (23a) and (23b), for non-masculine personal and masculine personal plural nouns, respectively, appear to exhibit homogeneous syntax, in contrast to higher numerals: in subject position, the quantified noun stands in the nominative case (as opposed to the genitive). If this accurately represented the behavior of numbers such as ‘two’, ‘three’, ‘four’, and ‘both’, we would conclude that there are no lower numerals in Polish, that these numbers are in fact adjectival, like ‘one’. However, (23c), based on a masculine personal quantified noun, presents a different picture, analogous to that of the higher numerals. As in (19a) and (20b), we do not know *prima facie* whether we are looking at homogeneous syntax with a mysterious (genitive-)accusative assigned to subject position or heterogeneous syntax with a mysterious (genitive-accusative) numeral form.

It seems methodologically preferable to minimize the contrast between higher and lower numerals, so we opt for the latter interpretation, which is what we have arrived at for higher numerals. That is, the genitive of the quantified noun in (23c) is the result of the operation Agree applying between the numeral and that noun. This same methodological desideratum drives us to ask if there is any way to assimilate (23a) and (23b) to the pattern of the higher numeral phrase (20a) as well.

Let us begin with the paradigm represented by (23a) and (23c), representing non-masculine personal and masculine personal nominals, respectively. The problem is that non-masculine personal nouns following lower numerals stand in the nominative, whereas such nouns following higher numerals stand in the genitive. Thus, we cannot assume the same quantitative case is involved in both contexts, since we would have no reasonable conditioning for a syncretism with the nominative after some numerals, but with the genitive after others. We can propose, however, that lower numerals in Polish are optionally associated with a different value of the case feature. This case, very simply, is the accusative, which independently necessary morphological rules replace in the plural by the observed genitive for masculine plural nouns and nominative for others. This simple option was not available for higher numerals, because non-masculine personal nouns take the genitive form in subject and object positions (cf. (3b) and (14)). The difference in case value aside, the analysis developed above for higher numerals applies to the lower, thereby accounting for the interdependence of internal grammar and external context.

Let us return to (23b). This lower numeral construction has no counterpart among the higher numerals. All indications point to the simplest analysis: homogeneous morphosyntax in a direct case position. This was blocked among higher numerals by the absence of a nominative case form (cf. **pięc chłopcy śpią* ‘Five_{NOM} boys_{NOM.PL} are sleeping’). The fact that such constructions are not blocked for lower numerals indicates that these numerals have nominative case forms: *dwaj* ‘two’, *trzej* ‘three’, *cztery* ‘four’, and *obaj, obydwaj* ‘both’. There is no form corresponding to *dwaj* etc. in the accusative, because a masculine personal noun in the accusative is spelled out by syncretism with the genitive. This analysis automatically accounts for the impossibility of a numeral of the form **dwudziestu trzech* ‘twenty three’ as a case conflict: since *dwudziestu* ‘twenty’ is in the genitive, it can only be followed by a genitive form such as *trzech* ‘three’.

Again, Russian is different, more complicated, as it turns out, since lower numerals in this language trigger more opaque morphology. The noun following the numeral is spelled out as the genitive singular and the adjective can be genitive or nominative, depending upon the gender:

- (24) a. [Dva mal’čika (*mal’čiki)]_{NOM} spjat.
 ‘[Two_{NOM} boys_{GEN.SG} (*boys_{NOM.PL})] are sleeping.’
 b. [Moi dve žený (*žěny)]_{NOM} spjat.
 ‘[My_{NOM.PL} two_{NOM} wives_{GEN.SG} (*wives_{NOM.PL})] are sleeping.’

In this language, lower numerals are associated with an abstract case, but one distinct from the quantitative. The analysis is developed in Rappaport (2002, 2003), where

the case, following earlier work, is called the *paucal* case, whose spell-out is subject to fairly arcane conditions of variation. Polish is much simpler: the distribution of the genitive vis-à-vis the nominative after lower numerals conforms to the conditions of the genitive-accusative, so that we can leverage that rule without introducing the new machinery that is appropriate for Russian.

The analysis here has, we submit, interesting consequences for the theory of grammatical categories. The numeral represents a hybrid, sharing some properties with nouns, and others with adjectives. In particular, the essence of grammatical concord, that adjectives agree with nouns, follows from the fact that nouns have inherent values for gender and number, but adjectives do not (acquiring them from nouns under concord). In this respect, numerals are like adjectives. Nouns are assigned structural case in direct positions because they can be Merged in syntactic structure with unvalued case features, but they are visible to the selectional restrictions of individual lexical items if they are Merged with a valued case feature. Numerals share with nouns this optional case valuation. The consequences of this property for numerals and nouns differ, as has been detailed above, because of the difference in syntactic function: numerals are modifiers, while nouns are (typically) arguments. Polish has a set of quantificational expressions which behave like numerals in direct case positions, but, unlike the latter, cannot appear in oblique case positions. Since we attribute the possibility of numeral phrases in oblique case positions to an unvalued case feature, the distributional limitations on such words indicate that they are lexically (read, exceptionally) required to contain a valued case feature. These observations can be summarized as follows:

(25)

	Inherent values for gender and number	Inherent value for case
Noun	Yes	Yes/No
Numeral	No	Yes/No
<i>dużo</i> ‘a lot’, <i>malo</i> ‘a little’, ...	No	Yes
Adjective	No	No

We presented a working definition of the numeral in (5) in terms of the dual morphosyntactic patterns exhibited and have seen that these patterns result from the inherent feature properties summarized in (25). In both Polish and Russian we have found it appropriate to distinguish two categories of numerals: higher and lower, with the latter differing in its properties in the two languages. In particular, while all numerals are optionally associated with a case, we distinguish two cases that they can be associated with. Higher numerals in both Polish and Russian can be associated with the quantitative case, spelled out by syncretism with the genitive. Lower numerals in the two languages differ. We propose here that lower numerals in Polish may be lexically associated with the accusative. Reference was made to argument presented elsewhere that lower numerals in Russian are associated with another ‘abstract’ case like the quantitative, but distinct from it, which we have called paucal. The distinction between higher and lower numerals just discussed can be summarized as follows:

(26)

(Optional) inherent case	Polish	Russian
Higher	quantitative	quantitative
Lower	accusative	paucal

Quantitative case is replaced by the genitive on the noun and its adjectival modifiers. On the numeral, it is expressed directly in Russian by a spell-out rule, while in Polish, there is a syncretism rule for masculine personals (with the genitive); otherwise a direct spell-out rule as in Russian is employed. Paucal case in Russian is spelled out as genitive singular on the noun, and as genitive or nominative on adjectives.

While higher numerals in Polish do not have a nominative case form (what is normally considered that form is actually the direct spell-out of the quantitative case), lower numerals do, as seen in the contrast between (23b)–(23c). We present these two distinctions in Polish as independent of each other. It might seem a priori desirable to associate them, but we have no reason to do so, and assume that it is a historical coincidence, now encoded in the lexicon, that they co-occur.

5 Conclusion

The results of this paper may be summarized as follows:

- Numerals constitute a category intermediate between nouns and adjectives: numerals share with nouns the optional lexical valuation of the case, and while resembling adjectives in being associated in the lexicon with unvalued gender and number features.
- Because a numeral performs the syntactic role of a modifier, as opposed to the typical argument function of nouns, the effect of a numeral being cased on merger is different from that of a noun being so cased.
- Homogeneous morphosyntax is the result of the numeral being merged without lexically-specified case value; heterogeneous morphosyntax is the result of a lexically-specified case value. In the latter instance, the quantified noun receives its case value by the phrase-internal application of Agree between that noun (and its modifiers) and the numeral, in lieu of the clause-level application of Agree between that noun and a functional category.
- ‘Abstract’ case and a theory of syncretism and spell-out achieve the required effects. Quantitative case on a higher numeral accounts for heterogeneous morphosyntax. Lower numerals can be associated with the accusative case in Polish, in contrast with a paucal case in Russian.
- The optionality of case valuation interacts with the distinct case assignment mechanisms for inherent and structural case so as to determine the distribution of homogeneous and heterogeneous morphosyntax, linking the internal properties of numeral phrases with their external context.
- Some lexical peculiarities of Polish have been explored. There are quantitative expressions (e.g., *malo* ‘a little’) which resemble numerals, but their syntactic

distribution is limited, because their case feature is obligatorily valued. Also, higher numerals differ from lower numerals in that the former do not have a nominative form, while lower numerals do (e.g., *dwaj* 'two').

References

- Babby, Leonard H. (1987) "Case, prequantifiers, and discontinuous agreement in Russian". *Natural language and linguistic theory* 5, 91–138.
- Chomsky, Noam. (2000) "Minimalist inquiries: The framework". Roger Martin, David Michaels and Juan Uriagereka, eds. *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*. Cambridge, MA: MIT Press, 89–156.
- Chomsky, Noam. (2001a) "Beyond explanatory adequacy". Unpublished manuscript, MIT.
- Chomsky, Noam. (2001b) "Derivation by phase". Michael Kenstowicz, ed. *Ken Hale: A life in language*. Cambridge, MA: MIT Press, 1–52.
- Corbett, Greville G. (1993) "The head of Russian numeral expressions". G. G. Corbett, N. M. Fraser and S. McGlashan, eds. *Heads in grammatical theory*. Cambridge: Cambridge University Press, 11–35.
- Franks, Steven. (1995) *Parameters of Slavic morphosyntax*. New York: Oxford University Press (Oxford studies in comparative syntax).
- Halle, Morris. (1990) "An approach to morphology". *Proceedings of the Northeast Linguistic Society* 20, 150–84.
- Halle, Morris. (1995) "The Russian declension: An illustration of the theory of distributed morphology". Jennifer Cole and Charles Kisseberth, eds. *Perspectives in phonology*. Palo Alto: CSLI, 321–53.
- Mel'čuk, Igor'. (1985) *Poverxnostnyj sintaksis russkix čislitel'nyx vyraženij*. Vienna: Institut für Slawistik der Universität Wien (Wiener slawistischer almanac. Sonderband 16).
- Przepiórkowski, Adam. (1999) *Case assignment and the complement-adjunct dichotomy: A non-configurational constraint-based approach*. Doctoral dissertation, University of Tübingen.
- Przepiórkowski, Adam. (2003) "Podmiot liczebnikowy: Biernik czy mianownik?". Unpublished manuscript, available at <http://dach.ipipan.waw.pl/~adamp/Papers/>.
- Rappaport, Gilbert C. (2002) "Numeral phrases in Russian: A minimalist approach". *Journal of Slavic linguistics* 10, 329–42.
- Rappaport, Gilbert C. (2003) "The grammatical role of animacy in a formal model of Slavic morphology". Alan Timberlake and Alexander Schenker, eds. *American contributions to the Thirteenth International Congress of Slavists*. Bloomington: Slavica Publishers, 149–166.

Gilbert C. Rappaport
Department of Slavic Languages and Literatures
University of Texas at Austin
Austin, Texas 78713-7217
USA
grapp@mail.utexas.edu