Abstract • Streszczenie

Agreement and Case Assignment in Polish
An Attempt at a Unified Account

We present a unified HPSG account of two important phenomena in Polish, namely agreement and case assignment. The analysis of agreement is concerned with both NP-internal and subject-verb agreement, as well as with describing the exact repertory of agreeing categories in a non-redundant way. The analysis of case assignment deals with context-dependent case variation and with case assignment in numeral phrases. The interaction of these analyses leads to a successful account of such ill-behaved phenomena as third person singular neuter verb 'agreement' with numeral subject.

Uzgodnienie i nadawanie wartości przypadka
Próba jednolitej analizy

Przedstawiamy analizę w ramach HPSG dwóch ważnych zjawisk języka polskiego: uzgodnienia kategorii syntaktycznych i nadawania wartości przypadka. Nasza analiza uzgodnienia dotyczy związków wewnątrz fraz nominalnych, jak i uzgodnienia podmiotu z predykatem, oraz stanowi próbę określenia klasy kategorii syntaktycznych podlegających uzgodnieniu. Analiza nadawania wartości przypadka dotyczy natomiast przede wszystkim wahań wartości przypadka w zależności od kontekstu składniowego, oraz zagadnienia wartości przypadka fraz liczebnikowych. Wynikiem interakcji zaproponowanych rozwiązań jest m.in. analiza takich zjawisk jak 'zubożałe' uzgodnienie czasownika z liczebnikowym podmiotem.

Acknowledgements

Both authors would like to thank prof. Marek Świdziński for comments on the first version of the paper and Ania Kupść for helpful discussions. Adam Przepiórkowski would also like to thank Bob Borsley and Zelal Gungordu for helpful comments on a version of section 3. Needless to say, all remaining errors and omissions are our own.

Contents

1. Introduction

The phenomena of agreement and case assignment are ubiquitous across languages and receive much attention in various grammar formalisms. In this paper, a particular grammar formalism — Head-driven Phrase Structure Grammar — is considered. At the current stage of development, the HPSG framework lacks full-fledged theories of agreement and case. The general approach proposed in [PS94] offers a satisfactory account of agreement and case assignment in English which represents a rather uncomplicated case. Other proposals have appeared in the literature and solutions inspired by those put forward by [Katng], [HM94] and others regarding agreement and case assignment in German will be applied here to propose an account of agreement and case assignment in Polish.

1.1. Essential HPSG

This chapter presents the basic notions appearing in the remainder. We will also try to give additional explanations of HPSG basics in footnotes across the text. However, readers unfamiliar with feature structure logic and HPSG are strongly recommended to consult [PS94] ch. 1 and [Car92] for unknown definitions and facts.[1]

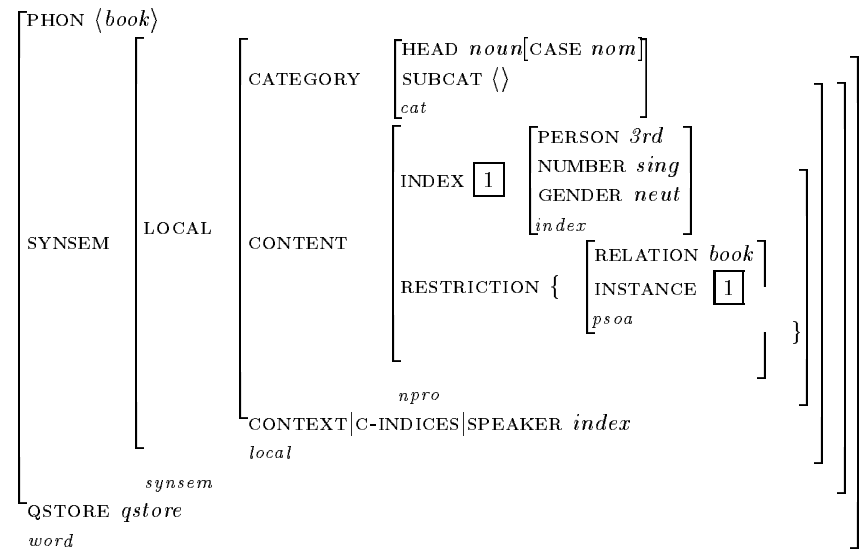
HPSG has been developed within the lexicalist and principle-based grammar traditions. It is characterized by a highly structured lexicon which interacts with a relatively small number of general principles. The lexicon is an inherent part of the grammar and — in the case of HPSG — it contains both syntactic and semantic information. General principles — Immediate Dominance Principle, Subcategorization Principle, Head Feature Principle, Control Theory, etc. — operate on this information in a declarative fashion, i.e., independent of the order in which they are applied. In general, the HPSG approach to language modelling can be described as constraint-based: both lexical entries and grammar principles[2] represent sets of constraints which must be met by well-formed linguistic expressions.

In HPSG, generalized, typed signs are the basic building elements of the grammar. Lexical entries, phrases and discourses are signs. To represent signs typed feature structures are employed. The commonly used AVM (attribute-value matrix) notation of feature structures will be applied in the remainder.

The feature structure associated with the sign corresponding to the word **book** is given below (in the formalism proposed in [PS94]):

[1] See also [MK94] for an introduction in Polish.
[2] With the exception of the Raising Principle.

(1)



In (1), the tag 1 appears as the value in SYNSEM|LOCAL|CONTENT|INDEX and also bears the tag 1. The same tag appears as the value in SYNSEM|LOCAL|CONTENT|RESTRICTION|INSTANCE, which exemplifies structure sharing that is extensively used in HPSG. The fact that two nodes are structure-shared means that they are token-identical. Structure sharing is among the features of HPSG which provide the framework with the power analogous to the power obtained by adopting mechanisms like movement in other formalisms, e.g., in GB ([Cow92], [Hae91]).

In HPSG, every sign has at least three attributes: PHON, SYNSEM and QSTORE. The value of the attribute PHON represents the unilateral part of the sign, i.e., the phonology, the orthography, etc., of the modelled word or phrase. SYNSEM (SYNTAX-SEMANTICS) contains the syntactic characteristics of the denoted word or phrase in the attribute LOCAL|CATEGORY and the respective semantic characteristics in the attribute LOCAL|CONTENT[3]. In the case of nominals, the attribute CONTENT contains the feature INDEX the value of which corresponds to a reference marker in DRT. The attribute LOCAL|CONTEXT contains information regarding discourse parameters, identifying the speaker, the hearer, the time and the place of the utterance, etc.; the information in the NONLOCAL attribute is used in the analysis of unbounded dependencies, wh-questions and relative questions; we do not deal with these phenomena in the

[3]We will use the simpler version of the attribute, i.e., without changes introduced in ch. 8 in [PS94].

article. QSTORE contains information about quantifiers and their scopes; this attribute is not applicable for most problems considered in the remainder. Additionally, phrasal signs, i.e., signs representing phrases, have the attribute DTRS in which the information about the daughters of the phrase, e.g., the head daughter HEAD-DTR and complement daughters COMP-DTRS, is stored.

As the name suggests, the notion of head plays the pivotal role in HPSG. In general, every phrase is supposed to have a head (i.e., one head per phrase[4]) which determines the syntactic and semantic properties of the phrase. In particular, the head contains syntactic and semantic specification of other objects of the grammar that can combine with the phrase as its complements or modifiers. The restrictions on the complements are stored in the feature CATEGORY|SUBCAT which contains specification of all complements of the given syntactic category ordered according to the increasing obliqueness, i.e., first the subject, then the least oblique complement and so on. In the newer versions of HPSG — due to work of Borsley — this feature corresponds to two features: SUBJ and COMPS, carrying information about the required subject of the phrase and its other complements, respectively. The value of the somewhat redundant SUBCAT attribute is simply a concatenation of the values of the attributes SUBJ and COMPS and is useful in formulating the Binding Theory. We will use the attribute SUBCAT as well as SUBJ and COMPS across the paper.

Information contained in the head of a phrase is percolated to the sign representing the phrase by means of the Head Feature Principle. This principle says that the HEAD value of any headed phrase is structure-shared with the HEAD value of the head daughter. Next to the Subcategorization Principle, which states that in a headed phrase the SUBCAT value of the head daughter is the concatenation of the phrase's SUBCAT list with the list of SYNSEM values of the complement daughters, the Head Feature Principle is among the Principles of Universal Grammar.[5] These principles as well as principles specific for a language constitute the grammar of the language.

By means of unification, feature structures representing syntactic categories can be combined to obtain other syntactic categories. This process occurs according to Immediate Dominance Schemata. Syntactic categories occurring in Immediate Dominance Schemata and other principles are underspecified which allows for a general formulation of principles.

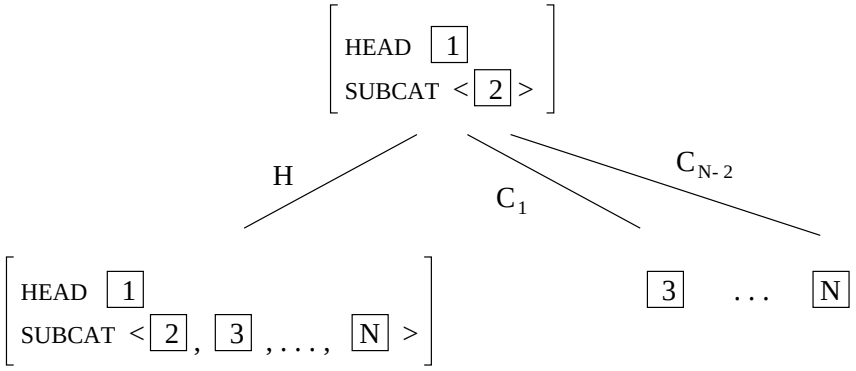
There are a few ID schemata used in HPSG. The result of the application of Schema 2 together with the Subcategorization Principle and the Head Feature Principle is presented in (2). The root of the tree represents the CAT attribute

[4]See also [PG87] for their multi-headed analysis of coordination.

[5]Head Feature Principle and Subcategorization Principle are mentioned only as examples of principles operating in HPSG.

of the feature structure corresponding to a phrase licensed by the schema. The branches correspond to the CAT attributes of the signs stored in the DTRS attribute of the phrasal sign: branch H to the HEAD-DTR sign, branches C_1, ..., C_{N-2} to the signs on the COMP-DTRS list.

(2)



Example (5) (p. 8) depicts the sign corresponding in HPSG to the phrase: "Kim walks". This sign has been obtained on the basis of Schema 1 (not presented here), Schema 2, Subcategorization Principle and Head Feature Principle (only part of the information contained in the full sign is shown here).

2. Agreement

2.1. Introduction

In the first part of the paper we will concentrate on agreement. The notion of agreement encompasses a broad range of phenomena found in many languages. Agreement can be shortly characterized as follows: a grammatical object (e.g., a syntactic category) X agrees with a grammatical object Y in property (agreement feature) P if both X and Y bear the same obligatory marking for P. Consider the following examples from Polish:

(3) a. *mała* *dziewczynka*
small$_{fem}$ girl$_{fem}$
'a small girl'

b. *mały* *chłopiec*
small$_{masc}$ boy$_{masc}$

'a small boy'

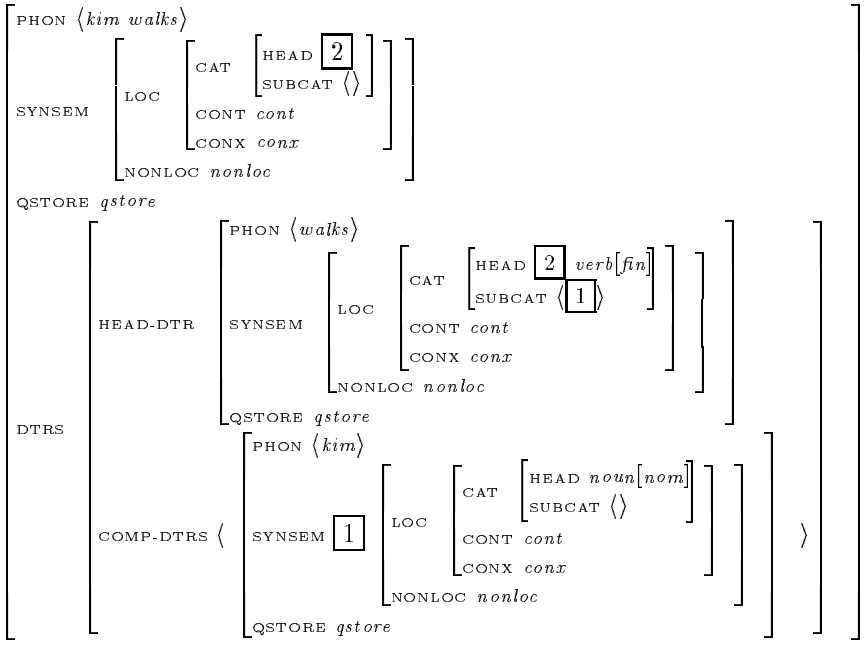
c. **mała* *chłopiec*
small$_{fem}$ boy$_{masc}$

(4) a. *Anna* *idzie.*
Anna$_{sing}$ go$_{3rd,sing}$.
'Anna goes'.

b. **Anna* *idą.*
Anna$_{sing}$ go$_{3rd,plur}$.

In (3a) and (3b) the noun and the respective attributive adjective agree in gender. This is not the case, however, in (3c), which makes the phrase ungrammatical. Similarly, in (4a), the subject agrees with the verb in number; the agreement does not hold in (4b) and the sentence in ungrammatical.

(5)



As in most Slavonic languages, rich agreement is characteristic of Polish and the above examples present only its simplest instances. In our attempt to characterize agreement in Polish we will follow the general schema suggested

by [FB88]. First, we will try to outline the domain of agreement in Polish, i.e., to state what grammatical elements agree in what kinds of grammatical configurations. We will describe the agreement features relevant for Polish. This will be influenced by a nondirectional approach to agreement we will advocate following [FB88]. We will also propose a way of avoiding the syntactic/semantic agreement conflict which occurs in personal sentences in the contemporary Polish. Finally, we will suggest a partition of simple sentences in Polish and a way of accounting for some special cases of agreement feature marking of verbal forms.

We start with a short outline of some agreement theories, then proceed to the original theory proposed by [PS94]. This is followed by a revision due to Kathol ([Katng]) and our discussion of agreement features and agreement in Polish.

2.1.1. Agreement in Other Linguistic Theories

Various linguistic theories addressed agreement phenomena. Both semantic (e.g., [DJ89]) and syntactic (e.g., in GB, cf. below) approaches have been adopted.

Semantic approaches have not offered satisfactory accounts (see, e.g., [PS94] for critical remarks on [DJ89]). Accounting for the antecedent-pronoun agreement in sentences like:

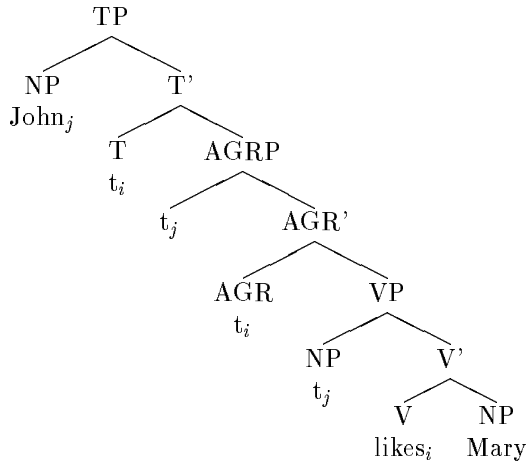
(6) a. *The dog*$_1$ *is so ferocious, it*$_1$ *even tried to bite itself*$_1$.

b. * *The dog*$_1$ *is so ferocious, it*$_1$ *even tried to bite himself*$_1$.

proved most troublesome. As a dog can be referred to by either *he* or *it*, the required matching of the pronouns and reflexives *he–himself* or *it–itself* is difficult to obtain in a purely semantic way.

In the derivational framework of GB, the phenomenon of subject-verb agreement is modelled by proposing the existence of a separate node INFL which dominates tense and all verbal inflection (in English; see [Hae91] p. 101). The INFL node of a finite clause is assumed to be able to assign (under government) nominative case to its subject. Somewhat different accounts (like the one proposed by Pollock in [Pol89], see also [Cow92] p. 174) involve splitting INFL node in two: T bearing tense and AGR bearing agreement. The theories assume that (in English, for nonauxiliary verbs) there is affix movement from AGR position to V position where the verb receives inflection and from T position to V position where the verb receives tense.

(7)



The verb itself does not carry any agreement information — no agreement information is specified in lexical entries of verbs. The agreement process is directional — from the subject to the verb — and it involves "movement of bundles of agreement features from a nominal onto something that agrees with the nominal" ([PS94]).

However, this kind of approach is faced with (at least) proliferation problems when applied to Polish (see also [BF88]). Consider:

(8)

a. *Ja byłem szczęśliwy.*
I was$_{sing,masc}$ happy$_{sing,masc}$.
'I was happy.'

b. *Ja byłam szczęśliwa.*
I was$_{sing,fem}$ happy$_{sing,fem}$.
'I was happy.'

c. *Ty byłeś szczęśliwy.*
you were$_{sing,masc}$ happy$_{sing,masc}$.
'You were happy.'

d. *Ty byłaś szczęśliwa.*
you were$_{sing,fem}$ happy$_{sing,fem}$
'You were happy.'

As (8) shows, if agreement information is to be copied as a bundle of features, multiple lexical entries for personal pronouns in Polish are needed. Similarly, the data in (9) requires that multiple null elements differing in the values

of the respective agreement features be proposed in order to account for the so-called pro-drop behaviour of Polish:

(9) a. *Byłem* *szczęśliwy.*
was$_{1st,sing,masc}$ happy$_{sing,masc}$
'I was happy.'

b. *Byłam* *szczęśliwa.*
was$_{1st,sing,fem}$ happy$_{sing,fem}$
'I was happy.'

c. *Byłeś* *szczęśliwy.*
were$_{2nd,sing,masc}$ happy$_{sing,masc}$
'You were happy.'

d. *Byłaś* *szczęśliwa.*
were$_{2nd,sing,fem}$ happy$_{sing,fem}$
'You were happy.'

Thanks to underspecification, the theory of agreement adopted below gives a more satisfactory account of these facts.

2.1.2. Pollard and Sag's Theory of Agreement

In this section we will present a short account of the agreement theory as proposed in ch. 2 of [PS94] which is the original theory of agreement in the HPSG tradition. HPSG is set within the unification-based framework. This facilitates a nondirectional account of agreement, which is the main difference in comparison with derivational approaches.[6] The nondirectionality is, however, limited to partial specification of agreement features on the agreement target (the argument) and on the SUBCAT list of the agreement source (the selector). The correct account of agreement is obtained thanks to the requirement that the partial information in both syntactic categories has to be compatible, i.e., unifiable.

Consider the lexical entry for the verb *walks* in (10). In this entry, there is no specification of agreement feature markings on the verb. The verb requires, however, that the potential argument (subject) is nominative and its index contains appropriate person and number information. Thus, in this approach, "to be a third singular verb is nothing more than to assign third-singular agreement to the index associated with one's subject" ([PS94], p. 86).

[6] Compare, however, the next section where an even "more nondirectional" approach is advocated.

(10)
$$\left[\begin{array}{ll} \textsc{phon} & \langle walks \rangle \\ \textsc{synsem|local} & \left[\begin{array}{ll} \textsc{cat} & \left[\begin{array}{l} \textsc{head}\ [\textsc{vform}\ fin] \\ \textsc{subcat}\ \langle \mathrm{NP}[nom]\boxed{1} \rangle \end{array}\right]_{category} \\ \textsc{cont} & \left[\begin{array}{l} \textsc{relation}\ walking \\ \textsc{walker}\ \boxed{1}\ \left[\begin{array}{l} \textsc{person}\ 3rd \\ \textsc{number}\ sing \end{array}\right]_{index} \end{array}\right]_{content} \end{array}\right]_{local} \end{array}\right]_{sign}$$

In (10), the index tagged with $\boxed{1}$ mediates in the agreement relation providing an instance of index agreement. In index agreement, the settings of features on the index of the subject must be equal to the settings required by the verb. This is obtained by structure-sharing of the index of a respective NP on the verb's SUBCAT list and the index corresponding to a semantic (thematic) role introduced by the verb.

It should be remarked here that indices play a double role in HPSG: on the one hand, indices provide an "interface" to extralinguistic properties of a referent by means of anchoring conditions, which is clearly a semantic function. On the other hand, they mediate in person and number agreement and play a crucial role in the Binding Theory, which is a syntactic function. Thus the account of agreement is of mixed origin and it makes use of both semantic and syntactic knowledge about linguistic objects.

In particular, indices participate in antecedent-pronoun agreement which gives a correct analysis for (6) where the indices on the noun *dog*, pronoun *he/it* and reflexive *itself/himself* must be the same (as a result of the Binding Theory, [PS94] ch. 6) leading to the desired grammaticality of (6a) and ungrammaticality of (6b). As mentioned above, this was difficult to obtain in a purely semantic account of agreement.

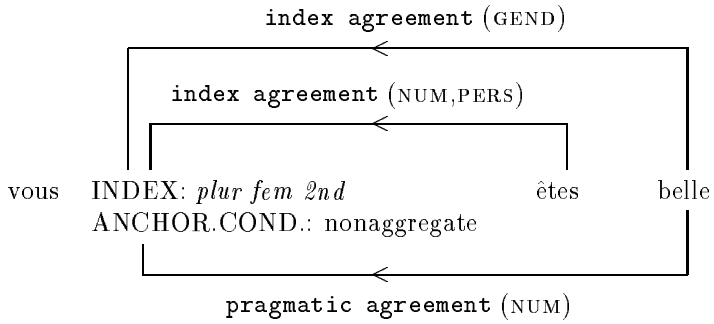
A more complicated picture is proposed by [PS94] to account for agreement in sentences like the following:

(11) *Vous* *êtes* *belle.*
you$_{plur}$ are$_{plur}$ beautiful$_{sing,fem}$
'You are beautiful.'

In (11), number of the predicative adjective *belle* — singular — does not agree with number of its subject *vous* — plural. The singular form of the adjective reflects the fact that the referent of *vous* is a nonaggregate entity.

This entity is referred to by a morphosyntactically plural pronoun expressing respect. Also, there is no direct agreement in gender between the subject *vous*, which is applicable for any gender, and the predicative adjective *belle*, which is feminine. The explanation [PS94] offer is "hybrid agreement" depicted in (12) ([Katng]). Firstly, the index of *vous* is assumed to bear the feminine and 2nd person markings. The number setting on the index is plural but the usual condition is relaxed and the pronoun is allowed to designate a nonaggregate entity. Secondly, the agreement between the subject and the copula is assumed to be indexical (as above for English), which explains the 2nd person plural form of the copula. Also indexical agreement with respect to gender obtains between the subject and the predicative adjective rendering the latter feminine. Thirdly, pragmatic agreement with respect to number between the subject *vous* and the predicative adjective *belle* is assumed to obtain. Pragmatic agreement makes a reference to anchoring conditions: as the entity denoted by *vous* is nonaggregate, the form of the predicative adjective is singular.

(12)



Finally, the "concord" type (syntactic type) of agreement is discussed in [PS94] and exemplified by NP-internal case agreement in German. Case concord is distinguished from index agreement as coindexed elements do not have to show agreement in case; thus, case is not part of the index structure and case concord is taken care of separately.

(13)

$$\begin{bmatrix} \text{PHON } \langle \textit{Mädchen} \rangle \\ \text{SYNSEM|LOCAL} \begin{bmatrix} \text{CATEGORY} & \begin{bmatrix} \text{HEAD } \textit{noun}[\text{CASE } \boxed{2} \neg \textit{gen}] \\ \text{SUBCAT } \langle\ \text{DET}[\text{CASE } \boxed{2}]\boxed{1}\ \rangle \end{bmatrix} \\ \text{CONTENT} & [\text{INDEX } \boxed{1}] \end{bmatrix} \end{bmatrix}$$

In (13) case of the determiner and case of the noun are forced to be identical by means of the constraint imposed by structure-sharing depicted by the tag $\boxed{2}$. In the concord type of agreement both syntactic categories bear the same value of an attribute (CASE), which follows closely the commonly accepted nondirectional view of agreement (see [BF88]).

Although we will not consider morphology in this paper, it is worth mentioning that the theory of [PS94] does not offer any special treatment of the morphology-syntax interface. The morphology-syntax interface is very closely related to agreement as in many languages agreement feature marking results in a marked, inflected form of a lexeme. In [PS87], the third person singular morphology is accounted for by means of a lexical rule[7] ([PS87], p. 213):

(14)
$$\begin{bmatrix} \textsc{phon} \boxed{1} \\ \textsc{3rdsng} \boxed{2} \\ \textsc{syn|loc|subcat} \boxed{3} \\ \textsc{sem|cont} \boxed{4} \\ \textit{base} \end{bmatrix} \mapsto \begin{bmatrix} \textsc{phon} \; \mathrm{f}_{3rdsng}(\boxed{1}, \boxed{2}) \\ \textsc{syn|loc|subcat} \boxed{3} \\ \textsc{sem|cont} \boxed{4} \\ \textit{3rdsng} \end{bmatrix}$$

f_{3rdsng} is assumed to be a morphological operator which returns the correct 3rd singular form of a verb, i.e., either the irregular form like in *has* or the regular form like in *reads* depending on the verb. The restrictions imposed on the sort *3rdsng* look as follows:

(15)
$$\begin{bmatrix} \textsc{syn|loc} & \begin{bmatrix} \textsc{head} \; [\textsc{vform} \; \textit{fin}] \\ \textsc{subcat} \; \langle \; \mathrm{NP}[\textit{nom}] \; \begin{bmatrix} \textsc{pers} \; \textit{3rd} \\ \textsc{num} \; \textit{sing} \\ \textit{index} \end{bmatrix} \; \rangle \\ \textit{local} \end{bmatrix} \\ \textit{3rdsng} & \end{bmatrix}$$

Such a simple approach cannot be applied to Polish. A more systematic account is required in order to capture effectively much more complicated paradigms of Polish verbs, nouns and adjectives.

2.1.3. Kathol's Approach to Agreement

The theory shortly presented above suffers from some deficiencies of both conceptual and empirical nature. We will state them here following [Katng] and, where relevant, support the criticism with data from Polish (see [Katng] for a detailed discussion).

[7] The rule presented here uses the old formalism of [PS87].

As we pointed out above, the framework of HPSG is well-suited for a nondirectional approach to agreement in which both the agreement selector and target show agreement feature marking. This possibility is not exploited by [PS94] in full as, e.g., verbs lack feature marking of this kind. The picture of subject-verb agreement resulting from the theory is closer to government (the verb assigns certain features to the subject) rather than agreement (both the verb and the subject show the same marking on agreement features).

The original theory cannot explain the fact that many languages exhibit agreement in features of the selecting and selected category which is not mediated by indices and is not an instance of case concord, e.g., in Polish:

(16)

te	*dwie*	*młode*
these$_{nom,plur,fem}$	two$_{nom,fem}$	young$_{nom,plur,fem}$
kobiety		
women$_{nom,plur,fem}$		
'these two young women'		

In (16), the agreement in gender, number and person occurs parallel to the agreement in case. As NP-internal agreement is not assumed to be mediated by indices, the covariation remains unexplained.

According to [Katng], impersonal passives in German constitute a problematic case for the original theory: as there seems to be no evidence of the existence of an appropriate null element functioning as the subject in the following example:

(17)

An jenem Abend	*wurde*	*viel*	*gelacht.*
during that evening	was$_{3rd,sing}$	much	laughed.
'There was much laughter that evening.'			

it is difficult to explain the 3rd person singular form of the auxiliary verb *wurde*.[8]

A similar problem cause meteorological verbs in Polish. They occur on their own, i.e., without an overt (expletive) subject, showing the 3rd singular neuter agreement:[9]

(18) a. *Grzmiało.*
thunder$_{past,3rd,sing,neut}$

[8] An alternative account could assume that the form *wurde* bears special agreement markings different from 3rd singular; similar proposals have been put forward for analogous phenomena in Polish.

[9] Alternative accounts have been proposed which do not assume 3rd singular neuter agreement on finite forms of meteorological verbs; we will return to this problem in 2.4.4..

'It thundered.'

b. *Mżyło.*
drizzle$_{past,3rd,sing,neut}$
'It drizzled.'

As there is no (overt) subject, the source of this marking is unclear. We think that proposing null elements for sentences like these in (18) is rather unmotivated.[10] According to [PS94], the verb morphology reflects certain feature settings on the subject. This makes an analysis assuming an empty subject in (18) impossible. We will, however, argue for such an analysis further.

Another question is raised by impersonal sentences in Polish exemplified below:

(19) *Mówiono, że będzie zimno.*
it was said that be$_{future}$ cold
It was said that it would be cold.

Again, the impersonal (the verbal form ending in *-no* or *-to*) always occurs without an overt subject. One way of solving the problem is to propose an empty element, which we would rather avoid (see above). Thus, the whole class of sentences with impersonals remains unaccounted for.

The quirky behaviour of Polish numerals and subject-verb agreement where the subject is a numeral phrase are certainly among the most difficult questions Pollard and Sag's theory is faced with. The reader is referred to section 3.3. for the discussion. In short, we think that in order to be able to account for case assignment by Polish numerals the case theory should be more fine-grained than it is proposed by [PS94].

The pragmatic agreement between the subject and the predicative adjective proposed by [PS94] and illustrated in (12) will not be sufficient for the following Polish data:

(20) a. *Pięciu mężczyzn było wysokich.*
five men was$_{sing,3rd,neut}$ tall$_{plur,gen}$
'Five men were tall.'

[10] One argument against it would be the potential need of introducing a null element different from the one required to account for pro-drop behaviour of Polish; this element would have to be specified in the way which would prohibit its occurrence in pro-drop sentences and, vice-versa, the null element occurring in pro-drop sentences would not be allowed in sentences like in (18). In practice, this would mean that expletive pronouns/indices would be required for Polish, for which we have found no other evidence so far.

b. *Dwaj mężczyźni byli wysocy.*
Two men were $\text{tall}_{plur,nom}$.
'Two men were tall.'

where the genitive case marking in the form of the predicative attribute occurs depending on the subject. In (20b), certain (rather rare) pragmatic and stylistic restrictions allow also for the instrumental form of the predicative adjective. However, the choice in (20a) with the numeral *pięć* ('five') is restricted to genitive. In our opinion, this fact should be handled by syntax rather than pragmatics, which is not possible in the [PS94] proposal.

A closer look at NP-internal agreement (or any other kind of agreement, for that matter) will lead to the conclusion that certain language-specific patterns of agreement exist that determine which features participate in agreement and what kind (or kinds, i.e, indexical, concord or pragmatic) of agreement a certain syntactic category can enter. For example, NUMBER, CASE and GENDER are relevant for NP-internal agreement in Polish; these features are not relevant for NP-internal agreement in English. The current theory does not provide a single place where abstract types of agreement capturing such regularities could be defined, which leads to redundancy.

Some of the above problems are solved by the account that follows which is based on the solution put forward by [Katng]. The basic difference in comparison with [PS94] is the presence of a new head feature called AGR which is represented by the respective attribute in the HEAD attribute of every lexical and phrasal category. AGR contains agreement information relevant for a given syntactic category, i.e., the attributes which can participate in any instance of agreement the category can enter. A similar solution, shortly described in the section 2.1.1. above, has been also introduced in GB. We emphasize, however, that the analogy is very remote as there is no movement involved and the account of agreement here is nondirectional.

Some instances of agreement can be modelled by means of structure sharing of values of relevant attributes contained in AGR. This kind of agreement is referred to as morphosyntactic agreement. In (16), repeated here as (21), the covariation in gender, case and number can be easily explained by means of morphosyntactic agreement.[11]

(21) *te* *dwie* *młode*
$\text{these}_{nom,plur,fem}$ $\text{two}_{nom,plur,fem}$ $\text{young}_{nom,plur,fem}$
kobiety
$\text{women}_{nom,plur,fem}$

[11] From this point, the subscripts in glosses represent the values of attributes in AGR.

'these two young women'

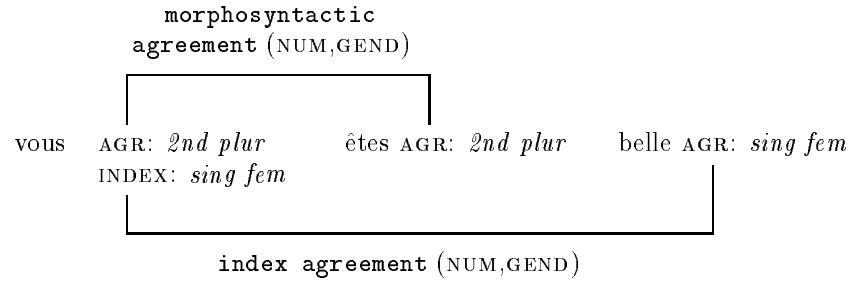
In (21), it can be assumed that the features CASE, NUMBER and GENDER collected in AGRs of all the syntactic categories are structure-shared, which accounts for the observed agreement.

An analogue of the indexical pattern of [PS94] is also present in Kathol's solution. Indices retain their semantic function and their role in antecedent-pronoun agreement. Due to their semantic function, indices provide an interface to anchoring conditions. This fact can be used in order to avoid the pragmatic agreement in sentences like (11) repeated here as (22):

(22) *Vous* *êtes* *belle.*
you are$_{plur}$ beautiful$_{fem,sing}$
'You are beautiful.'

Thanks to the clustering of agreement features in AGR and separating them from indices bearing semantic information an entirely new picture emerges. The new agreement structure can be depicted as follows ([Katng]):

(23)



In (23), the feature settings on the INDEX and AGR of *vous* are different. The settings on the INDEX reflect the way in which the referent is individuated in the discourse: here *vous* denotes a nonaggregate entity, which is "stored" on the INDEX and results in the *singular* value of NUMBER. The settings on the AGR of *vous* reflect the fact that morphosyntactically *vous* is a plural nominal; hence the *plural* value of NUMBER in AGR. The required agreement between the predicative adjective *belle* and the subject *vous* is easily explained by means of indexical agreement: the values of the attributes PERSON and GENDER on AGR of the adjective are structure-shared with the respective values on the INDEX of the subject. In (12), the same result was achieved by means of the

rather unclear pragmatic agreement. Simultaneously, morphosyntactic agreement obtains between the subject and the copula *êtes* with respect to number and person.

(23) illustrates the dichotomy of agreement which offers a flexible framework in which to account for various agreement patterns occurring across languages. The agreement patterns partition can be stated as follows:

(24)
1. morphosyntactic agreement: AGR(selector) $\approx$ AGR(arg)
2. semantic (index) agreement: AGR(selector) $\approx$ INDEX(arg)

where $\approx$ denotes structure sharing of relevant parts, i.e., the features that participate in a given instance of agreement. In this way the role of pragmatics in the grammar has changed: pragmatics determines the use of certain kinds of agreement but does not intervene in the agreement structure.

In this formulation the theory of agreement clearly differs from government. Agreement information is specified (at least partially) on both categories entering agreement, either on AGR or INDEX. The phenomenon of agreement is non-directional as it is based on structure sharing of attributes present on both categories and not on selectional restrictions as it is in [PS94]. The only directional "trait" here is the fact that it is the index of the argument that plays the role in agreement and not the index of the selector.[12]

A more general and systematic approach to indices emerges from Kathol's proposal. Consider the French *vous* again. Depending on the situation it can be anchored to a singular entity, namely when it expresses politeness, or to a multiple entity. Gender of the anchor can also vary: it can be either feminine or masculine. This gives rise to the following INDEX and AGR structure for *vous*:

(25)
$$\begin{bmatrix} \ldots \text{ AGR} & \begin{bmatrix} \text{NUMBER } plur \\ \text{PERSON } 2nd \end{bmatrix} \\ \ldots \text{ INDEX} & \begin{bmatrix} \text{NUMBER } number \\ \text{GENDER } gender \end{bmatrix} \end{bmatrix}$$

The attributes on INDEX remain unset, which illustrates the fact that the index features for *vous* are determined by anchoring conditions. Within this approach there is no need to assume that polite plurals can be anchored to nonaggregate entities and bear plural indices. In contrast, the lexical entry

[12] We are not convinced that the partition in (24) represents all kinds of agreement which are possible in natural language. It is sufficient to account for all instances of agreement discussed so far in the literature but there seems to be no general constraint which would prevent other combinations of AGR and INDEX of the selector and argument to appear as possible agreement patterns.

of a typical French noun can contain the following data where the respective attributes on INDEX and AGR are the same:

(26)
$$\begin{bmatrix} \ldots \text{ AGR} & \begin{bmatrix}\text{NUMBER } \textit{sing}\\ \text{PERSON } \textit{3rd}\\ \text{GENDER } \textit{fem}\end{bmatrix} \\ \ldots \text{ INDEX} & \begin{bmatrix}\text{NUMBER } \textit{sing}\\ \text{PERSON } \textit{3rd}\\ \text{GENDER } \textit{fem}\end{bmatrix} \end{bmatrix}$$

Here, the index settings are derived from the morphosyntactic features of the noun.

Thus, both anchoring conditions and morphosyntax can influence index features, which is depicted in (27) ([Katng]):

(27) $\boxed{\text{anchoring conditions}} \rightarrow \boxed{\text{index}} \leftarrow \boxed{\text{morphosyntax}}.$

Which factors "win" in various syntactic categories is language-specific. As can be seen from the above, in French anchoring conditions are crucial for polite pronominals whereas morphosyntax is to be taken into account for other nouns.

Let us mention another very desirable consequence of introducing the AGR feature. In the current formulation of the agreement theory, verbs carry agreement features of their own. This fact can be useful for an explanation of the occurrence of subjectless verbs without any need for multiple empty elements: even when a verb occurs without a subject, it can bear agreement feature marking; which agreement feature marking is appropriate for subjectless verbs should be independently predicted by the grammar. This will provide a way to account for sentences like (18) and (19).

The clustering of agreement information in AGR allows also for defining abstract agreement types capturing regularities like the above mentioned agreement of PERSON and NUMBER in English subject-verb agreement. HPSG subsorts, in turn, provide a means necessary to capture the resulting agreement abstract types and hierarchical dependencies among them.

Finally, [Katng] proposes a more flexible approach to the morphology-syntax interface than [PS87] and [PS94]. In general, the interface is similar to this proposed by [PS87] which we have seen above for the 3rd singular verbs in English and it is realized as a paradigm function. In short, a paradigm function PF takes two arguments which are the root of the paradigm and the complete and fully specified matrix of morphosyntactic features associated with a certain form. The morphology corresponding to this form is obtained as the result of the paradigm function. The root of the paradigm is taken from the feature STEM characteristic of the subtype *morph-complex* of the type *sign*.

This feature is a record of all steps which are required in order to derive a certain form from more basic elements. The reader is referred to [Katng] for details.

In the remainder of this section, we will attempt to give an account of the agreement in Polish within the framework defined by the new theory.

2.2. Gender and Number in Polish

We begin our discussion with an overview of the repository of inflectional categories in Polish. Some of them can be assumed to behave "as expected", e.g., the category of person in Polish fully corresponds to the category of person in English and German. We will not discuss such categories in the remainder and concentrate rather on inflectional categories which require a special treatment in order to reduce the redundancy of the description or provide a correct account of empirical data. In this subsection we will concentrate on the categories of number and gender, section 3. is devoted to the inflectional category of case.

The complex structure of gender in Polish has been well accounted for in the linguistic literature ([Mań56], [SŚ85]). However, these accounts have mostly been set within the descriptive framework and did not offer suitable mechanisms in which to express the observed regularities in a compact and nonredundant way. HPSG offers subsorts and unification which we will use to propose a hierarchical structure of gender. Our classification will be based on, albeit different from, the one in [SŚ85]. The crucial difference will be the combination of gender and number information in one category.[13] This leads to a different partition of the sort *gender* and provides a way of removing a considerable number of multiple lexical entries from the lexicon.

Compare the following examples in which the noun phrase subcategorized for by the verb *lubić* ('like') is accusative (from [SŚ85]):

(28) a. *Lubię* *dobrego* *chłopca.*
like$_{1st}$ good$_{masc,sing,acc}$ boy$_{masc,sing,acc}$
'I like a good boy.'

b. *Lubię* *dobrego* *psa.*
like$_{1st}$ good$_{masc,sing,acc}$ dog$_{masc,sing,acc}$
'I like a good dog.'

c. *Lubię* *dobry* *stół.*
like$_{1st}$ good$_{masc,sing,acc}$ table$_{masc,sing,acc}$

[13] As M. Świdziński has pointed out to us, an idea of combining the gender and number information, although in a much simpler form, is also proposed in [Świ92].

'I like a good table.'

d. *Lubię* *dobrą* *dziewczynę.*
like$_{1st}$ good$_{fem,sing,acc}$ girl$_{fem,sing,acc}$
'I like a good girl.'

e. *Lubię* *dobre* *okno.*
like$_{1st}$ good$_{neut,sing,acc}$ window$_{neut,sing,acc}$
'I like a good window.'

f. *Lubię* *dobrych* *chłopców.*
like$_{1st}$ good$_{masc,plur,acc}$ boy$_{masc,plur,acc}$
'I like good boys.'

g. *Lubię* *dobre* *psy.*
like$_{1st}$ good$_{masc,plur,acc}$ dog$_{masc,plur,acc}$
'I like good dogs.'

h. *Lubię* *dobre* *stoły.*
like$_{1st}$ good$_{masc,plur,acc}$ table$_{masc,plur,acc}$
'I like good tables.'

i. *Lubię* *dobre* *dziewczyny.*
like$_{1st}$ good$_{fem,plur,acc}$ girl$_{fem,plur,acc}$
'I like good girls.'

j. *Lubię* *dobre* *okna.*
like$_{1st}$ good$_{neut,plur,acc}$ window$_{neut,plur,acc}$
'I like good windows.'

In (28), the values of the attribute GENDER have been assigned on the basis of the ability of the nouns to be coindexed with personal pronouns: *chłopiec* ('boy'), *pies* ('dog') and *stół* ('table') can be coindexed with *on* ('he'), *dziewczyna* ('girl') with *ona* ('she'), *okno* with *ono* ('it'). The partition of Polish nouns into the three gender classes: *masc*, *fem* and *neut* is, however, not sufficient in order to account for the correct forms of the attributive adjectives in (28). Although *chłopiec* and *stół* are both masculine, the forms of the adjective in (28a) and (28c) are different. Also, the adjective forms in (28f) and (28g) differ, although the same form is used for the singular forms *chłopiec* and *pies* in (28a) and (28b).

On the basis of adjective-noun agreement, [SŚ85] propose five different, disjoint, gender classes:

1. *m1*: masculine-human nouns; e.g., *chłopiec* ('boy');

2. *m2*: masculine-animate nouns; e.g., *pies* ('dog');

3. *m3*: masculine-inanimate nouns; e.g., *stół* ('table');

4. *f*: feminine nouns; e.g., *dziewczyna* ('girl');

5. *n*: neuter nouns; e.g., *okno* ('window')

Additionally, the so-called *plurale tantum* nouns (e.g., *spodnie* ('trousers'), *wujostwo* ('uncle and aunt')) are considered in [SŚ85]. *Plurale tantum* nouns have only plural forms and are recognized as a separate gender class in order to account for their inability to occur with singular adjective forms. The class is divided further into masculine-human *plurale tantum* nouns and non-masculine-human *plurale tantum* nouns in order to explain the occurrence of distinct adjective forms in sentences like the following:

(29) *Mam* *dobrych* *wujostwa.*
have$_{1st}$ good$_{masc-hum}$ uncle and aunt$_{plurale-tantum}$
'I have a good uncle and aunt.'

(30) *Mam* *dobre* *spodnie.*
have$_{1st}$ good$_{non-masc-hum}$ trousers$_{plurale-tantum}$
'I have good trousers.'

Thus, two more gender classes are added to the classification:

6. *p1*: masculine-human *pluralia tantum*; e.g., *wujostwo* ('uncle and aunt');

7. *p-1*: non-masculine-human *pluralia tantum*; e.g., *nożyce* ('scissors').

Finally, the paradigm of the adjective *dobry* ('good') can be written as follows (the rows represent different *case* values, the columns - *gender* values):

(31)

	m1	*m2*	*m3*	*n*	*f*	*p1*	*p-1*
sing							
nom	dobry	dobry	dobry	dobre	dobra	—	—
gen	dobrego	dobrego	dobrego	dobrego	dobrej	—	—
dat	dobremu	dobremu	dobremu	dobremu	dobrej	—	—
acc	dobrego	dobrego	dobry	dobre	dobrą	—	—
ins	dobrym	dobrym	dobrym	dobrym	dobrą	—	—
loc	dobrym	dobrym	dobrym	dobrym	dobrą	—	—
plur							
nom	dobrzy	dobre	dobre	dobre	dobre	dobrzy	dobre
gen	dobrych	dobrych	dobrych	dobrych	dobrych	dobrych	dobrych
dat	dobrym	dobrym	dobrym	dobrym	dobrym	dobrym	dobrym
acc	dobrych	dobre	dobre	dobre	dobre	dobrych	dobre
ins	dobrymi	dobrymi	dobrymi	dobrymi	dobrymi	dobrymi	dobrymi
loc	dobrych	dobrych	dobrych	dobrych	dobrych	dobrych	dobrych

A closer look at the above paradigm leads to an immediate conclusion that it contains significant redundancy. Firstly, the forms for genders *p1* and *p-1* are identical to the plural forms for genders *m1* and *m2*, respectively. Moreover, in the plural only two distinct sets of adjective forms occur, namely these for genders *m1* and *m2*. Also, the paradigm for the singular is redundant for genders *m1*, *m2*, *m3* and *n* although the respective sets of adjective forms differ at least at one place.

Let us next overview the way in which nouns of the particular genders combine with verbs. Consider the following sentences (the past tense has been chosen as past verbal forms show (overt) gender marking; the meaning is irrelevant):

(32) a. *Chłopiec/pies/stół* *stał.*
boy$_{m1}$/dog$_{m2}$/table$_{m3}$ stood
'A boy/dog/table stood.'

b. *Okno* *stało.*
window$_n$ stood
'A window stood.'

c. *Dziewczyna* *stała.*
girl$_f$ stood
'A girl stood.'

d. *Chłopcy/wujostwo* *stali.*
boys$_{m2}$/uncle and aunt$_{p1}$ stood
'Boys/uncle and aunt stood.'

e. *Psy/okna/dziewczyny/nożyce* *stały.*
dogs$_{m2}$/windows$_n$/girls$_f$/scissors$_{p-1}$ stood
'Dogs/windows/girls/scissors stood.'

The examples in (32) prove that the verbal paradigms would also be highly redundant should the above gender classification be adopted. The overlaps have been noticed before and simple abbreviations in the form of disjunctions of *gender* values (e.g., $m = m1 \vee m2 \vee m3$) have been proposed. Such disjunctions, however, usually lead to multiple lexical entries in an implementation. Nevertheless, the main cause of redundancy is the assumption that values of the category *gender* are the same in the singular and plural.

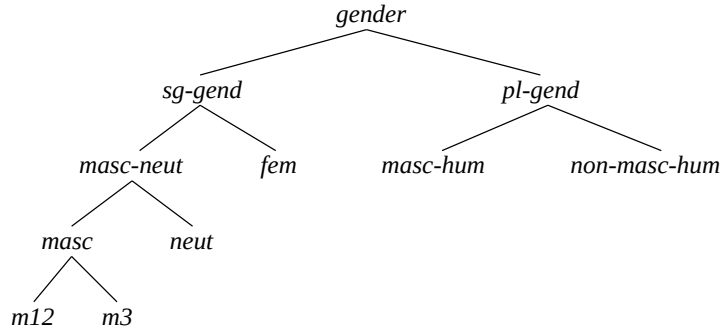
As our account is set in HPSG, expressing the above generalizations in a nondisjunctive way is relatively easy. We will propose a hierarchy of subtypes which captures the above facts and allows for a very compact form of the lexicon. In order to make the picture simple, we propose to give up having

the same gender values for the singular and plural. This also means that the category of number is not needed any more: since different genders correspond to the singular and plural forms, the number information is part of the gender information. Giving up the category of number simplifies paradigms of most syntactic categories but also leads to redundancies which we will deal with at the end of the section.

As we saw in (31), no more than four gender values are needed in the singular: *plurale tantum* nouns do not occur in the singular and the declensions in genders *m1* and *m2* are identical. In the plural only two gender values would be sufficient: one corresponding to the set of forms parallel to the one for *m1* and another one like for *m2*. Also, *m1*, *m2*, *m3* and *n* have multiple common forms. (32) confirms that it is desirable to place *m1*, *m2* and *m3* as subsorts of one gender and that in the plural two genders are sufficient.

The preliminary structure of the sort *gender*[14] based on the facts we have presented so far looks as follows:

(33)



Nodes in the tree in (33) represent subtypes of the type *gender*, daughters of a node represent (disjoint) subtypes of the type represented by the node.

With such a definition, the paradigm in (31) can be reduced to the following:[15]

[14] It should be noted that from now on the values of the sort *gender* are "more informative" than usually; we decided, however, to keep the traditional type name.

[15] In fact, only the adjective forms for the following (*case*, *gender*) pairs should be given in the lexicon:

(nom, masc)	(nom, neut)	(nom, fem)	(nom, masc-hum)	(nom, non-masc-hum)	
(gen, masc-neut)		(gen, fem)	(gen, pl-gend)		
(dat, masc-neut)		(dat, fem)	(dat, pl-gend)		
(acc, m12)	(acc, m3)	(acc, neut)	(acc, fem)	(acc, masc-hum)	(acc, non-masc-hum)
(ins, masc-neut)			(ins, fem)	(ins, pl-gend)	
(loc, masc-neut)			(loc, fem)	(loc, pl-gend)	

A similar reduction can be observed for the verbal paradigm.

(34)

	m12	*m3*	*neut*	*fem*	*masc-hum*	*non-masc-hum*
nom	dobry	dobry	dobre	dobra	dobrzy	dobre
gen	dobrego	dobrego	dobrego	dobrej	dobrych	dobrych
dat	dobremu	dobremu	dobremu	dobrej	dobrym	dobrym
acc	dobrego	dobry	dobre	dobrą	dobrych	dobre
ins	dobrym	dobrym	dobrym	dobrą	dobrymi	dobrymi
loc	dobrym	dobrym	dobrym	dobrą	dobrych	dobrych

Personal and possessive pronouns provide additional support for the above gender structure. The pronouns *ja* ('I') and *ty* ('you$_{sing}$') do not distinguish genders "below" *sg-gend*, *my* ('we') and *wy* ('you$_{plur}$') do not distinguish genders "below" *pl-gend* in the tree in (33). The possessive pronoun *jego* ('his, its') provides additional support for the sort *masc-neut*: it can be used for both neuter and masculine possessors. We will return to pronouns in 2.3.3.

With the first outline of the gender structure in Polish at hand we can consider further examples. More data on gender come from various instances of numeral-noun agreement. In order to avoid additional attributes for nouns and numerals in Polish, we will account for these data using different *gender* subsorts.

Consider:

(35) a. *dwaj* *chłopcy*
two boys$_{masc-hum}$

b. *dwoje* *wujostwa*
two$_{coll}$ uncle and aunt$_{masc-hum}$

In (35), the *masc-hum* nouns *chłopcy* and *wujostwo* combine with different numerals. The noun *chłopcy* requires an ordinary numeral, whereas the noun *wujostwo* requires a collective numeral. Thus, we propose two subsorts of the sort *masc-hum*: *mh1* and *mh2* on the basis of the above distinction. The sort *mh1* corresponds to plural *m1* forms and the sort *mh2* to *p1* in [SŚ85].

The sort *non-masc-hum* has to be split further in a similar fashion. Compare:

(36) a. *dwa* *psy/stoły/okna*
two dogs/tables/windows$_{non-masc-hum}$

b. *dwie* *kobiety*
two women$_{non-masc-hum}$

c. *dwoje* *skrzypiec/dzieci*
two$_{coll}$ violins/children$_{non-masc-hum}$

d. *dwie* *pary* *spodni*
two pairs trousers$_{non-masc-hum,gen}$
'two pairs of trousers'

We propose four subsorts of the sort *non-masc-hum*:

1. *nmh1*: for nouns requiring ordinary numerals, like in (36a); this class corresponds to the plural *m2*, *m3* and *n* nouns in [SŚ85];
2. *nmh2*: for nouns requiring ordinary numerals, like in (36b); this class corresponds to the plural *f* nouns in [SŚ85];
3. *nmh3*: for nouns requiring collective numerals, like in (36c); this class contains some *n* nouns, like *dzieci* ('children') and some non-masculine *plurale tantum* nouns;
4. *nmh4*: for nouns which cannot directly combine with numerals, like in (36d).

In HPSG, the fact that *plurale tantum* nouns have no singular forms is "automatically" accounted for by the lack of the corresponding "singular" lexical entries. In combination with adjectives and verbs, *plurale tantum* nouns behave like either *masc-hum* or *non-masc-hum* nouns and the special requirements towards numerals are also characteristic for some other nouns. Thus, we do not introduce a separate class for *plurale tantum* nouns. We are also aware that the above classification requires disjunctions of the form *mh1* $\vee$ *nmh3* in lexical entries of collective numerals, which could be a reason for recognizing a separate class of *plurale tantum* nouns. Such a class would, however, dim the picture of noun-verb agreement.[16]

In the current formulation the theory misses the fact that, e.g., *chłopiec* ('boy') and *chłopcy* ('boys') are forms of the same lexeme and are strongly semantically related, which is modelled by the category of number in traditional accounts. On our account, we propose a set of lexical rules which will capture these facts and account for derivational and semantic dependencies between genders.

In order to provide such a set of unambiguous rules, we will have to add new *gender* subsorts. Firstly, the sort *m12* has to be split into two: *m1* and *m2*, which correspond to the respective genders in [SŚ85]. This is necessary as only

[16] Also, the class of collective numerals is closed; thus the redundancy is fairly limited.

some nouns in *m12* have their traditional plural form in *mh1*; the remaining nouns collected in *m2* have their plural forms in *nmh2*. Also *neut* nouns have to be split into two classes, cf.:

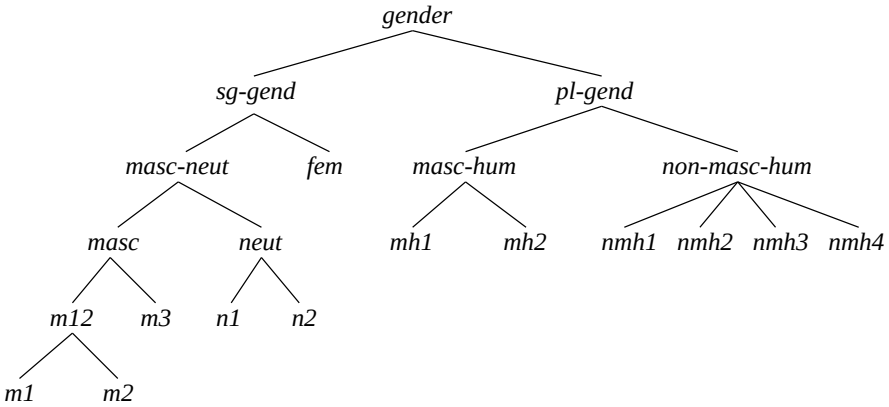
(37) a. *dwoje* *dzieci*
two$_{coll}$ children$_{nmh3}$

b. *dwa* *okna*
two windows$_{nmh1}$

In (37) both nouns *dzieci* and *okna* have their singular forms in *neut*. The respective plural forms, however, require different numerals. Thus, two subtypes of *neut* are proposed: *n1* for nouns like *okno* ('window') with the traditional plural in *mh2* and *n2* for nouns like *dziecko* ('child') with the traditional plural in *mh3*.

The final structure of the sort *gender* in Polish is presented in (38):

(38)



This structure is augmented with the following set of 6 lexical rules which cover all cases corresponding to the formation of plural forms in accounts were the category of number is present:[17]

[17] This set of rules can be easily modelled by one lexical rule with a PF function such that only for some pairs of the genders a value exists.

(39)

a. $\begin{bmatrix}\text{STEM}|\text{PHON}\ \boxed{1}\ \textit{mężczyzna}\\ \text{HEAD}|\text{AGR}|\text{GENDER}\ \boxed{2}\ \textit{m1}\end{bmatrix} \mapsto \begin{bmatrix}\text{PHON PF}(\boxed{1},\boxed{2})\ \textit{mężczyźni}\\ \text{HEAD}|\text{AGR}|\text{GENDER}\ \textit{mh1}\end{bmatrix}$

b. $\begin{bmatrix}\text{STEM}|\text{PHON}\ \boxed{1}\ \textit{pies}\\ \text{HEAD}|\text{AGR}|\text{GENDER}\ \boxed{2}\ \textit{m2}\end{bmatrix} \mapsto \begin{bmatrix}\text{PHON PF}(\boxed{1},\boxed{2})\ \textit{psy}\\ \text{HEAD}|\text{AGR}|\text{GENDER}\ \textit{nmh1}\end{bmatrix}$

c. $\begin{bmatrix}\text{STEM}|\text{PHON}\ \boxed{1}\ \textit{stół}\\ \text{HEAD}|\text{AGR}|\text{GENDER}\ \boxed{2}\ \textit{m3}\end{bmatrix} \mapsto \begin{bmatrix}\text{PHON PF}(\boxed{1},\boxed{2})\ \textit{stoły}\\ \text{HEAD}|\text{AGR}|\text{GENDER}\ \textit{nmh1}\end{bmatrix}$

d. $\begin{bmatrix}\text{STEM}|\text{PHON}\ \boxed{1}\ \textit{kobieta}\\ \text{HEAD}|\text{AGR}|\text{GENDER}\ \boxed{2}\ \textit{fem}\end{bmatrix} \mapsto \begin{bmatrix}\text{PHON PF}(\boxed{1},\boxed{2})\ \textit{kobiety}\\ \text{HEAD}|\text{AGR}|\text{GENDER}\ \textit{nmh2}\end{bmatrix}$

e. $\begin{bmatrix}\text{STEM}|\text{PHON}\ \boxed{1}\ \textit{okno}\\ \text{HEAD}|\text{AGR}|\text{GENDER}\ \boxed{2}\ \textit{n1}\end{bmatrix} \mapsto \begin{bmatrix}\text{PHON PF}(\boxed{1},\boxed{2})\ \textit{okna}\\ \text{HEAD}|\text{AGR}|\text{GENDER}\ \textit{nmh1}\end{bmatrix}$

f. $\begin{bmatrix}\text{STEM}|\text{PHON}\ \boxed{1}\ \textit{dziecko}\\ \text{HEAD}|\text{AGR}|\text{GENDER}\ \boxed{2}\ \textit{n2}\end{bmatrix} \mapsto \begin{bmatrix}\text{PHON PF}(\boxed{1},\boxed{2})\ \textit{dzieci}\\ \text{HEAD}|\text{AGR}|\text{GENDER}\ \textit{nmh3}\end{bmatrix}$

Before we close this section, let us note that all gender values we introduced are meaningful and useful. The subsorts *sg-gend* and *pl-gend* represent the level of granularity appropriate for personal pronouns in the 1st and 2nd person and present verbal forms. Sorts *masc*, *fem*, *neut*, *masc-hum*, *non-masc-hum* are distinguished by personal pronouns in the 3rd person and simple past and analytic future verbal forms. Sorts *m12*, *m3*, *fem*, *masc-hum* and *non-masc-hum* are the basis of the adjective paradigm. Multiple syncretisms occurring in the adjective paradigm for genders *masc* and *neut* can be captured by the *masc-neut* sort. Subsorts *mh1*, *mh2*, *nmh1*, *nmh2*, *nmh3* and *nmh4* are crucial to the correct account of the ability of nouns to combine with numerals. Sorts *m1*, *m2* and *n1*, *n2* are necessary to provide an unambiguous set of rules accounting for semantic and derivational dependencies modelled by the category of number which we do not use.

Finally, here are examples of nouns belonging to the *gender* classes proposed above:

1. *m1*: *mężczyzna* ('man');
2. *m2*: *pies* ('dog');
3. *m3*: *stół* ('table');
4. *fem*: *kobieta* ('woman'), *grupa* ('group'), *książka* ('book');
5. *n1*: *okno* ('window');
6. *n2*: *dziecko* ('child'), *prosię* ('piglet');

7. *mh1*: *mężczyźni* ('men');
8. *mh2*: *państwo* ('Mr and Mrs'), *wujostwo* ('uncle and aunt');
9. *nmh1*: *psy* ('dogs'), *stoły* ('tables');
10. *nmh2*: *kobiety* ('women')
11. *nmh3*: *drzwi* ('door, doors'), *children* ('dzieci');
12. *nmh4*: *spodnie* ('trousers').

2.3. NP-internal Agreement

In section (2.1.3.), all instances of agreement were classified into two classes corresponding to the morphosyntactic and indexical agreement patterns. This was captured by (24) repeated here as (40):

(40)
1. morphosyntactic agreement: AGR(selector) $\approx$ AGR(arg)
2. semantic (index) agreement: AGR(selector) $\approx$ INDEX(arg)

Our analysis will be based on (40): for a syntactic category, data illustrating its behaviour in constructions in which the category is the selector will be used in order to propose its AGR structure. Agreement features relevant for these data, i.e., the features with respect to which an obligatory agreement obtains, will be included into AGR of the selector on the basis of (40).

After a short discussion of the index structure in Polish, we start off with the NP-internal agreement. Strictly speaking, the NP-internal agreement in Polish falls into neither case mentioned in (40) as the selector/argument distinction is not applicable here. However, a morphosyntactic pattern can be applied to this instance of agreement.

2.3.1. Index Structure in Polish

The structure of indices in Polish plays a crucial role in the adopted theory of agreement. The attributes required in indices can be determined on the basis of pronoun-antecedent agreement.

As can be expected, case is not present in indices:

(41)

Tego	*psa*$_1$	*się*	*wszyscy*	*bali,*	*choć*	*nie*	*był*
this$_{gen}$	dog$_{gen}$	REFL	everybody	feared	although	not	was

on$_1$ *duży.*
he$_{nom}$ big

'Everybody was afraid of this dog although it was not big.'

In (41), the genitive form *psa* is coindexed with the nominative pronoun *on*, which proves that structure-sharing of the values of CASE does not obtain.

Compare:

(42) a. *Ten* *pies*$_1$ *jest* *groźny,* *choć* *nie* *jest* *on*$_1$
this dog$_{masc}$ is dangerous although not is he$_{masc}$
duży.
big
'This dog is dangerous although it is not big.'

b. * *Ten* *pies*$_1$ *jest* *groźny,* *choć* *nie* *jest* *ona*$_1$
this dog$_{masc}$ is dangerous although not is she$_{fem}$
duża.
big

c. * *Ten* *pies*$_1$ *jest* *groźny,* *choć* *nie* *jesteś* *ty*$_1$
this dog is dangerous although not are you
duży.
big

In grammatical (42a), the agreement with respect to both gender and person obtains in contradistinction to ungrammatical (42b) and (42c) were the agreement in gender and person, respectively, does not hold. The attributes of GENDER and PERSON are thus parts of index structure. We posit that only these two attributes are needed and propose the following index structure:

(43) $\begin{bmatrix} \text{PERSON } person \\ \text{GENDER } gender \\ index \end{bmatrix}$

In contrast to English, we do not assume the existence of any expletive indices (i.e., similar to indices *it* and *there* introduced for English); all indices are referential. (Non-overt) expletives and expletive indices could be useful in an analysis of sentences containing meteorological verbs like those in (18) which are not cases of pro-drop. In section 2.4.4. we propose a subjectless analysis of sentences of this type which does not require expletives and avoids unmotivated null elements in the grammar. The grammar will provide additional constraints on agreement feature marking on verbs in these sentences. Thus, we have not found any use for nonreferential indices in Polish so far.

2.3.2. Attributive Adjectives

Agreement features which should be included in AGR of nouns and adjectives can be determined on the basis of NP-internal and subject-predicative adjective agreement. Consider the following examples:

(44) a. *Ja jestem duża/duży.*
I am big$_{fem}$/big$_{masc}$
'I am big.'

b. *Ty jesteś duża/duży.*
you are big$_{fem}$/big$_{masc}$.
'You are big.'

In (44), the form of the adjective *duży* ('big') depends on gender of the subject but does not change with its person. In (45), the predicative adjective occurs with different case depending on the subject:

(45) a. *Dwóch chłopców było leniwych.*
Two boys$_{gen}$ was$_{3rd,neut}$ lazy$_{gen}$.
'Two boys were lazy.'

b. *Dwie dziewczynki były leniwe.*
Two girls$_{nom}$ were$_{3rd,non-masc-hum}$ lazy$_{nom}$.
'Two girls were lazy.'

Similarly, in (46) illustrating case concord in Polish, case of the adjective *duża* ('big') is consistent with case of the noun *piłka* ('ball'):

(46)
duża$_{nom}$	*piłka*$_{nom}$	'(a) big ball'
dużej$_{gen}$	*piłki*$_{gen}$	
dużej$_{dat}$	*piłce*$_{dat}$	
dużą$_{acc}$	*piłkę*$_{acc}$	
dużą$_{ins}$	*piłką*$_{ins}$	
dużej$_{loc}$	*piłce*$_{loc}$	

Hence, the structure of AGR for Polish predicative and attributive adjectives we propose is as follows:

(47)
$$\begin{bmatrix} \text{GENDER } gender \\ \text{CASE } case \\ adj\text{-}agr \end{bmatrix}$$

Let's have a look at the following NPs:

(48) a. *jego szacowna wysokość*
his respectable$_{fem}$ highness$_{fem}$
'his respectable highness'

b. *jej szacowna wysokość*
her respectable$_{fem}$ highness$_{fem}$
'her respectable highness'

(48a) and (48b) suggest that the attributive adjective-noun agreement pattern is of morphosyntactic nature. In both NPs, the adjective *szacowna* agrees with the syntactic gender of the noun *wysokość*, which is *fem*, and not with gender of the referent, which is *masc* in (48a) and *fem* in (48b). In general, the morphosyntactic pattern assuming case and gender concord between the adjective and the modified NP correctly predicts all instances of attributive adjective-noun agreement in Polish, and the following constraint on the HEAD attribute of attributive adjectives can be formulated:

(49)
$$\begin{bmatrix} \text{AGR} & \begin{bmatrix} \text{GENDER } \boxed{1} \\ \text{CASE } \boxed{2} \\ adj\text{-}agr \end{bmatrix} \\ \text{MOD N'} & \begin{bmatrix} \text{HEAD}\vert\text{AGR} & \begin{bmatrix} \text{GENDER } \boxed{1}\ gender \\ \text{CASE } \boxed{2}\ case \\ n\text{-}agr \end{bmatrix} \end{bmatrix} \\ \text{PRD} & - \\ attr\text{-}adj\text{-}pat & \end{bmatrix}$$

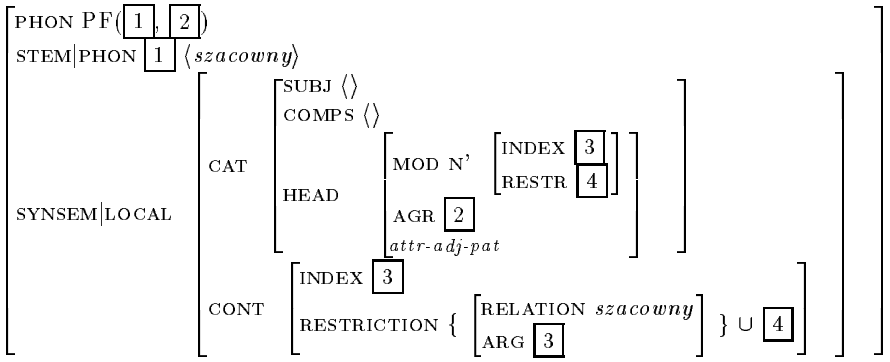
The pattern in (49) has been described as morphosyntactic because it involves structure sharing of the corresponding AGR features of the adjective and the NP it modifies as specified in MOD. We also assumed the following AGR structure for Polish nouns:

(50)
$$\begin{bmatrix} \text{CASE } case \\ \text{GENDER } gender \\ n\text{-}agr \end{bmatrix}$$

So far we have not seen any data which would require us to add any other features to AGR of nouns. We will argue further that this structure is sufficient, which contrasts with other accounts assuming that the category of person is a morphosyntactic characteristic of Polish nouns.

As an example we present the lexical entry for the attributive variation of *szacowny*:

(51)



2.3.3. Possessive Pronouns

Let us confront the above AGR of nouns with the facts taking possessive pronouns into account. (48a) and (48b) offer some data on Polish possessive pronouns. Consider additionally the following:

(52) a. *jego/jej* *książka*
his/her book$_{fem}$
'his/her book'

b. *jego/jej* *podręcznik*
his/her handbook$_{masc}$
'his/her handbook'

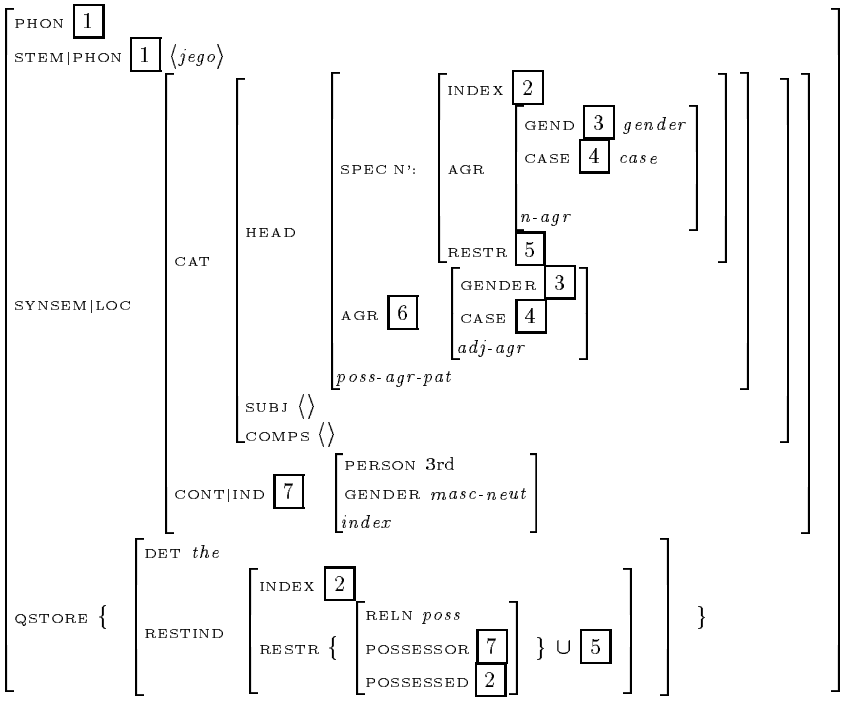
In (48a) and (48b), the form of the possessive is determined by gender of the possessor stored on the index: *jego* ('his/its') is used for masculine and neuter possessors and *jej* ('her(s)') is used for feminine possessors. As (52) shows, the form of *jego/jej* does not depend on gender of the possessee. Similarly, it does not depend on case:

(53) *o* *jego/jej* *podręczniku*
about his/her handbook$_{loc}$
'about his/her handbook.'

Also, the form of the remaining third person possessive pronoun *ich* ('their(s)') depends neither on gender of the possessee nor case. The following lexical entry for *jego* captures the above remarks; it is assumed that possessives in Polish are specifiers[18] and have the head feature SPEC filled:

[18] This assumption may raise some problems since in Polish specifiers/determiners are not obligatory, i.e., in most cases nouns occur on their own or only with modifiers. With the

(54)



In (54) the *poss-agr-pat* pattern represents a restriction on the HEAD attribute of possessives in Polish. It assumes that the AGR structure of possessive pronouns is identical to the AGR structure of adjective and expresses morphosyntactic agreement with respect to gender and case between the possessee and the possessor.

The last assumption has no influence on the form of the pronoun *jego*: the PHON value does not depend on 6. However, possessive pronouns for 1st and 2nd person possessors exhibit a different behaviour. Their form depends on gender of the possessee and does not depend on gender of the possessor:

current formulation of ID Schemata, specifiers are sanctioned by means of the SPEC principle "cooperating" with one of the ID Schemata. In English, the correct account is obtained by the requirement that a determiner phrase is obligatory, which is realized by a respective element on the SUBCAT list of every noun. This in turn sanctions the presence of the determiner in the sentence by means of schema 2. Such an account is rather unsuitable for Polish as it would require either two lexical entries for every noun — with and without a determiner — or null determiners which would have to be placed in front of most nouns. We will not pursuit these interesting issues here.

(55) a. *moja* *książka*
my$_{fem}$ book$_{fem}$
'my book'

b. *mój* *podręcznik*
my$_{masc}$ handbook$_{masc}$
'my handbook'

(56) a. *twoja* *książka*
your$_{fem}$ book$_{fem}$
'your book'

b. *twój* *podręcznik*
your$_{masc}$ handbook$_{masc}$
'your handbook.'

The form of 1st and 2nd person possessive pronouns changes also with case:

(57) *o* *moim* *podręczniku*
about my$_{loc}$ handbook$_{loc}$
'about my handbook.'

Below we give a lexical entry for *mój* ('my') which presents the relevant agreement dependencies:[19]

In (58), the same *poss-agr-pat* is used in order to account for the agreement between the possessive and the noun. Additionally, the index of the possessor is anchored to the speaker there.[20]

Summerizing, the data on possessive pronouns and NP-internal agreement provide further support for the AGR structure for nouns we proposed in 2.3.2. Also, the morphosyntactic pattern of agreement proved useful to describe the choice of possessive pronouns forms in Polish.

---

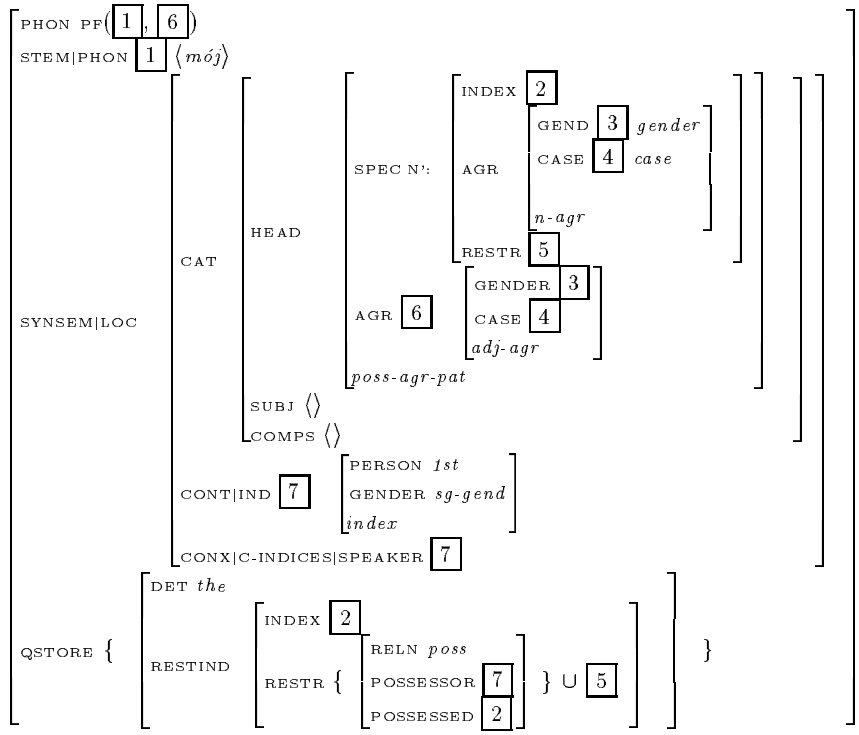
[19] In (58) and (54) we used the structure of QSTORE proposed for English, which might not be appropriate for Polish. As the theory of quantifiers in Polish has not yet been created and it is not the topic of the paper, we find our usage of QSTORE justified for illustrative purposes.

[20] The *poss-agr-pat* will also appear in the lexical entry for the possessive pronoun *swój* ('ones own'). *Swój* appears in contexts like in the following example:

(i) *Jan* *wziął* *swój* *zeszyt.*
Jan took own copy-book
'Jan took his copy-book.'

The form of the pronoun *swój* depends on gender of the possessee and case. Hence, the lexical entry for this pronoun resembles the lexical entry for *mój*.

(58)



2.4. Subject-Verb Agreement

2.4.1. AGR Structure for Polish Verbs

The following simple examples provide support for including GENDER and PERSON into the agreement features characteristic of verbs:

(59) a. *Ja poszedłem.*
I went$_{1st,past,masc}$
'I went.'

b. *Ona poszła.*
she went$_{3rd,past,fem}$
'She went.'

c. *On poszedł.*
he went$_{3rd,past,masc}$

'He went.'

d. * *On* *poszła.*
he went$_{3rd,past,fem}$

e. * *Ja* *poszła.*
I went$_{3rd,past,fem}$

In (59a) and (59b) the verb forms differ according to the person marking of the subject; in (59b) and (59c) the verbal forms bear a gender marking matching the gender marking of the subject. As (59d) and (59e) show, the matching in gender and person is obligatory.

On the face of the above, the structure of AGR for Polish verbs we propose is as follows:

(60) $\begin{bmatrix} \text{PERSON } person \\ \text{GENDER } gender \\ v\text{-}agr \end{bmatrix}$

The following set of examples suggests that additional information might be required in AGR of verbs:

(61) a. *Adam* *zaczyna* *czytać.*
Adam start$_{3rd,sg-gend,pres}$ read$_{inf,nonperfective}$
'Adam is starting to read.'

b. * *Adam* *zaczyna* *przeczytać.*
Adam start$_{3rd,sg-gend,pres}$ read$_{inf,perfective}$

The examples present verbs which can select infinitive forms of *nonperfective* verbs. So far selectional restrictions of a syntactic category have been modelled by the SUBCAT attribute which contained requirement on the AGR of potential complements. In order to preserve this picture, adding the MOOD agreement feature with two possible values: *perfective* and *nonperfective* could be helpful. As we will not discuss mood and tense in the remainder, we will further assume the structure in (60).

2.4.2. Morphosyntactic vs. Index Agreement

As noted above, Kathol's theory of agreement distinguishes two patterns of agreement: morphosyntactic and indexical. In this section we will discuss the distribution of these patterns in subject-verb agreement in Polish.

Consider the following examples:

(62) a. *Dzieci* *biegły.*
children$_{non-masc-hum,3rd}$ ran$_{non-masc-hum,3rd}$
'Children ran.'

b. * *Dzieci* *biegli.*
children$_{non-masc-hum}$ ran$_{masc-hum}$

c. * *Dzieci* *biegłyśmy.*
children$_{3rd}$ ran$_{2nd}$

In (62a), there is an obligatory agreement in gender and person between the subject and the verb. This is the regular pattern of subject-verb agreement in Polish (cf. [BKS71]). Within our theory, it can be modelled by the morphosyntactic pattern in which the attributes GENDER and PERSON of the subject's AGR and the verb's AGR are structure-shared. This would mean that the feature PERSON is a part of AGR for Polish nouns and should be added to the structure in (43). Such an approach would correspond to traditional accounts of the phenomenon.

However, the morphosyntactic pattern is not suitable for all instances of subject-verb agreement in Polish. Compare the following sentences:

(63) a. *Jego* *szacowna* *wysokość* *był*
his respectable$_{fem}$ highness$_{fem,3rd}$ was$_{masc,3rd}$
zmęczony.
tired$_{masc}$.
'His respectable highness was tired.'

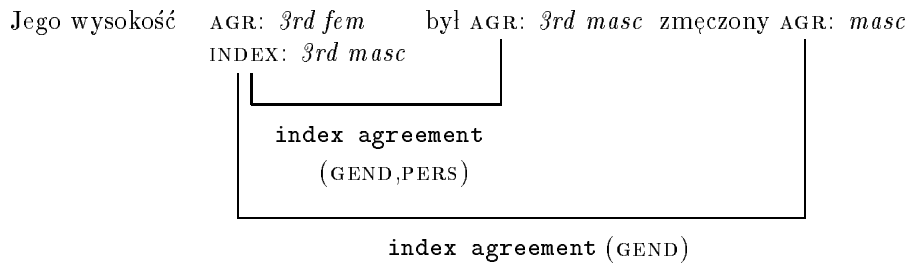
b. *Jej* *szacowna* *wysokość* *była* *zmęczona.*
Her respectable$_{fem}$ highness$_{fem,3rd}$ was$_{fem,3rd}$ tired$_{fem}$.
'Her respectable highness was tired.'

In (63a) the noun *wysokość* is morphosyntactically feminine but the copula *był* that follows it bears the masculine gender marking *masc*. Gender of the copula reflects the fact that the sentence refers to a male person. Similarly, (63b) can be used to assert a fact about a female person, which results in a different form of the copula; the covariation of GENDER of *wysokość* and *była* in (63b) is coincidental. Thus, the morphosyntactic subject-verb agreement pattern is not applicable to sentences like those in (63).

A correct analysis is obtained when the indexical pattern of subject-verb agreement is assumed for sentences (63a) and (63b). According to the theory, the index of the noun *wysokość* reflects its anchoring conditions. Thus, the

value of the GENDER attribute of the noun *wysokość* in (63a) and (63b) is *masc* and *fem*, respectively. By means of index agreement with respect to gender the value of GENDER of the nouns is structure-shared with the value of GENDER of the copula's AGR, which results in the masculine verb form in (63a) and the feminine verb form in (63b). The analysis of (63a) is presented in (64):

(64)



(63)–(64) illustrate the semantic pattern of subject-verb agreement which is very rare in contemporary Polish (see [BKS71]). In traditional accounts of agreement in Polish, this type of agreement is assumed to play a marginal role. Nevertheless, cases like the ones in (63a) and (63b) occur, especially with nouns expressing respect.

If we followed the assumption that subject-verb agreement is usually morphosyntactic and only sometimes semantic, we would obtain a disjunctive analysis, i.e., an analysis which allows for two subject-verb agreement patterns. In our opinion the undesirable disjunctive character can be easily shifted to the lexicon. This takes the burden of disjunction from parsing to lexicon construction resulting in more efficient applications.

2.4.3. AGR and INDEX in Polish Nouns Revisited

Let us first consider gender assignment in Polish nouns. Polish is a grammatical gender language. As mentioned above, this amounts to a partition of all Polish nouns into classes corresponding to the different *gender* values within the system of the Polish language. Only for some nouns (mostly those having creatures with biological sex as denotation) the *gender* value can be derived from extralinguistic properties of the referent. If we consider "typical" nouns like *stół* ('table'), we come to the conclusion that the settings of the GENDER attribute on the INDEX and the AGR of the noun are the same. This can be confirmed for the "typical" noun *stół* by the following example:

(65) *Widzę* *duży* *stół*$_1$. *On*$_1$ *jest* *czarny.*
see$_{first}$ big$_{acc,m3}$ table$_{acc}$ he$_{nom}$ is black
'I see a big table. It is big.'

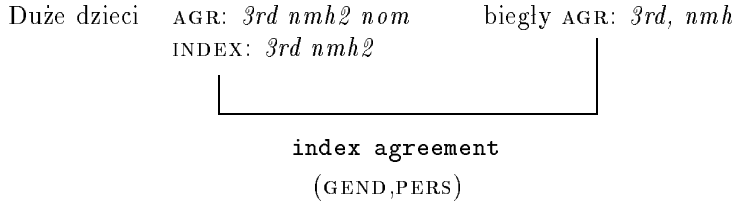
On the basis of the accusative NP *duży stół* we can assert that the GENDER value in AGR of the noun *stół* is *m3*. Observe that the assumption that the GENDER value on the INDEX of *stół* is *m3* leads to the correct account of the pronoun form: on the basis of our *gender* structure, *m3* is a subtype of *masc* for which the pronoun *on* is the right choice.

Let us now consider the PERSON attribute. As we showed in the previous section, this attribute is present on indices. It is not, however, present on AGR of adjectives. Thus it does not play any role in adjective-noun agreement. It can be observed that it does not play any role in NP-internal agreement in Polish at all. So far, the presence of the PERSON attribute on the AGR of nouns was required by our assumption that the subject-verb agreement in Polish is of morphosyntactic nature. We will drop this assumption, however, and posit that the subject-verb agreement in Polish is indexical. Hence, the PERSON attribute can be removed from AGR of nouns without causing the grammar to overgenerate. This results in the following, simpler structure which we have already proposed:

(66) $\begin{bmatrix} \text{CASE } case \\ \text{GENDER } gender \\ np\text{-}agr \end{bmatrix}$

Our assumption about the subject-verb agreement pattern has to be squared with data which were accounted for by the morphosyntactic pattern so far. Let us reanalyze example (62a) with the new assumption. The new analysis assumes also that *dzieci* ('children') is a typical noun. Thus the values of GENDER on the INDEX and AGR of the noun are the same and the indexical pattern explains the required gender agreement as well as the morphosyntactic pattern does. The noun *dzieci* can be coindexed with the pronoun *one*, which proves that its *person* attribute is set to *3rd*. Hence, the indexical pattern easily explains the requirement that the verb in (62a) bears the 3rd person marking, which concludes our proof that the indexical pattern is equally good as the morphosyntactic one at explaining the properties of subject-verb agreement in Polish. Since this pattern is the only possible one in cases like in (63), we will henceforth assume that all instances of subject-verb agreement in Polish are indexical. Finally, the new analysis of (62a) is depicted in (67).

(67)



Similarly to GENDER settings, the PERSON settings on the INDEX of a (typical) Polish noun seem to be highly restricted. This means that although indices provide an interface to the extralingual properties of the referent and are strongly connected with the way an object is individuated in the discourse, the values of their attributes are specified morphosyntactically. These values are set in the lexicon and it is up to the speaker to choose a noun with the correct INDEX settings in order to preserve grammaticality. Such a conclusion can be drawn from the fact that most nouns can be coindexed only with personal pronouns bearing the person value specific for the noun. Again, for nouns expressing politeness, the value of PERSON is underspecified in the lexicon as it can vary depending on the way the denoted object is individuated in the discourse, which in turn can be influenced, e.g, by the level of familiarity, cf.:

(68) a. *Państwo* *pójdą* *tędy.*
Madam and Sir go$_{3rd}$ this way
'Please, go this way.'

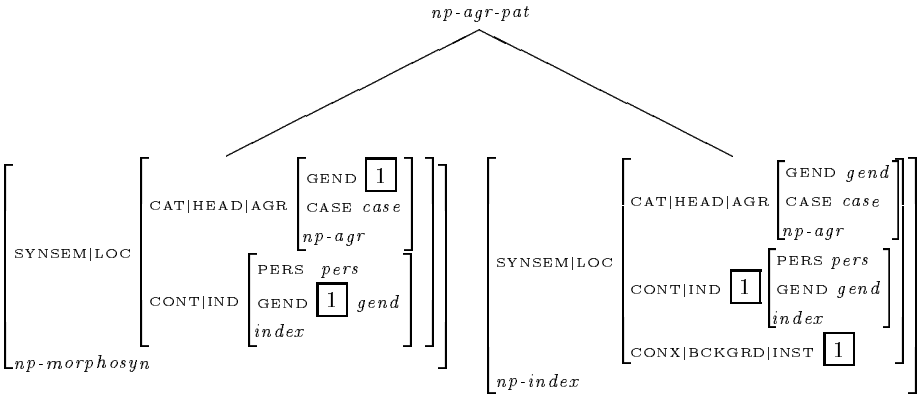
b. *Państwo* *pójdziecie* *tędy.*
Madam and Sir go$_{2nd}$ this way
'Please, go this way.'

where (68a) is slightly more familiar than (68b) (similarly for *Pan* ('Sir'), *Pani* ('Madam'), etc.).

In order to formalize our proposal, we posit two patterns to which Polish nouns adhere (which are restrictions on the SYNSEM values of the respective lexical signs): *np-morphosyn* that obtains for most nouns and features structure sharing of the relevant AGR and INDEX attributes, and *np-index* that obtains for personal nouns like *wysokość*, *mość*, etc. that make reference to anchoring

conditions in order to acquire the index settings.[21] For all nouns following the *np-index* pattern the AGR features are set in the lexicon arbitrarily, i.e., independently for every noun, often on the basis of the noun's morphology.

(69)



The partition of Polish nouns in (69) presents the current state in the language. From the diachronic point of view this partition is very recent: the choice of pattern has not always been clear-cut and the morphosyntactic pattern used to be applicable to nouns like *wysokość* whereas in the contemporary Polish only the indexical pattern is correct.

Before we close the section, we remark that the above analysis is clearly disjunctive but its disjunctive character is hidden in the lexicon. Thus, there is no additional overhead in parsing which would be unavoidable should two patterns of subject-verb agreement in Polish be assumed. The unique analysis was made possible by assuming the indexical pattern for subject-verb agreement. This assumption allowed us also to reduce the AGR structure for Polish nouns.

2.4.4. Subject-Verb Agreement Patterns

Personal sentences

So far we have considered only the core cases of subject-verb agreement in Polish. All examples contained sentences which we will call *personal*. All sentences with a personal form of the main or auxiliary verb belong to this

[21] It should be noted that personal pronouns which make a reference to anchoring conditions seem to adhere to the *np-morphsyn* pattern.

class. In personal sentences, the subject is usually nominative (we leave the discussion of sentences containing subjects with numerals to 3.3.).

The following examples (based on [BKS71]) show the covariation exhibited in personal sentences.

(70) a. *Chłopiec biegł.*
boy$_{masc}$ run$_{masc,past}$
'A boy ran.'

b. *Dziewczynka biegła.*
girl$_{fem}$ run$_{fem,past}$
'A girl ran.'

(71) a. *Ty piszesz.*
you$_{2nd}$ write$_{2nd}$.
'You are writing.'

b. *Ja piszę.*
I$_{1st}$ write$_{1st}$.
'I am writing.'

(70) shows agreement in gender and (71) agreement in person. In the previous section we decided that the subject-verb agreement is of indexical nature. Thus, the pattern of subject-verb agreement for personal forms in Polish can be formulated as the following restriction on the CATEGORY attribute of lexical entries of personal verbs:

(72)
$$\left[\begin{array}{l} \text{HEAD|AGR} \left[\begin{array}{l}\text{PERSON } \boxed{1} \\ \text{GENDER } \boxed{2} \\ v\text{-}agr\end{array}\right] \\ \text{SUBJ } \langle\ \text{NP}[\text{HEAD|AGR|CASE } nom] \left[\begin{array}{l}\text{PERSON } \boxed{1}\ person \\ \text{GENDER } \boxed{2}\ gender \\ index\end{array}\right] \rangle \\ personal \end{array}\right]$$

Pro-drop in Polish

Before we discuss other cases of subject-verb agreement, we need to clarify a potential problem. In the remainder, we will often refer to empty and null subjects. Null subjects are nonempty subjects filled with null elements, i.e., elements with no overt material, empty subjects correspond to the requirement that the SUBJ list is the empty list. These are not to be confused with

unexpressed or understood subjects which occur very often in Polish. This phenomenon, often called pro-drop in derivational approaches, is ubiquitous in Polish: it is natural to omit the subject in sentences in which it can be easily recovered from the context:

(73) a. *Powiedział, że jest chory.*
said$_{masc,3rd}$ that is$_{sg-gend,3rd}$ ill
'He said he was ill.'

b. *On powiedział, że on jest chory.*
he said$_{masc,3rd}$ that he is$_{sg-gend,3rd}$ ill
'He said he was ill.'

In (73a), the main clause and the subordinate clause have unexpressed subjects. Following traditional accounts, we assume the same analysis for (73a) and (73b) with a null element taking the place of the pronoun *on* ('he').[22] A plethora of examples can be given which show that a null element can occur as the subject in sentences with various personal verb forms. We will assume that only one null element fully unspecified for gender and person is needed in order to account for pro-drop in Polish. Additionally, we will assume that the feature AGR|CASE is set to *nom*. The structure of the null element is presented in (74).

(74)
$$\begin{bmatrix} \text{PHON } \langle\rangle \\ \\ \text{SYNSEM|LOCAL} & \begin{bmatrix} \text{CATEGORY|HEAD|AGR} & \begin{bmatrix} \text{GENDER } gender \\ \text{CASE } nom \\ n\text{-}agr \end{bmatrix} \\ np\text{-}morphsyn \end{bmatrix} \end{bmatrix}$$

The reader can easily check that such a definition will not lead to an ambiguity: the use of nominative is restricted to the subject position (cf. 3.); hence, the above empty element will not occur in nonsubject positions and will not cause overgeneration.

Impersonals

Let us start the analysis with sentences containing verb forms like in the examples:

(75) a. *Pito dużo piwa.*
There was drunk a lot of beer

[22] On one of the meanings, of course. The other meaning can be obtained for (73b) when the two occurrences of *on* are not coindexed.

'A lot of beer was drunk.'

b. *Mówiono, że Jan jest chory.*
It was said that Jan is ill
'Jan was said to be ill.'

The verb forms ending in *-no* or *-to* (which we will call the *impersonal* following Świdziński) can function as predicators in *impersonal* sentences. We posit that sentences of this type do not have subjects. This claim can be justified by the fact that impersonal sentences never occur with overt subjects. Thus proposing empty subjects is plausible and it prevents the introduction of an unmotivated null element in the grammar.

Because of the idiosyncratic verb form, we propose to analyze these sentences as a separate class. The appropriate agreement pattern is obtained by means of the following lexical rule which derives the impersonal from the infinitive:

(76)
$$\begin{bmatrix} \textsc{phon}\ \boxed{1} \\ \textsc{synsem|local|category}\ \begin{bmatrix} \textsc{head|vform}\ \mathit{inf} \\ \textsc{subj}\ \langle\ \mathrm{NP}[\mathit{nom}]\ \rangle \end{bmatrix} \\ \mathit{word} \end{bmatrix}$$
$$\mapsto$$
$$\begin{bmatrix} \textsc{phon}\ \mathrm{PF}(\boxed{1}, \boxed{2}) \\ \textsc{synsem|local|category}\ \begin{bmatrix} \textsc{head|vform}\ \boxed{2}\ \mathit{impersonal} \\ \textsc{subj}\ \langle\rangle \end{bmatrix} \end{bmatrix}$$

Note that we assume that the impersonal can be derived only for verbs which require a nominative NP as subject.

Constructions like:

(77) a. *Trzeba iść.*
should go
'One should go.'

b. *Można iść.*
allow go
'It is allowed to go.'

seem to constitute a similar lot with respect to the subject requirements. Lexemes **trzeba** ('one should'), **można** ('one may'), **warto** ('it is worth'), **wolno** ('it is allowed') are often assumed to be verbs. There are at least two arguments for such a treatment: (a) they can function as predicators in sentences like in (77), (b) they conjugate for mood and tense like verbs do.

For verbs like *można* it is plausible to assume that their subjects are empty. These verbs never occur with overt subjects and positing a null element is, in our opinion, unmotivated. Also, we will assume that these verbs have only finite forms and bear the same agreement features as 'normal' verbs and that they are fully underspecified for these features.

Meteorological verbs

Next, let's consider meteorological verbs. They appear in sentences similar to the following:

(78) a. *Mży.*
drizzle$_{3rd,neut,pres}$
'It drizzles.'

b. *Grzmiało.*
thunder$_{3rd,neut,past}$
'It thundered.'

In our account we propose that meteorological verbs are marked for person and gender in their finite forms. It can be also plausible to assume that no marking is involved, or that there is no marking for person (as in [SŚ85]), which could explain the fact that the subjects of meteorological verbs are empty. We will not follow the idea here and we will assume that finite forms of meteorological verbs can bear the same markings as "normal" verbs, i.e., verbs which can occur with nominal subjects. The main argument for such a stand is the fact that the special agreement feature marking of meteorological verbs can be explained in a general way useful for analyzing other related phenomena and it avoids assigning meteorological verbs a special character. On the basis of the morphology of the finite forms of meteorological verbs we assume that they are marked as 3rd person neuter (simple past forms of meteorological verbs end with *-o* like simple past forms of regular verbs).

Meteorological verbs do not occur with overt subjects. In other languages, for instance in English and German, (overt) expletive pronouns occur with such verbs. We do not find, however, any plausible motivation for introducing expletive pronouns for such sentences. Instead, we propose to analyze sentences with meteorological verbs as subjectless, avoiding null elements again.

The next step of our analysis is to provide a way of accounting for the special agreement marking of meteorological verbs. It can be easily captured by means of the following implication which is supposed to be viewed as a general principle holding for verbs in Polish:

(79)
$$\left[\begin{array}{l}\text{SYNSEM|LOCAL|CATEGORY}\ \left[\begin{array}{l}\text{HEAD|VFORM}\ \textit{fin}\\ \text{SUBJ}\ \langle\rangle\end{array}\right]\\ \textit{word}\end{array}\right] \Longrightarrow$$
$$\left[\text{SYNSEM|LOCAL|CATEGORY|HEAD|AGR}\ \left[\begin{array}{l}\text{PERSON}\ \textit{3rd}\\ \text{GENDER}\ \textit{neut}\end{array}\right]\right]$$

The above rule forces the correct setting of agreement features in finite forms of meteorological verbs: as these verbs occur with no subjects, the only finite form they can have is 3rd person neuter.

In our proposal, the fact that the subject is empty induces the special agreement. Clearly, introducing such a restriction on all verbs only in order to explain the form of meteorological verbs would be rather far-fetched. We will shortly present more sentence types in which rule (79) operates.

Before we do this, let us go back to example (77). As we assumed above that verbs like *można* have finite forms with no subjects, the implication (79) applies also to these verbs. This, however, does not lead to inconsistencies since we assumed also that these verbs are fully underspecified for agreement features.

Nonnominative subjects

Let's consider the following examples:

(80) a. *Pawłowi* *ubywało* **sił.**
Paweł$_{dat}$ decreased$_{3rd,neut}$ energy$_{gen}$
'Paweł was losing his energy.'

b. **Dnia** *przybywało.*
day$_{gen}$ increased$_{3rd,neut}$
'Days were getting longer.'

c. **Kasi** *było* *zimno.*
Kasia$_{dat}$ was$_{3rd,neut}$ cold
'Kasia was cold.'

On traditional accounts, the words in bold are analyzed as nonnominative subjects. Again, we do not find a sufficient motivation for such an analysis.[23] In order to obtain a clear picture of sentence types in Polish, we propose that these sentences are also subjectless. Then we can assume that the verbs which occur in (80) bear agreement features characteristic for regular verbs and, as

[23] Polish is characterized by a relatively free word order and the position of a phrase does not provide evidence for it being a subject.

in the case of meteorological verbs, we posit that the verb forms are marked 3rd person neuter. This restriction on the agreement feature marking is readily accounted for by the implication in (79).

On the face of the above, the traditional nonnominative subjects in (80) are to be analyzed as complements. It can be done in a very natural way assuming the following SUBCAT lists for the lexical entries of the respective verbs (for all of them, SUBJ ⟨⟩):

(81) a. ubywać: ⟨ NP[*dat*], NP[*gen*] ⟩
b. przybywać: ⟨ NP[*gen*] ⟩
c. być: ⟨ NP[*dat*], AdvP[*feeling*] ⟩

We also do not agree with some analyses (e.g., [SŚ85]) which assume that the form *nie ma* ('is not') is idiosyncratic and constitutes a separate lexeme. We prefer to treat it as the negative present form of the lexeme *być* ('be') which requires an empty subject, a genitive NP and an adverbial phrase as its complements:

(82) a. *Jasia* *jest/było* *pełno* *wszędzie.*
John$_{gen}$ is/was$_{3rd,neut}$ full everywhere.
'John is/was all over the place.'
b. *Jasia* *nie* *ma/było* *w* *domu.*
John$_{gen}$ not is/was$_{3rd,neut}$ in home.
'John is/was not at home.'

in contrast to the regular lexeme which we assume to have the regular negative form:

(83) a. *Jaś* *był* *w* *domu.*
John$_{nom}$ was$_{masc,3rd}$ in home
'John is at home.'
b. *Jaś* *nie* *był* *w* *domu* *(, tylko* *w* *szkole).*
John$_{nom}$ not was$_{masc,3rd}$ in home (but in school)
'John is not at home (but at school).'

The regular lexeme *być* conforms with the nominative subject-personal verb agreement pattern we discussed in 2.4.4.. This pattern can be built into the analysis by means of another implication which is complementary to (79):

(84)
$$\left[\begin{array}{l}\textsc{synsem|local|category}\ \left[\begin{array}{l}\textsc{head|vform}\ \mathit{fin}\\ \textsc{subj}\ \langle\ \mathrm{NP}[\mathit{nom}]\left[\textsc{person}\ \boxed{1}\right]\rangle\\ \qquad\qquad\qquad\quad\left[\textsc{gender}\ \boxed{2}\right]\end{array}\right]\\ \mathit{word}\end{array}\right]$$
$$\Longrightarrow$$
$$\left[\textsc{synsem|local|category|head|agr}\ \left[\begin{array}{l}\textsc{person}\ \boxed{1}\\ \textsc{gender}\ \boxed{2}\end{array}\right]\right]$$

Sentential subjects

The reader will surely notice that (79) and (84) are too restrictive in order to make correct predictions about all Polish simple sentences. A very numerous group of sentences with numeral phrase subjects will be incorporated into the analysis in 4.. Here we will add sentential subjects.

Let us first have a look at some motivation for introducing sentential subjects. Compare the following sentences:

(85) a. *Janka* *dziwiła* *ta* *podwyżka.*
John$_{acc}$ surprised$_{fem,3rd}$ this raise$_{fem,nom}$.
'This raise surprised John.'

b. *Janka* *dziwiło,* *że* *podniesiono* *ceny.*
John$_{acc}$ surprised$_{neut,3rd}$ that raised prices.
'That prices have been raised surprised John.'

In (85a), the nominative NP *ta podwyżka* ('this raise') is the subject. In (85b), the structure is parallel to (85a) but one of the arguments of *dziwi* is expressed by a *że*-sentence. There are at least two ways of analyzing the examples in (85). The first one would be to assume that there are two different lexical entries for *dziwi*: one which requires a nominative subject as in (85a) and another one with the empty subject and an NP[*acc*] and S[*że*] as complements. Another possibility, which we assume, is to posit that there is only one lexical entry which allows for the subject to be either NP[*nom*] or S[*że*].[24] The correct 3rd person neuter agreement of *dziwi* that we marked in (85b) is accounted for by the following revised version of (79):

(86)
$$\left[\begin{array}{l}\textsc{synsem|local|category}\ \left[\begin{array}{l}\textsc{head|vform}\ \mathit{fin}\\ \textsc{subj}\ \neg\langle\ \mathrm{NP}[\mathit{nom}]\rangle\end{array}\right]\\ \mathit{word}\end{array}\right]\quad\Longrightarrow$$

[24] It can be viewed as two lexical entries with the same argument structure; it is not the case for *dziwi* with an empty subject.

$$\left[\text{SYNSEM|LOCAL|CATEGORY|HEAD|AGR}\ \begin{bmatrix}\text{PERSON}\ \textit{3rd}\\ \text{GENDER}\ \textit{neut}\end{bmatrix}\right]$$

Note that (86) covers all cases (79) did allowing additionally for sentential subject. The negation in (86) is very general but it should be borne in mind that the subject is specified in the lexical entry of a verb and the only function of (86) is to determine the verb agreement features in some special cases.[25]

Verb Agreement Principle

Summerizing the above discussion of subject-verb agreement in Polish we formulate The Verb Agreement Principle operating in Polish which captures the above remarks.

---

[25] For completeness, sentences like in (i) should also be taken into account:

	W	*Polsce*	*często*	*jadło*	*się*	*ziemniaki.*
(i)	In	Poland	often	ate$_{3rd,neut}$	REFL	potatoes
	'Potatoes were often eaten in Poland.'					

In our opinion, the 3rd person neuter agreement of the verb form *jadło* is unquestionable. It is, however, slightly unclear what is the subject in (25). [SŚ85] account for sentences like in (25) by assuming that the reflexive *się* ('oneself') is the nominative subject specified as 3rd person neuter. We agree with this proposal only partially: the reflexive *się* can also occur in other contexts where assuming the nominative form of the above occurrence of *się* could lead to multiple lexical entries. Consider:

		Piotr	*myje*	*rękę.*
(ii)	a.	Peter	washes	hand$_{acc}$
		'Peter is washing his hand.'		
		Piotr	*myje*	*się.*
	b.	Peter	washes	oneself
		'Peter is washing (himself).'		
		Ja	*myję*	*się.*
	c.	I	wash	oneself
		'I am washing (myself).'		
		Ona	*myje*	*się.*
	d.	She	washes	oneself
		'She is washing (herself).'		

Although we have not explored the intricacies of Polish reflexives, we would rather assume that there is only one lexical entry for *się* specified as NP[acc]. In this way we leave open the possibility of using the same lexical entry in an analysis of (25b). Additionally we assume that *się* is fully unspecified for person and gender which can be useful in an analysis of (25c) and (25d). As for the correct agreement of the verb in (25), it is readily provided by the implication (86) and it follows from our assumption that *się* is accusative. Currently we are still researching this part of the theory.

(87) **Verb Agreement Principle:**

- $\left[\begin{array}{l}\textsc{synsem|local|category}\ \left[\begin{array}{l}\textsc{head|vform}\ \textit{fin}\\ \textsc{subj}\ \neg\langle\ \mathrm{NP}[\textit{nom}]\rangle\end{array}\right]\\ \textit{word}\end{array}\right]$

  $\Longrightarrow$

  $\left[\textsc{synsem|local|category|head|agr}\ \left[\begin{array}{l}\textsc{person}\ \textit{3rd}\\ \textsc{gender}\ \textit{neut}\end{array}\right]\right]$

- $\left[\begin{array}{l}\textsc{synsem|local|category}\ \left[\begin{array}{l}\textsc{head|vform}\ \textit{fin}\\ \textsc{subj}\ \langle\ \mathrm{NP}[\textit{nom}]\left[\textsc{person}\ \boxed{1}\ \textsc{gender}\ \boxed{2}\right]\rangle\end{array}\right]\\ \textit{word}\end{array}\right]$

  $\Longrightarrow$

  $\left[\textsc{synsem|local|category|head|agr}\ \left[\begin{array}{l}\textsc{person}\ \boxed{1}\\ \textsc{gender}\ \boxed{2}\end{array}\right]\right]$

2.5. Predicative Adjective and Noun Phrases

In comparison to languages like English and German, Polish predicative adjectives show additional case variation. Two distinct phenomena should be distinguished: pragmatic and stylistic variation between nominative and instrumental forms, and syntactic variation with the choice of form between nominative/instrumental and genitive.

Consider the following examples:

(88) a. *On jest nauczycielem.*
he is teacher$_{ins}$
'He is a teacher.'

b. *On jest nauczyciel.*
he is teacher$_{nom}$
'He is a teacher.'

(89) a. *On jest duży.*
he is big$_{nom}$
'He is big.'

b. ?*On jest dużym.*
he is big$_{ins}$
'He is big.'

Sentences (88a) and (89a) exemplify two very common patterns in Polish. In (88a), the noun *nauczyciel* ('teacher') is instrumental and in (89a) the adjective *duży* is nominative. Sentences (88b) and (89b) are clearly marked and their usage is restricted by multiple semantic, pragmatic and stylistic requirements which we will not discuss here. In general, in sentences like in (88) and (89) the more common form of predicative nouns is instrumental and the more common form of predicative adjectives is nominative.

The second kind of covariation which can be observed in Polish predicative adjectives and nouns is exemplified below:

(90) a. *Dwie kobiety były zmęczone.*
two women were tired$_{non-masc-hum,nom}$
'Two women were tired.'

b. *Dwie kobiety były matkami.*
two women were mothers$_{ins}$
'Two women were mothers.'

(91) a. *Pięć kobiet było zmęczonych.*
five women$_{gen}$ was$_{3rd,neut}$ tired$_{non-masc-hum,gen}$
'Five women were tired.'

b. *Pięć kobiet było matkami.*
five women$_{gen}$ was$_{3rd,neut}$ mothers$_{ins}$
'Five women were mothers.'

As discussed in section 3.3., Polish numerals require a special treatment in grammar in order to account for case of noun phrases in numeral phrases. As (91) shows a similar treatment is required for predicative adjectives: in (90a) the predicative adjective *zmęczone* is nominative whereas in (91a) it is genitive; this change parallels the change of the subject. Surprisingly, predicative nouns are not sensitive to the numeral form of the sentence subject, which is illustrated by (90b) and (91b).

As discussed in section 2.4.2. predicative adjectives in Polish show indexical agreement with the sentence subject. The following sign for the predicatively used adjective *zmęczony* can be inferred from the above discussion:[26]

[26] Although we find it very interesting, we will not discuss the case assignment to predicative adjectives in this paper.

(92)
$$\left[\begin{array}{l}\text{PHON PF}(\boxed{1},\boxed{2})\\ \text{STEM}|\text{PHON}\ \boxed{1}\ \langle zmęczony\rangle\\ \text{SYNSEM}|\text{LOCAL}\ \left[\begin{array}{ll}\text{CATEGORY} & \left[\begin{array}{l}\text{SUBJ}\ \langle\ \text{NP}\ \boxed{4}\ \rangle\\ \text{HEAD}|\text{AGR}\ \boxed{2}\\ pred\text{-}adj\text{-}pat\end{array}\right]\\ \text{CONTENT} & \left[\begin{array}{l}\text{INDEX}\ \boxed{4}\\ \text{RESTR}\ \{\ \left[\begin{array}{l}\text{RELN}\ tired\\ \text{ARG}\ \boxed{4}\end{array}\right]\ \}\end{array}\right]\end{array}\right]\end{array}\right]$$

where *pred-adj-pat* stands for the following structure, capturing the relevant indexical agreement pattern:

(93)
$$\left[\begin{array}{l}\text{COMPS}\ \langle\rangle\\ \text{SUBJ}\ \langle\ \text{NP}\ \left[\begin{array}{l}\text{GENDER}\ \boxed{3}\ gender\\ \text{PERSON}\ person\\ index\end{array}\right]\ \rangle\\ \text{HEAD}\ \left[\begin{array}{ll}\text{PRD} & +\\ \text{AGR} & \left[\begin{array}{l}\text{GENDER}\ \boxed{3}\\ \text{CASE}\ case\\ a\text{-}agr\end{array}\right]\end{array}\right]\\ pred\text{-}adj\text{-}pat\end{array}\right]$$

3. Case

3.1. Introduction

In this section, we will present a basic theory of case for Polish. This theory, developed in the framework of Head-driven Phrase Structure Grammar, is deeply embedded in the tradition of generative grammar (in the broad sense of the term, cf. [Pol93]) and it borrows freely from Chomsky's Government and Binding Theory.

In this work we will not deal with the relationship between meaning and case (if any); our approach will be — to use the terminology of [Mel86] — syntagmatically (rather than paradigmatically) oriented. We will also not succumb to the temptation of defining the notion of case. This (by no means trivial!) theoretical task is well outside the scope of this paper.[27]

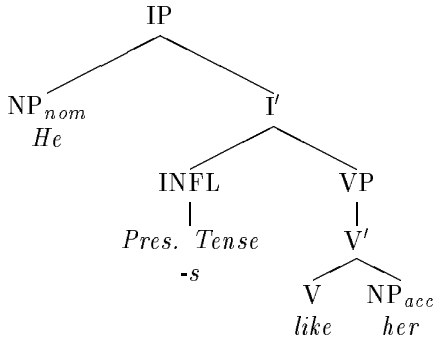
[27] See, however, [Mel86] and [Com86] for some attempts.

3.1.1. A Historical Note

It will not be an exaggeration to say that everything that has been written about case within the framework of HPSG stands in a strong relationship with the Case Theory of Government and Binding (GB). For [PS94] this relation is wholesale rejection, while for [HM94] it is development. This subsection will present some of the most prominent assumptions regarding case which exist within GB.

In GB, Case[28] is mainly a structural phenomenon: the Case an NP receives in a sentence depends on its position in the derivational tree of this sentence. Thus, complements of a verb get the accusative Case, while the subject receives the nominative Case:

(94)



This is explained by the fact that verbs which govern (i.e., are close to, in a certain configurational sense) their complements assign the accusative, while INFL nodes (representing a bundle of features related to tense and agreement; cf. section 2.1.1..), which govern subjects, assign the nominative Case.

This theory works quite well for English — it neatly explains some syntactic phenomena, in particular the phenomenon of passivization exemplified below.

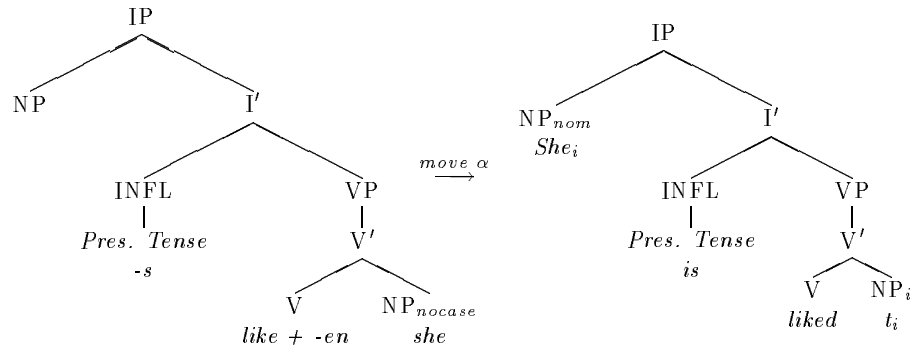
(95) He_{nom} likes her_{acc}. $\xrightarrow{pass}$ She_{nom} is liked.

According to GB, passivization is — simplifying a little — nothing more than adding the passive morpheme (*-en*) to the verb. This morpheme, in turn, has some very peculiar properties: it *absorbs* both the case assigned by the verb

[28] GB distinguishes between morphological case (written with the small 'c') and abstract Case (capital 'C'). Languages differ in the extent to which they exhibit case, but they are all assumed to have abstract Case.

in question, and the so-called external theta role.[29] The latter property simply means that thus affected verb does not sanction a subject, while the former that the complement of the verb does not receive Case. However, according to another principle of GB, the Case Filter, every (overt) NP *must* be assigned Case. Hence, in order to get Case, the complement has to move to a vacant position (leaving a trace behind) where Case can be assigned. Such a position is here the subject position which, according to the Extended Projection Principle, has to be present in the tree even if there is no subject (as in the case of passive verbs). But the Case it receives is not accusative anymore; as the complement is now governed by INFL, it receives the nominative Case. This is illustrated by (96) below:

(96)



Even though the Case Theory of GB might have worked well for English, it soon became clear that this purely configurational notion of Case cannot provide explanation for, e.g., more sophisticated course of passivization in German. The problem German poses is that passivization in this language affects Case assignment only in some instances. As the reader can easily verify, GB's Case Theory outlined above correctly accounts for (97), but fails in (98) below.[30] In the latter example *ihm* clearly retains its dative Case assigned by the verb.

(97) a. *Sie sieht ihn.*
She$_{nom}$ sees him$_{acc}$.
'She sees him.'

b. *Er wird gesehen.*
He$_{nom}$ is seen.

[29] The reader unfamiliar with GB is referred to [Hae91] for an excellent introduction to the theory.

[30] These examples come from [Hai85], cited here after [Hae91].

'He is seen.'

c. * *Ihn wird gesehen.*
He$_{acc}$ is seen.
'He is seen.'

(98) a. *Sie hilft ihm.*
She$_{nom}$ helps him$_{dat}$.
'She helps him.'

b. * *Er wird geholfen.*
He$_{nom}$ is helped.
'He is helped.'

c. *Ihm wird geholfen.*
He$_{dat}$ is helped.
'He is helped.'

These, and many other Case-related problems with GB's account (also in English, e.g., the problem of the so-called *of*-insertion) lead to substantial changes in the Case Theory and, as a result, in [Cho86a] Chomsky distinguishes between two types of Case assignment: structural, based — as before — on the position of NP at S-structure, and inherent, assigned by the lexical element at D-structure. Unlike structural Case, inherent Case is characterized by its stability: its morphological realization does not change with syntactic environment. Dative case in German (as well as, e.g., genitive in English) is analyzed as an instance of inherent case, though the extent to which a particular language realizes inherent case is a parameter of the theory; languages can differ significantly in this respect.

Many languages with rich declensional paradigms support this Case dichotomy in interesting ways. We will give examples of such supporting phenomena in the following sections.

3.1.2. Case in HPSG

There has been no separate theory of case within the framework of HPSG until very recently. [PS94] put considerable effort into rejecting GB's Case Theory and, by doing so, they announce that no theory of case (or Case) whatsoever is necessary: case is treated in HPSG only as a part of subcategorization requirements. Thus, for example, the value of the SUBCAT feature of the verb *like* would be:

(99) like: ⟨NP[*nom*], NP[*acc*]⟩

In this approach phenomena such as passivization are adjourned to the lexicon. More specifically, the Passive Lexical Rule takes care of permuting the complements within the SUBCAT list (cf. [PS87], p. 215 and [PS94], p. 130) and of changing the CASE values.

However, as [HM94] notice, "this approach fits well with arguments exhibiting the same case in all syntactic constructions [...] but makes it difficult to cope with complements showing variations of case depending on the syntactic context." Such variations exist in German and — appropriately enough — were first (within HPSG) taken into consideration in [NNP94] by [Pol94] and, especially, [HM94]. In this section, we will present some of the most prominent assumptions of Heinz and Matiasek concerning case; the reader is, however, referred to their article for details and some applications of their theory to German.

[HM94], following [Hai85], posit two kinds of case: lexical (an analogue of GB's inherent case) and structural. The morphological form of the former is determined via subcategorization requirements. This means that whenever a verb requires its complement to bear a lexical case, it also specifies the morphological realization of this case. Thus, the morphological case form of such a complement cannot vary with the syntactic environment.

Alternatively, a verb (or any other lexical element) can specify its complement as bearing a structural case. In this instance, the morphological form of the complement depends upon (and may alternate with) its syntactic environment, rather than on the lexical item (the verb) in question.[31]

According to [HM94], nominalization is an example of a test checking whether a given complement of a verb has lexical or structural case. This can be illustrated by the following examples:

(100) a. *Der Mann hilft mir.*
The man$_{nom}$ helps me$_{dat}$.
'The man is helping me.'

b. *das Helfen mir*
the helping me$_{dat}$
'the help for/*from me'

c. *das Helfen meiner*
the helping me$_{gen}$
'the help from/*for me'

[31] This lexical/structural case dichotomy parallels, of course, GB's inherent/structural case distinction. However, this division can be traced back to the direct vs. oblique case distinction of [Jak71] and grammatical vs. concrete case distinction of [Kur60].

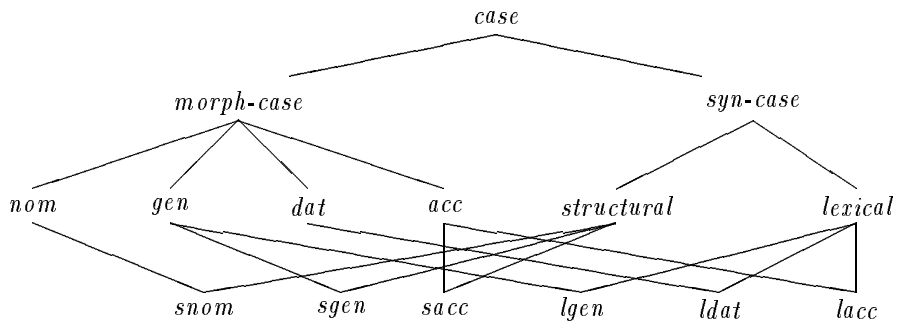
(101) a. *Der Mann unterstützt mich.*
The man$_{nom}$ helps me$_{acc}$.
'The man is helping me.'

b. *das Unterstützen meiner*
the helping me$_{gen}$
'the help for/from me'

Example (100) shows that the verb *helfen* requires a dative complement. The case of this complement does not change under nominalization, it remains dative. This should be contrasted with the behaviour of the complement of *unterstützen*; the case of this complement changes from accusative to genitive in the process (see (101)). Moreover, the case of the subject changes from nominative to genitive in both cases. These data suggest that dative is a lexical case, while nominative, genitive and accusative are structural: their distribution is a matter of configurational rules, rather than inherent properties of lexical items.

Of course, this does not mean that a given morphological case can only be either lexical or structural. [HM94], developing upon [Hai85], come up with the following type lattice depicting which morphological cases can be instances of which syntactic (i.e., lexical or structural) cases in German:

(102)



In this type lattice we follow the convention used in [HM94] of writing the most general type at the top of the lattice. So, the type *case* has as its subtypes *morph-case* and *syn-case*. The former determines the morphological cases German allows, i.e., its subtypes are *nom* (nominative), *gen* (genitive), *dat* (dative)

and *acc* (accusative). On the other hand, *syn-case* determines the syntactic properties of cases; each case marking on an NP is either *structural* or *lexical*. Structural cases are nominative (*snom*), genitive (*sgen*) or accusative (*sacc*), while lexical cases are genitive (*lgen*), dative (*ldat*) and accusative (*lacc*). This, of course, means, that NPs bearing genitive or accusative morphological case are either *structural* or *lexical*, depending on the case assigner. In the sequel of this paper we will usually abbreviate *structural* to *str* and *lexical* to *lex*.

[HM94] do not have much to say about lexical case: its morphological realization is entirely determined within the SUBCAT list, so it does not require a separate theory. On the other hand, it is assumed that lexical items do not specify the exact morphological realization of the NP[*str*] they subcategorize for. For example, verbs do not specify their subjects as NP[*nom*], but rather as NP[*str*]. It is the Case Principle, whose scope is somewhat analogous to GB's Case Theory, that determines the exact morphological realization of structural cases. Thus, for instance, values of SUBCAT for *helfen* and *unterstützen* look as follows:

(103) a. helfen: ⟨NP[*str*], NP[*ldat*]⟩
b. unterstützen: ⟨NP[*str*], NP[*str*]⟩

What morphological value a given instance of structural case gets is decided by the Case Principle which, in short, says that an NP[*str*] subject of a verb receives nominative (i.e., *snom*) case, structural complements of verbs receive accusative case (*sacc*), while structural complements of nouns receive genitive case (*sgen*). More precisely (cf. [HM94], p. 34), the following constraints have to be present in German grammar:

(104)
$$\begin{bmatrix} \text{SYNSEM|LOC|CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{VFORM } \textit{fin} \\ \textit{verb} \end{bmatrix} \\ \text{SUBCAT } \langle\ \rangle \\ \textit{cat} \end{bmatrix} \\ \text{DTRS} & \begin{bmatrix} \text{HEAD-DTR|} \ldots \text{|SUBCAT } \langle\ \text{NP}[\textit{str}], \ldots \rangle \\ \textit{h-c-str} \end{bmatrix} \\ \textit{phrase} \end{bmatrix}$$

$$\Longrightarrow \left[\text{DTRS|HEAD-DTR|} \ldots \text{|SUBCAT } \langle\ \text{NP}[\textit{snom}], \ldots \rangle\right]$$

(105)
$$\begin{bmatrix} \text{SYNSEM|LOC|CAT} & \begin{bmatrix} \text{HEAD } \textit{verb} \\ \text{SUBCAT } \langle\ \rangle \vee \langle\ \textit{synsem}\ \rangle \\ \textit{cat} \end{bmatrix} \\ \text{DTRS} & \begin{bmatrix} \text{HEAD-DTR|} \ldots \text{|SUBCAT } \langle\ \textit{synsem}, \text{NP}[\textit{str}], \ldots \rangle \\ \textit{h-c-str} \end{bmatrix} \\ \textit{phrase} \end{bmatrix}$$

$$\Longrightarrow \left[\text{DTRS}|\text{HEAD-DTR}|\ \ldots|\text{SUBCAT}\ \langle\ \mathit{synsem}, \text{NP}[\mathit{sacc}], \ldots\rangle\right]$$

(106)
$$\left[\begin{array}{ll}\text{SYNSEM}|\text{LOC}|\text{CAT} & \left[\begin{array}{l}\text{HEAD}\ \mathit{noun}\\ \text{SUBCAT}\ \langle\ \rangle \vee \langle\ \mathit{synsem}\ \rangle\end{array}\right]_{cat}\\ \text{DTRS} & \left[\text{HEAD-DTR}|\ \ldots|\text{SUBCAT}\ \langle\ \mathit{synsem}, \text{NP}[\mathit{str}], \ldots\rangle\right]_{h\text{-}c\text{-}str}\\ \multicolumn{2}{l}{}\end{array}\right]_{phrase}$$

$$\Longrightarrow \left[\text{DTRS}|\text{HEAD-DTR}|\ \ldots|\text{SUBCAT}\ \langle\ \mathit{synsem}, \text{NP}[\mathit{sgen}], \ldots\rangle\right]$$

The reader will recall that the symbol '$\Longrightarrow$' is used to denote *constraints*, i.e., *conditional feature structures* (cf. [PS87], p. 43). Thus, (104) should be understood as a condition imposed upon every feature structure of type *phrase*, modelling a finite verb phrase (VFORM fin) whose DTRS value is a structure of type *head-complement-structure*, and whose head-daughter specifies its first complement as NP[*str*]. This condition says, that the first complement of the head-daughter of such a phrase has to bear the nominative case (*snom*).

In the subsequent sections we will follow [HM94] in abbreviating conditional feature structures such as the ones above in the following form:

Case Principle
In a *head-complement-structure* of category

- **verb[fin]**: the structural subject has a CASE value of *snom*,
- **verb**: the structural object has a CASE value of *sacc*,
- **noun**: the structural object has a CASE value of *sgen*.

These are the only saturated or almost saturated *head-complement-structures* with structural arguments.[32]

Note, that, given the SUBCATs of *helfen* and *unterstützen* as in (103), this formulation of Case Principle correctly predicts the nominalization facts shown in the examples (100) and (101) on page 58. Of course, the underlying assumption here is that the only change which nominalization (implemented for example as a lexical rule) brings to the SUBCAT list is making each argument optional and possibly specifying that at most one argument is present.

In the remainder of this article we will examine to what extent the lexical vs. structural case dichotomy can be argued for Polish. We will also attempt to formulate an analogous Case Principle for this language.

[32] Structural subject should be understood as an NP element of the SUBJECT list (in the sense of [PS94], chapter 9) if it (is present and) bears *str* case. Similarly, by structural object we mean any structural NP element of COMPS. We find these notions more intuitive than, respectively, *external argument* and *internal argument* inherited from GB and used by [HM94]. See, however, their article for some motivation for this nomenclature.

3.2. Structural vs. Lexical Case in Polish

Polish, a language with rich inflectional morphology, shows considerable support for the lexical vs. structural case distinction. The only formal analysis (in the framework of GB) of how this distinction functions in Polish that we know of is [Wil90].[33] We will draw from this work freely, though our analysis will substantially differ from (the translation into HPSG of) that of Willim in many respects.

3.2.1. Morphological Case in Polish

There are seven morphological cases in Polish, though vocative can be argued not to be a case in the strict sense: it is used in isolation, mainly for getting attention and for addressing.[34] Of the remaining six, nominative never appears outside sentential subject position,[35] accusative is realized by verbal and prepositional arguments, genitive, dative and instrumental occur as arguments of all main lexical categories, and locative is restricted to the prepositional arguments.

3.2.2. Nominalization

As far as nominalization is concerned, Polish parallels German.[36] Consider the following data:

(107) a. *Janek* *pomaga* *Tomkowi.*
John$_{nom}$ helps Tom$_{dat}$.

[33] Some work has been done on analysis of case dichotomy in other Slavic languages, mainly Russian. The reader is referred to [Bab86], [Fra86], [Fra90], [Fra94] and references cited therein.

[34] See a.o. [Wil90], [Pol93] (p. 578, entry for *vocativus*) and [Str93] for arguments for this position, but also [SŚ85] p. 137 for important arguments against it.

[35] This statementy, as it stands, is too strong; cf. examples like (i) (pointed to us by Bob Borsley) or (ii) (from [SŚ85], p. 118) below:

(i) Jan, rozmawiałem z nim.
John$_{nom}$, talked$_{1st,sg}$ with him$_{ins}$.
'John, I talked to him.'

(ii) Przyjaciele wołają go Grubas.
Friends$_{nom}$ call him$_{acc}$ Fat$_{nom}$.
'The friend call him Fatty.'

[36] Cf. section 3.4.3. where we actually argue against this statement.

'John is helping Tom.'

b. *pomaganie Tomkowi*
helping Tom$_{dat}$
'the help for/*from Tom'

c. *pomaganie Janka*
helping John$_{gen}$
'the help from/*for John'

(108) a. *Janek pogardza Tomkiem*
John$_{nom}$ scorns Tom$_{ins}$.
'John scorns Tom.'

b. *pogardzanie Tomkiem*
scorning Tom$_{ins}$
'the scorn for/*from Tom'

c. *pogardzanie Janka*
scorning John$_{gen}$
'the scorn from/*for John'

(109) a. *Janek wspiera Tomka*
John$_{nom}$ supports Tom$_{acc}$.
'John is supporting Tom.'

b. *wspieranie Tomka (Janka)*
helping Tom$_{gen}$ (John$_{gen}$)
'the help for/from Tom (John)'

Examples (107) and (108) suggest that dative and instrumental cases are here instances of lexical case: they do not change under nominalization. On the other hand, as (109) shows, accusative is structural: the case changes to genitive in the process of nominalization. Of course, nominative and genitive are also structural cases here, just like in German.

On the basis of the above observations we can postulate the first version of the Case Principle for Polish:

Case Principle
In a *head-complement-structure* of category

- **verb**: the structural subject has a CASE value of *snom*,
  the structural object has a CASE value of *sacc*,
- **noun**: any structural argument (subject or object) has a CASE value of *sgen*.

3.2.3. Prepositions: Initial Remarks

It should be emphasized that the foregoing remarks are not relevant for those arguments which are prepositional phrases:[37]

(110) a. *Janek czeka na Marię.*
John$_{nom}$ waits on Mary$_{acc}$.
'John is waiting for Mary.'

b. *czekanie na Marię*
waiting on Mary$_{acc}$
'the waiting for Mary'

As the above example shows, accusative NPs which are arguments of prepositions do not change their case under nominalization.

Heinz and Matiasek analyze prepositions devoid of their inherent (locational) meaning as 'markers'. For them, prepositional phrases are just 'marked' nominal phrases, i.e., NP[+marked]. Hence, Case Principle can be applied to prepositional phrases. This, in turn, means that the accusative complement in example (110) cannot be structural: if it were structural then — according to Case Principle — it would have to change into genitive in the process of nominalization.

On the basis of analogous observations for German, [HM94] posit that prepositional (i.e., 'marked' in their terminology) arguments exhibiting accusative case are instances of lexical accusative (and exactly for this reason such 'marked' NPs do not change case under nominalization). Note that this failure of 'marked' (i.e., prepositional) arguments to change case under nominalization seems to be the only reason for allowing lexical accusatives; all other (i.e., 'unmarked') occurrences of accusative phrases are structural. Hence, Heinz and Matiasek miss the generalization that all 'unmarked' (i.e., true NP) accusative phrases are structural, while all 'marked' (i.e., prepositional) accusative phrases are lexical.

On our account there are no such coincidences: we do not follow [HM94] in analyzing prepositional phrases as 'marked' nominal phrases, but rather, traditionally, as true prepositional phrases.[38] This treatment of prepositional phrases explains the failure of prepositional accusative arguments to become genitive under nominalization (cf. (110a) and (110b)) — according to the Case Principle only *nominal* structural phrases alternate with environment, not prepositional phrases.

[37] Neither are they relevant for adverbial modifiers but here facts are less clear-cut. Unfortunately, discussion of case assignment to adverbial modifiers is outside the scope of this paper.

[38] Not much hinges on this decision, though.

We will also postulate that accusative case is always structural in Polish. This position is a consequence of the observation that there are no verbs in Polish subcategorizing for a *stable* accusative (or rather, structural) complement; accusative complements always become genitive under nominalization (and under genitive of negation, see below).[39] Of course, since prepositional arguments can bear accusative case, and we assume (cf. page 60) that lexical items never specify the morphological case of their structural complements, we have to add one more clause to the Case Principle:

Case Principle
In a *head-complement-structure* of category

...	...
• **noun**:	any structural argument (subject or object) has a CASE value of *sgen*,
• **preposition**:	the structural object has a CASE value of *sacc*.

We invite the reader to check that thus revised Case Principle and the lexical entry for the preposition *na* given (partially) below account for the example (110).

(111)
$$\begin{bmatrix} \text{PHON } \langle na \rangle \\ \\ \text{SYNSEM|LOC|CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{PREP } `na'+str \\ prep \end{bmatrix} \\ \text{SUBCAT } \langle \text{NP}[str] \rangle \\ category \end{bmatrix} \\ \\ word \end{bmatrix}$$

The next section shows that prepositional arguments are also not affected by Genitive of Negation.

3.2.4. Genitive of Negation

Another phenomenon of case variation is the so-called Genitive of Negation (GoN): an accusative object of a verb appearing in a declarative sentence changes its case marking to genitive under sentential negation. This is illustrated by the following example:

(112) a. *Janek lubi Marię.*
John$_{nom}$ likes Mary$_{acc}$.

[39] See section 3.4. for other arguments for the structurality of the accusative of prepositional arguments.

'John likes Mary.'

b. *Janek nie lubi Marii.*
John$_{nom}$ not likes Mary$_{gen}$.
'John doesn't like Mary.'

GoN does not exist in German but it is widespread in Slavic and exists also in some other languages (e.g., Finnish). GoN is a very unstable phenomenon: in many Slavic languages the accusative case expands rapidly taking the place of genitive in many constructions, not least in sentential negation. For example, in Czech GoN has practically ceased to exist; only the older generations use it in some restricted environments. In Russian, on the other hand, both accusative and genitive are allowed under sentential negation[40], while in Polish, even though accusative replaces genitive in many syntactic environments, genitive remains the only possibility under sentential negation.[41]

It is worth noticing that Genitive of Negation, just as nominalization, does not affect dative and instrumental complements. The examples below should be compared with (107)–(109) above:

(113) a. *Janek pomaga Tomkowi.*
John$_{nom}$ helps Tom$_{dat}$.
'John is helping Tom.'

b. *Janek nie pomaga Tomkowi.*
John$_{nom}$ not helps Tom$_{dat}$.
'John is not helping Tom.'

(114) a. *Janek pogardza Tomkiem*
John$_{nom}$ scorns Tom$_{ins}$.
'John scorns Tom.'

b. *Janek nie pogardza Tomkiem*
John$_{nom}$ not scorns Tom$_{ins}$.
'John doesn't scorn Tom.'

(115) a. *Janek wspiera Marię*
John$_{nom}$ supports Mary$_{acc}$.

[40] The reader is referred to [Tim86] for an analysis of distribution of accusative and genitive under negation and for defence of the hypothesis that GoN is in the state of withdrawal in Russian.

[41] Actually, this rule has a few exceptions. [BKS71] give two conditions when accusative is allowed. The first is semantical in nature: accusative is allowed when the sentence has a positive meaning despite its apparent negation. The second, which is structural, says that accusative is allowed when the complement is "far" from the finite verb. We do not try to model these exceptions in this paper.

'John is supporting Mary.'

b. *Janek nie wspiera Marii*
John$_{nom}$ not supports Mary$_{gen}$.
'John is not supporting Mary.'

Also prepositional arguments are not affected by negation. Again, the examples below parallel (110):

(116) a. *Janek czeka na Marię.*
John$_{nom}$ waits on Mary$_{acc}$.
'John is waiting for Mary.'

b. *Janek nie czeka na Marię.*
John$_{nom}$ not waits on Mary$_{acc}$.
'John is not waiting for Mary.'

These data independently confirm the distinction between structural and lexical case made in the previous section.[42] They also call for splitting one of the clauses of the Case Principle:

Case Principle
In a *head-complement-structure* of category

...	...
• **verb**:	the structural subject has a CASE value of *snom*,
• **verb**[−**neg**]:	the structural object has a CASE value of *sacc*,
• **verb**[+**neg**]:	the structural object has a CASE value of *sgen*,
...	...

3.2.5. Case Lattice for Polish

So far we have said nothing really new about structural/lexical case dichotomy; we have simply illustrated the distinction postulated by several authors working

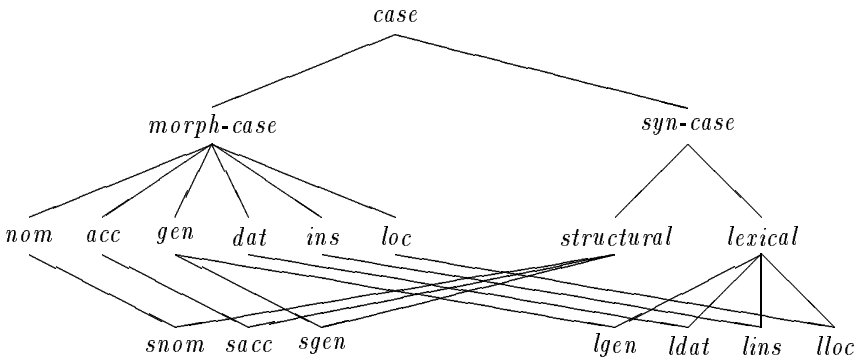
[42] We avoid here answering the question whether negation should be best analyzed as an essentially lexical phenomenon, or whether it should be dealt with in the syntax. A careful account of negation is needed in order to treat examples such as (ii) below (pointed to us by Bob Borsley), where the +NEG value seems to be shared between the matrix verb and its VP[inf] complement.

(i) Jan chiał widzieć Marię.
John$_{nom}$ wanted see$_{inf}$ Mary$_{acc}$.
'John wanted to see Mary.'

(ii) Jan nie chciał widzieć Marii.
John$_{nom}$ not wanted see$_{inf}$ Mary$_{gen}$.
'John didn't want to see Mary.'

within GB (Franks, Babby) to Polish. We have done this in the spirit of [HM94]. On the basis of the foregoing discussion we are ready to (tentatively) posit the following case lattice for Polish:

(117)



In the next sections we will try to analyze within the framework established so far some more ephemeral and idiosyncratic issues concerning case assignment in Polish, namely those of the case of numerals (section 3.3.) and so-called indefinite numerals (section 3.4.). We will also make a few remarks on passivization (section 3.5.).

3.3. Numerals

The complexity of numerals in Slavic languages is really daunting; Polish is no exception here. In general, this complexity is thought to be caused by the transitional character of numerals in Polish, see for example [BKS71] for a short assessment of the changes the system of numerals is undergoing currently. We will try not to forget about these diachronic considerations in what follows. However, our account will be mainly synchronical; we will attempt to analyze the phenomena involved as they stand.[43]

3.3.1. Basic Facts

In this subsection, we will deal with the most typical relationship basic numerals establish with noun phrases.

[43] For some analyses of numerals in other Slavonic languages the reader is referred to [Cor78] and [Fra94] (and references cited therein).

There are no nominative numeral phrases!

Consider the declension patterns shown below.[44]

(118) Non-masculine-human declension:

	these	*five*	*women*
NOM	te$_{nom/acc}$/tych$_{gen}$	pięć$_{nom/acc}$	kobiet$_{gen}$
GEN	tych$_{gen}$	pięciu$_{gen}$	kobiet$_{gen}$
DAT	tym$_{dat}$	pięciu$_{dat}$	kobietom$_{dat}$
ACC	te$_{nom/acc}$/tych$_{gen}$	pięć$_{nom/acc}$	kobiet$_{gen}$
INS	tymi$_{ins}$	pięcioma$_{ins}$	kobietami$_{ins}$
LOC	tych$_{loc}$	pięciu$_{loc}$	kobietach$_{loc}$

(119) Masculine-human declension:

	these	*five*	*men*
NOM	tych$_{gen/acc}$	pięciu$_{nom/gen/acc}$	mężczyzn$_{gen/acc}$
GEN	tych$_{gen}$	pięciu$_{gen}$	mężczyzn$_{gen}$
DAT	tym$_{dat}$	pięciu$_{dat}$	mężczyznom$_{dat}$
ACC	tych$_{gen/acc}$	pięciu$_{gen/acc}$	mężczyzn$_{gen/acc}$
INS	tymi$_{ins}$	pięcioma$_{ins}$	mężczyznami$_{ins}$
LOC	tych$_{loc}$	pięciu$_{loc}$	mężczyznach$_{loc}$

Note first that in both patterns there is total and unambiguous case concord between the determiner, the numeral and the noun phrase in four cases: genitive, dative, instrumental and locative (i.e., in the lexical cases). Notice also that in both declensions the nominative phrase is the same as the accusative one. This is expected as far as non-masculine-human declension is concerned, but quite surprising with respect to the masculine-human declension; in Polish, nominative and accusative cases are normally (i.e., in phrases with no numerals) different in masculine-human declension. Our account explains these facts.

[44] This is how such patterns should be understood: the NOM row contains these forms which can appear as subjects of typical verbs (such as *jeść*, 'eat', or *lubić*, 'like'); the ACC row contains these forms which can appear in an object position of typical transitive verbs (such as *lubić*, 'like'); the DAT row contains these forms which can appear as second objects of typical ditransitive verbs (such as *dać*, 'give'), etc. It is important to bear this point in mind as we will analyze some sentential subjects (i.e., forms appearing in the NOM row) as *accusative* phrases. Hence, there is no contradiction in a pattern containing the following row (cf. (125)):

(i) NOM tych$_{gen/acc}$ pięciu$_{acc}$ mężczyzn$_{gen}$

The subscripts in such patterns indicate the case values (of the nominal forms in question) which we initially deem possible (and relevant). We might argue against some of these case values in the ensuing discussion.

The crux of our analysis concerns the case ambiguities indicated in the NOM and ACC rows of examples (118) and (119) (cf. footnote 44). In order to try to resolve these case ambiguities we will make the natural assumption that in both declension patterns the elements in corresponding slots have the same case. For example, since in the non-masculine-human declension (cf. (118)) the case of *kobiet* in the NOM and ACC rows is unambiguously genitive, we will assume that the case of *mężczyzn* in the corresponding slots of the masculine-human pattern (cf. (119)) is also genitive (rather than accusative).

Before we proceed with resolving other case ambiguities, including the crucial one in the NOM row, we have to explicate what exactly we mean by case ambiguities here. Let us start with the NOM row of the non-masculine-human declension (118). The determiner *te* is marked as ambiguous between nominative and accusative. This means that *te* can occur only with nominative or accusative nouns (i.e., only in NOM and ACC rows), e.g.:

(120) a. *te* *kobiety*
these women$_{nom/acc}$

b. * *te* *kobiet/kobietom/kobietami/kobietach*
these women$_{gen/dat/ins/loc}$

Similarly, by marking *pięć* as ambiguous with respect to nominative and accusative case, we mean that it can appear only in the NOM and ACC rows of declension patterns. On the other hand, *tych* is unambiguously marked as genitive in the NOM and ACC; by that we indicate that it can co-occur only with genitive nouns (we consider the fact that it can also occur with locative nouns irrelevant here).

As far as the masculine-human declension pattern (119) is concerned, we have already seen that *mężczyzn* (NOM and ACC rows) has to be genitive by analogy to the non-masculine-human pattern (118). The crucial fact in the masculine-human declension pattern is that *tych* can normally choose only between genitive and accusative cases:

(121) a. *tych* *mężczyzn*
these men$_{gen/acc}$

b. * *tych* *mężczyźni/mężczyznom/mężczyznami/mężczyznach*
these men$_{nom/dat/ins/loc}$

Unlike in (118), the numeral in masculine-human pattern (119) is ambiguous with respect to three case values: nominative, accusative and genitive. Applying the same method that we have already used above, we can establish — again, by analogy with non-masculine-human pattern — that the case values

really at issue here are nominative and accusative. The considerations so far are summarized below:

(122) Non-masculine-human declension (extract):

	these	*five*	*women*
NOM	te$_{nom/acc}$/tych$_{gen}$	pięć$_{nom/acc}$	kobiet$_{gen}$
ACC	te$_{nom/acc}$/tych$_{gen}$	pięć$_{nom/acc}$	kobiet$_{gen}$

Masculine-human declension (extract):

	these	*five*	*men*
NOM	tych$_{gen/acc}$	pięciu$_{nom/acc}$	mężczyzn$_{gen}$
ACC	tych$_{gen/acc}$	pięciu$_{nom/acc}$	mężczyzn$_{gen}$

In order to show that the numeral phrases in the NOM (and ACC) row are really *accusative*, we will concentrate on the apparent disparity between the possible determiners in both patterns.

Note first that the non-masculine-human declension allows any of the two determiners *te*$_{nom/acc}$ (which agrees with the numeral *pięć*$_{nom/acc}$) and *tych*$_{gen}$ (which agrees with the noun *kobiet*$_{gen}$). In fact, *te* cannot be nominative; if it were nominative, then analogous (but masculine) nominative determiner should be also allowed in the masculine-human declension pattern. As the judgement below shows, this is definitely not the case:

(123) * *ci* *pięciu* *mężczyzn*
these$_{nom}$ five men

This means that, since *te* cannot be nominative, it has to be accusative. But if it is accusative, the numeral it agrees with also has to be accusative; in Polish determiners always agree (with respect to case) with the phrases they modify. So, the NOM and ACC rows of the non-masculine-human declension pattern finally look as follows:

(124) Non-masculine-human declension (extract):

	these	*five*	*women*
NOM	te$_{acc}$/tych$_{gen}$	pięć$_{acc}$	kobiet$_{gen}$
ACC	te$_{acc}$/tych$_{gen}$	pięć$_{acc}$	kobiet$_{gen}$

But this, in turn, means that the numeral in the masculine-human declension also has to be accusative:

(125) Masculine-human declension (extract):

	these	*five*	*men*
NOM	tych$_{gen/acc}$	pięciu$_{acc}$	mężczyzn$_{gen}$
ACC	tych$_{gen/acc}$	pięciu$_{acc}$	mężczyzn$_{gen}$

Note that the foregoing analysis has been conducted without any assumption as to the real structure of the nominal phrases as the ones above; i.e., we have not committed ourselves to any decision on what constitutes head of such phrases: it might be either a numeral (*pięciu*), or a noun phrase (*mężczyzn*). So far any analysis seems plausible, as far as we postulate a linear precedence (LP) rule stating that in (Polish) numeral phrases determiners precede numerals which, in turn, precede noun phrases. This LP rule, when applied to *pięć* $tych_{gen}$ $kobiet_{gen}$ gives $tych_{gen}$ *pięć* $kobiet_{gen}$.

However, now we commit ourselves to a particular analysis of numeral phrases: we will analyze them as *true* numeral phrases, i.e., headed by a numeral. This stance is consistent not only with the rows corresponding to lexical cases (GEN, DAT, INS and LOC), but also with the ones for ACC; indeed, if we analyzed such phrases as headed by a noun, we would have to explain why an accusative phrase is headed by a genitive noun. The ensuing sections will provide us with more arguments for such analysis of numeral phrases.

The most important conclusion of the foregoing discussion is this: numeral phrases in sentential positions cannot be nominative. Moreover, since they are headed by a numeral (and we have established that the case of the numeral in NOM row is accusative), such phrases have to bear the accusative case.[45] This conclusion explains a number of facts. One of them we have already seen: numeral phrases are the same in NOM and ACC rows even in the masculine-human declension, an apparent breach in a rule that masculine-human nominative phrases differ from masculine-human accusative phrases. Now the explanation of this deviation is obvious: nominative and accusative numeral phrases are the same because... there are no nominative numeral phrases! It is the *accusative* numeral phrases that fill the 'nominal' (sentential subject) positions.

Another crucial phenomenon our theory explains in an elegant way is the fact that whenever a numeral phrase occupies the sentential subject position, the verb of this sentence is marked as 3rd person singular neuter:

(126) a. *Dwóch facetów jadło jabłko.*
Two guys $\text{eat}_{3rd,sing,neut,past}$ apple.
'Two guys were eating an apple.'

b. *Pięć kobiet poszło do kina.*
Five women $\text{go}_{3rd,sing,neut,past}$ to cinema.

[45] It has to be emphasized that this is a very unorthodox result, although it has been signalled a.o. by [Zab89] and [Fra94]. The traditional grammarians analyze numeral phrases in subject position as headed by a genitive noun (cf. [Kle86], p. 121), while the formal Polish grammar by Saloni and Świdziński (cf. [SŚ85]) analyzes them as headed by a nominative numeral.

'Five women went to the cinema.'

Now, in view of our theory of agreement (cf. section 2.), the 3rd person singular neuter marking on the verb follows from the fact that the numeral phrase (in the examples above: *dwóch facetów* and *pięć kobiet*) is not nominative, and hence triggers the 3rd person singular neuter agreement pattern.

What are numerals?

The natural question that arises here is why numeral phrases have such heterogeneous declension patterns. In order to answer this question we will first of all posit that numerals are nouns, i.e., numeral phrases are really noun phrases. This position has two practical advantages over any other analyses of numerals: lexical items do not have to subcategorize separately for noun phrases and numeral phrases, and — more importantly — numeral phrases are in the scope of the Case Principle (as far as they are structural). Our modelling of numerals in terms of *sign*s will be based on the observation that there is no case agreement between the numeral and the NP it subcategorizes for only in NOM and ACC rows (cf. (118) and (119)), i.e., exactly in these cases which have to be structural (i.e., which do not have lexical counterparts; cf. case lattice for Polish (117) on page 68).

This observation leads us to the conclusion that there are two kinds of numerals, or rather that each numeral has to have two entries in the lexicon, one with lexical and one with structural case. This is exemplified below:[46]

(127) a.
$$\begin{bmatrix} \textsc{phon} & \textsc{pf}(\boxed{1},\boxed{2}) \\ \textsc{stem}|\textsc{phon} & \boxed{2}\langle pięć\rangle \\ \textsc{synsem}|\textsc{loc}|\textsc{cat} & \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textsc{numeral} & + \\ \textsc{agr} & \boxed{1}\ [\textsc{case}\ \boxed{3}\ lex] \end{bmatrix}_{nom} \\ \textsc{comps} & \langle \textsc{np}[\textsc{agr}|\textsc{case}\ \boxed{3}]\rangle \end{bmatrix}_{cat} \\ word \end{bmatrix}$$

b.
$$\begin{bmatrix} \textsc{phon} & \textsc{pf}(\boxed{1},\boxed{2}) \\ \textsc{stem} & \boxed{2}\langle pięć\rangle \\ \textsc{synsem}|\textsc{loc}|\textsc{cat} & \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textsc{numeral} & + \\ \textsc{agr} & \boxed{1}\ [\textsc{case}\ str] \end{bmatrix}_{nom} \\ \textsc{comps} & \langle \textsc{np}[\textsc{agr}|\textsc{case}\ gen]\rangle \end{bmatrix}_{cat} \\ word \end{bmatrix}$$

[46] Notice that the PHON feature is a function of STEM and AGR; we follow here [Katng].

Note that in our account numerals are simply nouns with their NUMERAL feature set to '+'. (All other nouns are specified as −NUMERAL.) Note also that in fact one lexical entry for each numeral will suffice in the lexicon in the technical sense as a simple lexical rule can be used to derive the other — this way we will be able to express the systematic relation between the two signs in (127). These signs will receive morphological case when combined with other words: lexical numerals will receive their case from the subcategorizing element via Subcat Principle, while structural numerals will get it via the Case Principle.[47] Of course, the Case Principle cannot remain as it is now or it would require nominative numeral phrases (which do not exist!) in the sentential subject position. Numeral phrases (and *only* numeral phrases) in subject positions are accusative, so the Case Principle has to depend on the feature NUMERAL. The revised version of this principle is shown below:

Case Principle
In a *head-complement-structure* of category

- **verb**: the structural subject has a CASE value of *snom* if −NUMERAL or *sacc* if +NUMERAL,
- **verb[−neg]**: the structural object has a CASE value of *sacc*,
- **verb[+neg]**: the structural object has a CASE value of *sgen*,
- **preposition**: the structural object has a CASE value of *sacc*,
- **noun**: any structural argument has a CASE value of *sgen*.

These are the only saturated or almost saturated *head-complement-structures* with structural arguments.

3.3.2. More Facts

The previous section describes the behaviour of numerals from *pięć* ('five') onwards (with exceptions, see below). Numerals *dwa* ('two') to *cztery* ('four') behave in Polish in a slightly different way; they adhere to the usual declension patterns and always agree with the noun phrases they govern:

(128) Non-masculine-human declension:

[47] We hope that the reader will not be deceived by the 'transformational' language that we use here to describe 'declarative' constraints.

	these	*three*	*women*
NOM	$\text{te}_{nom/acc}$	$\text{trzy}_{nom/acc}$	$\text{kobiety}_{nom/acc}$
GEN	tych_{gen}	trzech_{gen}	kobiet_{gen}
DAT	tym_{dat}	trzem_{dat}	kobietom_{dat}
ACC	$\text{te}_{nom/acc}$	$\text{trzy}_{nom/acc}$	$\text{kobiety}_{nom/acc}$
INS	tymi_{ins}	trzema_{ins}	kobietami_{ins}
LOC	tych_{loc}	trzech_{loc}	kobietach_{loc}

(129) Masculine-human declension:

	these	*three*	*men*
NOM	ci_{nom}	trzej_{nom}	mężczyzni_{nom}
GEN	$\text{tych}_{gen/acc}$	$\text{trzech}_{gen/acc}$	$\text{mężczyzn}_{gen/acc}$
DAT	tym_{dat}	trzem_{dat}	mężczyznom_{dat}
ACC	$\text{tych}_{gen/acc}$	$\text{trzech}_{gen/acc}$	$\text{mężczyzn}_{gen/acc}$
INS	tymi_{ins}	trzema_{ins}	mężczyznami_{ins}
LOC	tych_{loc}	trzech_{loc}	mężczyznach_{loc}

Considerations similar to those above lead us to the conclusion that these numerals behave like nouns as far as declension patterns are concerned. In particular, they can bear the nominative case, and the NOM and ACC rows in masculine-human declension (129) differ. Hence, we will analyze them as 'normal' (i.e., −NUMERAL) nouns:

(130)
$$\begin{bmatrix} \textsc{phon}\ \text{PF}(\boxed{1},\boxed{2}) \\ \textsc{stem}|\textsc{phon}\ \boxed{2}\langle trzy\rangle \\ \textsc{synsem}|\textsc{loc}|\textsc{cat} \begin{bmatrix} \textsc{head} \begin{bmatrix} \textsc{numeral} - \\ \textsc{agr}\ \boxed{1} \\ nom \end{bmatrix} \\ \textsc{comps}\ \langle \text{NP}[\textsc{agr}\ \boxed{1}]\rangle \\ cat \end{bmatrix} \\ word \end{bmatrix}$$

In fact (cf. [BKS71]), there is a tendency in modern Polish to adopt a uniform system of numerals. One of the symptoms of these changes is the behaviour of numerals *dwa* ('two') to *cztery* ('four'). Apart from the declension patterns (128) and (129) shown above, these numerals have an alternative masculine-human declension which parallels that of other numerals (such as *pięć*, cf. (119)):

(131) Masculine-human declension:

	these	*three*	*men*
NOM	$\text{tych}_{gen/acc}$	trzech_{acc}	mężczyzn_{gen}
GEN	tych_{gen}	trzech_{gen}	mężczyzn_{gen}
DAT	tym_{dat}	trzem_{dat}	mężczyznom_{dat}
ACC	$\text{tych}_{gen/acc}$	trzech_{acc}	mężczyzn_{gen}
INS	tymi_{ins}	trzema_{ins}	mężczyznami_{ins}
LOC	tych_{loc}	trzech_{loc}	mężczyznach_{loc}

As [BKS71] claim (p. 343), there is a rapid shift in the contemporary Polish towards the usage of the latter form of the masculine-human declension patterns. That is, in terms of our *sign* feature structures, there is a rapid expansion of lexical entries such as (127).

One more example of this tendency is given by the numerals such as *tysiąc* ('thousand'), *milion* ('million'), etc. Traditionally, they are analyzed as nouns which always assign genitive case to their NP complements:

(132)

	thousand	*men*	(*women*)
NOM	$\text{tysiąc}_{nom/acc}$	mężczyzn_{gen}	(kobiet_{gen})
GEN	tysiąca_{gen}	mężczyzn_{gen}	(kobiet_{gen})
DAT	tysiącu_{dat}	mężczyzn_{gen}	(kobiet_{gen})
ACC	$\text{tysiąc}_{nom/acc}$	mężczyzn_{gen}	(kobiet_{gen})
INS	tysiącem_{ins}	mężczyzn_{gen}	(kobiet_{gen})
LOC	tysiącu_{loc}	mężczyzn_{gen}	(kobiet_{gen})

The reader will immediately notice that numeral phrases involving *tysiąc*, etc. crucially have to be analyzed as true numeral phrases headed by *tysiąc*: the case of the whole phrase is the same as the case of the numeral, while the subcategorized NP is always genitive. This provides us with one more argument, an argument of uniformity, for analyzing numerals in *all* numeral phrases as heads.

Notice also that *tysiąc* cannot be analyzed here just as a −NUMERAL noun for the reasons we give presently. In Polish *tysiąc* has the masculine gender. If, when in sentential subject position, it were really just a normal (i.e., −NUMERAL) nominative noun, it would take part in gender agreement with the past tense verb as all other nominative nouns do (see sections 2.4.4. and 2.4.4. for details). This is, however, not the case:

(133) a.

Tysiąc	*mężczyzn*	*poszło*	*do*	*pracy.*
$\text{One thousand}_{nom/acc,\text{masc}}$	men	$\text{went}_{3rd,sing,\text{neut}}$	to	work.

'One thousand men went to work.'

b. * *Tysiąc* *mężczyzn* *poszedł* *do*
One thousand$_{nom/acc,\mathrm{masc}}$ men went$_{3rd,sing,\mathrm{masc}}$ to
pracy.
work.
'One thousand men went to work.'

This[48] strongly suggests that the noun (numeral) phrase *tysiąc mężczyzn* is not assigned nominative case. But this fits well in the picture drawn so far: we will posit that *tysiąc*, *milion*, etc. are +NUMERAL nouns which combine with genitive (plural) NPs but which are not themselves specified for case:

(134)
$$\begin{bmatrix} \textsc{phon}\ \mathrm{PF}(\boxed{1},\boxed{2}) \\ \textsc{stem}|\textsc{phon}\ \boxed{2}\langle \textit{tysiąc}\rangle \\ \textsc{synsem}|\textsc{loc}|\textsc{cat} \begin{bmatrix} \textsc{head} \begin{bmatrix} \textsc{numeral} + \\ \textsc{agr}\ \boxed{1} \\ nom \end{bmatrix} \\ \textsc{comps}\ \langle \mathrm{NP}[\textsc{agr}|\textsc{case}\ \textit{gen}]\rangle \\ cat \end{bmatrix} \\ word \end{bmatrix}$$

Notice, that phrases headed by this sign can appear both in environments requiring lexical case and in those requiring structural case.[49] As usual, in the former instance case values will be assigned by a lexical element, while in the latter — by the Case Principle. In particular, numeral phrases in subject position headed by *tysiąc*, being +NUMERAL and not specified for the CASE value (i.e., potentially structural), will be assigned the structural accusative case and, hence, trigger the 3rd person singular neuter agreement pattern on the verb as predicted by our analysis of agreement in section 2. This explains the ungrammaticality of (133b) where the verb *poszedł* is marked as masculine.

3.3.3. Collective Numerals

There is one more kind of numerals that we have said nothing about: the so-called collective numerals.[50] This is a group of numerals used with noun phrases describing people of mixed sex (e.g., *pięcioro studentów*, 'five students (of mixed sex)'), children (*czworo dzieci*, 'four children'), small animals (*troje*

[48] Some speakers find (133b) also acceptable. This suggests that their lexical entries for *tysiąc* are ambiguous with respect to the NUMERAL feature.

[49] Actually, in this respect *tysąc* behaves like a 'normal' (−NUMERAL) noun.

[50] In this section we will draw heavily on observations made by [BKS71].

kurcząt, 'three chickens'), and with some *plurale tantum* nouns (*pięcioro drzwi*, 'five doors'). However, as these numerals are to some extent redundant and their declension is very ill-behaved (see below), the 'regular' numerals take over in contemporary Polish.[51] Thus, in the examples below, the (b) form, although still not sanctioned by the linguistic norm, becomes more and more widespread:[52]

(135) a. *W klasie było dwadzieścioro dzieci.*
In class were$_{3rd,sing,neut}$ twenty$_{coll,acc}$ children$_{gen}$.
'There were twenty children in the class.'

b. *W klasie było dwadzieścia dzieci.*
In class were$_{3rd,sing,neut}$ twenty$_{reg,acc}$ children$_{gen}$.
'There were twenty children in the class.'

(136) a. *Zapukał do pięciorga drzwi.*
Knocked$_{3rd,sing,masc}$ to five$_{coll,gen}$ doors$_{gen}$.
'He knocked at five doors.'

b. *Zapukał do pięciu drzwi.*
Knocked$_{3rd,sing,masc}$ to five$_{reg,gen}$ doors$_{gen}$.
'He knocked at five doors.'

Simultaneously, apart from this quantitative change, collective numerals undergo another, qualitative change in declension pattern. Consider first the current declension pattern of these numerals:

(137)

	five	*children*
NOM	pięcioro$_{acc}$	dzieci$_{gen}$
GEN	pięciorga$_{gen}$	dzieci$_{gen}$
DAT	pięciorgu$_{dat}$	dzieciom$_{dat}$
ACC	pięcioro$_{acc}$	dzieci$_{gen}$
INS	pięciorgiem$_{ins}$	dzieci$_{gen}$
LOC	pięciorgu$_{loc}$	dzieciach$_{loc}$

[51] They have already taken over to a large extent: collective numerals were used with any noun phrases once.

[52] These examples are taken from [BKS71], p. 30. She also notes that the supersession of collective numerals by regular numerals becomes visible in the linguistic norm which allows both following constructions:

(i) *W przedszkolu jest trzydzieścioro/trzydzieści czworo dzieci.*
In kindergarten is$_{3rd,sing}$ thirty$_{\text{coll/reg},acc}$ four$_{\text{coll},acc}$ children$_{gen}$.
'There are thirty four children in the kindergarten.'

Notice that this pattern differs from that of 'regular' numerals (cf. (118) and (119)) in the INS row: the numeral requires a genitive (rather than instrumental) noun phrase here. Apparently, this awkward behaviour of collective numerals is caused by their transitional status from the 'agreeing' pattern[53] (like that of regular numerals, e.g., *pięć*) to the 'governing' pattern (like that of *tysiąc*). The existence of this process is confirmed by the fact that a steady shift towards the regular governing pattern (especially in locative) can be observed in contemporary Polish.[54]

Of course, these diachronic considerations do not relieve us from the duty of modelling the current state of the language. Thus, on the basis of the foregoing discussion, we will posit the following lexical entries for collective numerals:

(138) a.

$$\begin{bmatrix} \textsc{phon}\ \mathrm{PF}(\boxed{1},\boxed{2}) \\ \textsc{stem}|\textsc{phon}\ \boxed{2}\langle \textit{pięcioro}\rangle \\ \textsc{synsem}|\textsc{loc}|\textsc{cat} \begin{bmatrix} \textsc{head} \begin{bmatrix} \textsc{numeral} + \\ \textsc{agr}\ \boxed{1}\ [\textsc{case}\ \boxed{3}\ \textit{lex} \wedge \neg \textit{lins}] \end{bmatrix}_{nom} \\ \textsc{comps}\ \langle \mathrm{NP}[\textsc{agr}|\textsc{case}\ \boxed{3}]\rangle \end{bmatrix}_{cat} \\ \textit{word} \end{bmatrix}$$

b.

$$\begin{bmatrix} \textsc{phon}\ \mathrm{PF}(\boxed{1},\boxed{2}) \\ \textsc{stem}|\textsc{phon}\ \boxed{2}\langle \textit{pięcioro}\rangle \\ \textsc{synsem}|\textsc{loc}|\textsc{cat} \begin{bmatrix} \textsc{head} \begin{bmatrix} \textsc{numeral} + \\ \textsc{agr}\ \boxed{1}\ [\textsc{case}\ \textit{str} \vee \textit{lins}] \end{bmatrix}_{nom} \\ \textsc{comps}\ \langle \mathrm{NP}[\textsc{agr}|\textsc{case}\ \textit{gen}]\rangle \end{bmatrix}_{cat} \\ \textit{word} \end{bmatrix}$$

These feature structures differ from those for 'regular' numerals (cf. 127) minimally; the only difference is slightly more complex values of AGR|CASE feature in (138).

[53] Of course, the numeral 'agrees' with the noun phrase only in lexical cases, while it 'governs' it (i.e., requires genitive case) in structural cases.

[54] The interesting question is what rules — if, indeed, any — govern the changes described above, that is, why the transition from 'agreeing' to 'governing' valency of collective numerals started in instrumental case, why the transition from Genitive of Negation to the lack of it in Russian takes place as described in [Tim86], etc. These matters are, of course, well outwith the scope of this work.

3.4. Indefinite Numerals

3.4.1. The Analysis

In this section we will show that the vast majority of the so-called indefinite numerals patterns the numerals described in the previous section. We will also present an intriguing puzzle, the highly idiosyncratic behaviour of an indefinite numeral *dużo* ('a lot of'), and argue that this idiosyncrasy confirms in a striking way the account of case in Polish given so far.

Morphologically, indefinite numerals constitute a very heterogeneous class. They are traditionally (cf. [BKS71], pp. 341–343) divided into pronominal numerals (*tyle*, *ile*, etc.), adjectival numerals (*dużo*, *wiele*) and nominal numerals (*szereg*, *część*), but this will not concern us here. What is important for us is their valency. From this point of view we can split indefinite numerals into three classes:

CLASS 1 Indefinite numerals which parallel 'ordinary' numerals (such as *pięć*, 'five'). These are mainly pronominal numerals and some adjectival numerals, e.g., *wiele* ('many'), *kilka* ('a few'), *ile* ('how many'), *tyle* ('that many'), *parę* ('a couple'), etc. Their lexical entries will be almost identical with those of 'ordinary' numerals (see declension patterns (118)–(119) and feature structures (127)).[55]

CLASS 2 Indefinite numerals which behave like *tysiąc*, *milion*, etc. (See declension patterns (132) and feature structure (134).) These are mainly nominal numerals such as *mnóstwo*, *mnogość* ('lots of'), *szereg* ('series'), *część* ('part of'), etc.

Due to the common (in Polish) phenomenon of numeralization, CLASS 2 is currently the most actively expanding group of numerals. Numeralization is the process of transforming nouns into numerals. As we noted above, numerals (apart from 1–4) do not bear nominative case; instead the Case Principle assigns structural accusative case to those which function as sentential subjects. This, in turn, triggers the 3rd person singular neutral agreement pattern rather than the usual subject-verb agreement pattern. Hence, in practice, numeralization can be witnessed when what used to be a nominal phrase co-occurs with 3rd person singular neuter verb. Examples of nouns that seem to be undergoing the process currently are (cf. [BKS71], p. 347) *szereg* ('series'), *moc* ('plenty'), *część* ('part of'). The usage is shifting from patterns such as (139a) towards (139b).

[55] We will not deal here with the exact form of the CONTENT value.

(139) a. *Szereg osób wiedział o tym.*
Series$_{nom/acc,sing,\text{masc}}$ people$_{gen}$ knew$_{3rd,sing,\text{masc}}$ about this.
'A series of people knew about this.'

b. *Szereg osób wiedziało o tym.*
Series$_{nom/acc,sing,\text{masc}}$ people$_{gen}$ knew$_{3rd,sing,\text{neut}}$ about this.
'A series of people knew about this.'

Of course, in terms of our feature structure numeralization is simply a change of value of NUMERAL from '−' to '+'. We do not have much to say about lexical entries of CLASS 2 indefinite numerals as they closely match that of *tysiąc* (cf. (134)).[56]

The most interesting class of indefinite numerals is, however, CLASS 3:

CLASS 3 Indefinite numerals which are traditionally analyzed as having only nominative and accusative forms (cf. [Dor80]), e.g., *dużo* ('a lot'), *mało* ('little'), *trochę* ('a little'), *sporo* ('quite a lot'), etc.

Numerals such as *dużo* ('a lot') do not decline, they always have the same nominative/accusative form and always combine with genitive NPs. Below we present the defective declension pattern for *dużo*:[57]

(140)

	a lot of	*men*	(*women*)
NOM	dużo$_{acc}$	mężczyzn$_{gen}$	(kobiet$_{gen}$)
GEN	—	—	—
DAT	—	—	—
ACC	dużo$_{acc}$	mężczyzn$_{gen}$	(kobiet$_{gen}$)
INS	—	—	—
LOC	—	—	—

The puzzle concerning these numerals is that they are grammatical in some positions which normally require genitive case, but not in others:

(141) a. *Nie mam w domu (zbyt) dużo chleba.*
Not have$_{1st,sing}$ in home (too) a lot of$_{nom/acc}$ bread$_{gen}$.

[56] It is perhaps worth noting here that the same behaviour is also exhibited by the so-called fractional numerals (*półtora*, 'one and a half', *dwie trzecie*, 'two thirds', etc.) and, to some extent, collective numerals (see section 3.3.3.). Again, these numerals have to be analyzed as heads of the nominal phrases they occur in.

[57] *Dużo*-phrases, when subjects of sentences, always trigger the 3rd person singular neuter agreement patters. This means, that just as other numeral phrases, they should be analyzed as accusative, rather then nominative, phrases.

'I don't have (too) much bread at home.'

b. *Nie mam w domu chleba.*
Not have$_{1st,sing}$ in home bread$_{gen}$.
'I don't have bread at home.'

c. * *Nie mam w domu chleb.*
Not have$_{1st,sing}$ in home bread$_{acc}$.
'I don't have bread at home.'

(142) a. *Nie lubię dużo osób.*
Not like$_{1st,sing}$ a lot of$_{nom/acc}$ people$_{gen}$.
'I don't like a lot of people.'

b. *Nie lubię tych osób.*
Not like$_{1st,sing}$ these$_{gen}$ people$_{gen}$.
'I don't like these people.'

c. * *Nie lubię te osoby.*
Not like$_{1st,sing}$ these$_{acc}$ people$_{gen}$.
'I don't like these people.'

(143) a. * *Boję się dużo osób.*
Fear$_{1st,sing}$ REFL a lot of$_{nom/acc}$ people$_{gen}$.
'I am afraid of a lot of people.'

b. *Boję się tych osób.*
Fear$_{1st,sing}$ REFL these$_{gen}$ people$_{gen}$.
'I am afraid of these people.'

c. * *Boję się te osoby.*
Fear$_{1st,sing}$ REFL these$_{nom/acc}$ people$_{nom/acc}$.
'I am afraid of these people.'

In the examples above, the (b) and (c) sentences show that a genitive NP is required by *nie mam* ('I don't have'), *nie lubię* ('I don't like') and *boję się* ('I am afraid') (see (b)), and that it cannot be realized by an accusative phrase (see (c)). However, in (141a) and (142a) *dużo*-phrases are allowed, while in (143a) they are not.[58]

The careful reader will have noticed that these examples themselves suggest an answer to the quandary: *dużo*-phrases are allowed under the Genitive

[58] Actually, some speakers feel uncomfortable with (142a), but they always deem it more grammatical than (143a).

of Negation, but not as a genitive complement of a verb.[59] This, and the underlying assumption we made implicitly, namely that verbs requiring genitive complements specify them as *lexical* genitive, suggests that the indefinite numerals of CLASS 3 can only be assigned structural case, never lexical. Thus, the ungrammaticality of (143a) stems from the fact that *boję się* subcategorizes for a lexical genitive phrases (which cannot be realized by *dużo*-phrases), while the grammaticality of (141a) and (142a) is a consequence of the fact, that *mieć* ('have') and *lubić* ('like') require a structural complement (which can be realized by *dużo*-phrases).

These considerations lead us to postulating the following lexical entry for *dużo*:

(144)
$$\begin{bmatrix} \text{PHON } \langle \textit{dużo} \rangle \\ \\ \text{SYNSEM|LOC|CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{NUMERAL } + \\ \text{AGR|CASE } \textit{str} \\ \textit{nom} \end{bmatrix} \\ \text{COMPS } \langle \text{NP}[\text{AGR|CASE } \textit{gen}] \rangle \\ \textit{cat} \end{bmatrix} \\ \\ \textit{word} \end{bmatrix}$$

Notice that the interaction of the Case Principle, agreement patterns, and simple lexical entries accounts in a very elegant way for the quirky behaviour of CLASS 3 indefinite numerals. Being +NUMERAL, structural and nominal, these indefinite numerals get (via the Case Principle) accusative case (*sacc*) when in sentential subject position. As they bear a case different from nominative, they trigger the 3rd person singular neuter agreement pattern. This in turn means that the verb has the 3rd person singular neuter agreement features:

(145) *Dużo osób poszło do domu.*
A lot of$_{acc}$ people$_{gen}$ went$_{3rd,sing,neut}$ to home.
'A lot of people went home.'

3.4.2. Some Ramifications

Our analysis of *dużo* supports many of the decisions we have taken in the previous sections. In this subsection we will point out two of them: the analysis of numeral phrases as headed by a numeral (see p. 72), and the analysis of accusative complements of prepositions as structural (see section 3.2.3., p. 64). We will start with the latter.

[59] [SŚ85] seem to simplify things suggesting (p. 83) that *dużo*-phrases are allowed with verbs and disallowed as complements of nouns: examples like (143a) are clearly ungrammatical for all the native speakers we have consulted.

The crucial observation that we will employ here is that *dużo*-phrases are specified as bearing an arbitrary structural case, but *only* structural case. This analysis has allowed us to explain the extremely idiosyncratic behaviour of *dużo*-phrases, esp. the fact that these phrases seem to be allowed in some genitive environments, while disallowed in others. In other words, we have found ourselves in possession of a convenient test for checking structurality of any given environment.

This test confirms our analysis of accusative prepositional arguments:

(146) a. *Maria czeka na dużo osób.*
Mary waits on a lot of people.
'Mary is waiting for a lot of people.'

b. *Janek przejeżdżał przez dużo wsi i miasteczek.*
John went through a lot of villages and towns.
'John went through a lot of villages and towns.

On the other hand, if we analyzed prepositions as markers, we would have two options. One would be to assume that all 'marked' nominal phrases are lexical (cf. [HM94]), but this would contradict the judgements above (because *dużo*-phrases are structural). The other would be to allow structural 'marked' phrases, but then we would have to add several clauses for 'marked' nominal phrases to the Case Principle (each corresponding to a different category of *head-complement-structure*: verb, noun, etc.). By contrast, our analysis allows us to add just one clause to the Case Principle (that for prepositions) and to get rid of lexical accusative altogether.

The other point we want to emphasize here is more fundamental. We have already given several arguments for analyzing numeral phrases as *true* numeral phrases, i.e., as phrases headed by a numeral. Our analysis of *dużo*-phrases provides us with one more, essentially an argument of uniformity (similar to the one mentioned in section 3.3.2.). It is crucial that in phrases such as *dużo osób* ('a lot of people') it is the numeral that heads the phrase. If it were the genitive noun (*osób*), then it would be very difficult to account for the following judgements:

(147) a. * *Boję się dużo osób.*
Fear$_{1st,sing}$ REFL a lot of$_{nom/acc}$ people$_{gen}$.
'I am afraid of a lot of people.'

b. *Boję się wielu osób.*
Fear$_{1st,sing}$ REFL a lot of$_{gen}$ people$_{gen}$.
'I am afraid of a lot of people.'

c. *Boję się tych osób.*
Fear$_{1st,sing}$ REFL these$_{gen}$ people$_{gen}$.

'I am afraid of these people.'

If *osób* were to be the head here, then, in order to explain the ungrammaticality of (147a) and the grammaticality of (147b) and (147c), we would have to postulate that *bać się* subcategorizes for a genitive NP *not* modified by any indefinite numerals of CLASS 3 (cf. (147a)) but possibly modified by some other numeral (cf. (147b)), or not modified at all (cf. (147c)); a highly *ad hoc* explanation to give. On the other hand, upon our account the NP *dużo osób* is headed by the numeral[60] whose case is specified simply as *str*. This means that the case value of the whole NP is *str* (by the Head Feature Principle), and, thus, it cannot fulfill syntactic requirements of the verb *bać się* subcategorizing for a lexical NP (specifically, for NP[CASE *lgen*]). On the other hand, the indefinite numerals *wielu* and *tych* behave like most numerals and can bear either lexical (as in the examples above) or structural case.

3.4.3. Nominalization Revisited

In section 3.2.2. we have stated that "Polish parallels German" as far as nominalization is concerned. Now, in view of some foregoing results, we will have to change our view on the matter.

We will again apply the test on structurality of a given environment provided by *dużo*-phrases, this time to investigate complements of nominalized verbs. We illustrate our considerations with the transitive verb *zjeść* ('eat').

(148) *Janek zjadł dużo rodzynek.*
John$_{nom}$ ate a lot of raisins$_{gen}$.
'John has eaten a lot of raisins.'

(149) *Janek nie zjadł dużo rodzynek.*
John$_{nom}$ not ate a lot of raisins$_{gen}$.
'John has not eaten a lot of raisins.'

The above examples show that the object of the verb is structural and — just as predicted by the Case Principle and the lexical entry for *dużo* (cf. (144)) — *dużo rodzynek* is allowed as an object.

However, judgements such as the one below seem to contradict the Case Principle:

(150) * *Zjedzenie dużo rodzynek przez Janka mogło mu*
Eating a lot of raisins by John might have he$_{dat}$
zaszkodzić.
harm.

[60] Remember that numerals *are* nouns in our account!

'John's eating many raisins might have harmed him.'

The unacceptability of the above sentence cannot be a matter of semantic restrictions as the sentence below having the same meaning as (150) is perfectly grammatical.

(151) *Zjedzenie* *wielu* *rodzynek* *przez* *Janka* *mogło* *mu* *zaszkodzić.*
Eating a lot of raisins by John might have he$_{dat}$ harm.
'Eating many raisins might have harmed John.'

Note that the implicit assumption in our (as well as that of [HM94]) rendering of nominalization is that this process, realized as a lexical rule, does not change CASE values of SUBCAT elements. In other words, structural arguments of a verb stay structural as arguments of deverbal nouns. Examples like (151) argue against maintaining this assumption. Instead, we will assume that the nominalization lexical rule changes all structural CASE values of SUBCAT elements to *lgen*.

Such an analysis should not seem ad hoc as there are independent reasons for the nominalization lexical rule to make changes in SUBCAT. The most conspicuous change concerns sentential subjects. As the examples below show, they can be realized either by NP[*gen*] (cf. (152)), or by PP[PREP '*przez*' + *acc*] (cf. (153)):

(152) a. *Maria* *czeka* *na* *Janka.*
Mary$_{nom}$ waits on John$_{acc}$.
'Mary is waiting for John.'

b. *czekanie* *Marii* *na* *Janka*
waiting Mary$_{gen}$ on John$_{acc}$
'Mary's waiting for John'

(153) a. *Jan* *je* *rodzynki.*
John$_{nom}$ eats raisins$_{acc}$.
'John is eating raisins.'

b. *jedzenie* *rodzynek* *przez* *Jana*
eating raisins$_{gen}$ by John$_{acc}$
'John's eating raisins'

Thus, in the process of nominalization the SUBCAT list changes considerably.

We will not attempt to formally state the nominalization lexical rule here, as its technical characterization could only distort the picture drawn above. Instead, we will give examples of operation of this rule:

(154)
$$\begin{bmatrix} \textsc{phon} \langle czekać \rangle \\ \textsc{synsem|loc|cat} \begin{bmatrix} \textsc{head} \ verb \\ \textsc{subcat} \langle \mathrm{NP}_{\boxed{1}}[str], \ \mathrm{PP}_{\boxed{2}}[\textsc{prep} \ `na' + str] \rangle \\ category \end{bmatrix} \\ word \end{bmatrix}$$
$$\mapsto$$
$$\begin{bmatrix} \textsc{phon} \langle czekanie \rangle \\ \textsc{synsem|loc|cat} \begin{bmatrix} \textsc{head} \ noun \\ \textsc{subcat} \langle (\mathrm{NP}_{\boxed{1}}[lgen]), (\mathrm{PP}_{\boxed{2}}[\textsc{prep} \ `na' + str]) \rangle \\ category \end{bmatrix} \\ word \end{bmatrix}$$

(155)
$$\begin{bmatrix} \textsc{phon} \langle jeść \rangle \\ \textsc{synsem|loc|cat} \begin{bmatrix} \textsc{head} \ verb \\ \textsc{subcat} \langle \mathrm{NP}_{\boxed{1}}[str], \ \mathrm{NP}_{\boxed{2}}[str] \rangle \\ category \end{bmatrix} \\ word \end{bmatrix}$$
$$\mapsto$$
$$\begin{bmatrix} \textsc{phon} \langle jedzenie \rangle \\ \textsc{synsem|loc|cat} \begin{bmatrix} \textsc{head} \ noun \\ \textsc{subcat} \langle (\mathrm{NP}_{\boxed{1}}[lgen]), (\mathrm{PP}_{\boxed{2}}[\textsc{prep} \ `przez' + str]) \rangle \\ category \end{bmatrix} \\ word \end{bmatrix}$$

(156)
$$\begin{bmatrix} \textsc{phon} \langle pomagać \rangle \\ \textsc{synsem|loc|cat} \begin{bmatrix} \textsc{head} \ verb \\ \textsc{subcat} \langle \mathrm{NP}_{\boxed{1}}[str], \ \mathrm{NP}_{\boxed{2}}[ldat] \rangle \\ category \end{bmatrix} \\ word \end{bmatrix}$$
$$\mapsto$$
$$\begin{bmatrix} \textsc{phon} \langle pomaganie \rangle \\ \textsc{synsem|loc|cat} \begin{bmatrix} \textsc{head} \ noun \\ \textsc{subcat} \langle (\mathrm{NP}_{\boxed{1}}[lgen]), (\mathrm{NP}_{\boxed{2}}[ldat]) \rangle \\ category \end{bmatrix} \\ word \end{bmatrix}$$

The first two examples ((154) and (155)) correspond to the nominalization examples (152) and (153). In these examples all structural nominal phrases become lexical genitive, all the arguments become optional, and the nominative subject is changed to a prepositional phrase (example (155)). The last example

shows that lexical complements (in this case *ldat*) do not change their case in the process of nominalization.

By positing such a lexical rule we have transferred part of the scope of Case Principle to the lexicon. The revised Case Principle will be rid of the noun clause:

Case Principle
In a *head-complement-structure* of category

- **verb**: the structural subject has a CASE value of *snom* if −NUMERAL or *sacc* if +NUMERAL,
- **verb[−neg]**: the structural object has a CASE value of *sacc*,
- **verb[+neg]**: the structural object has a CASE value of *sgen*,
- **preposition**: the structural object has a CASE value of *sacc*.

These are the only saturated or almost saturated *head-complement-structures* with structural arguments.

The reader familiar with Chomsky's GB will note that now the Case Principle is compatible with the independently motivated Case Assignment Principle of GB which states that "an NP receives Case at S-structure if it is governed by and adjacent to [−N]. [−N] elements are INFL[+tense], V and P" (cf. [Cow92, p. 102]).[61]

3.5. Passive

In this — very short — section we will show that (unlike in German or Russian) passivization in Polish does not seem amenable to an analysis in terms of structural vs. lexical case dichotomy.

First of all, note that there are verbs requiring lexical objects which can be nevertheless passivized.

(157) a. *Jan kieruje fabryką.*
John$_{nom}$ manages factory$_{ins}$.
'John manages a factory.'

b. *Fabryka jest kierowana przez Jana.*
Factory$_{nom}$ is managed by John$_{acc}$.
'A factory is managed by John.'

In this example it is the (lexical) instrumental object that gets passivized. According to the case lattice for Polish (cf. (117) on page 68) instrumental cannot be an instance of structural case. This observation is confirmed by

[61] See also [Cho86b, p. 36], [Fra90], [Fra94] and [Net94].

the inability of the object in (157a) to change its case (to genitive) under nominalization or negation (i.e., by its failure to pass the two tests of structural environment):

(158) a. *kierowanie fabryką/*fabryki*
managing factory$_{ins/gen}$
'managing a factory'

b. *Jan nie kieruje fabryką/*fabryki.*
John$_{nom}$ not manages factory$_{ins/gen}$.
'John does not manage a factory.'

Moreover, only some verbs subcategorizing for instrumental complements can be passivized:[62]

(159) a. *Jan macha chorągiewką.*
John$_{nom}$ waves banner$_{ins}$.
'John waves a banner.'

b. * *Chorągiewka jest machana przez Jana.*
Banner$_{nom}$ is waved by John$_{acc}$
'A banner is being waved by John.'

In the example above, *macha* ('waves') seems to have the same syntactic subcategorization requirements as *kieruje*, but it cannot passivize.

Contrasts such as (157) vs. (159) suggest that passivization in Polish is independent of the syntactic characterization of the SUBCAT arguments. This conclusion is further supported by the observation that, contrary to the generalization often made, not all verbs which are transitive (in the sense that their SUBCAT value is ⟨NP[*str*], NP[*str*]⟩) can be passivized. Some exceptions are given below:

(160) a. *Brzuch boli Jana.*
Stomach$_{nom}$ aches John$_{acc}$.
'John has a stomach ache.'

b. * *Jan jest bolony przez brzuch.*
John$_{nom}$ is ached by stomach$_{acc}$.
'John has a stomach ache.'

[62] In Polish, unlike in German, the passivized object always receives the nominative case. Thus the sentence:

(i) * *Chorągiewką jest machana przez Jana.*
Banner$_{ins}$ is waved by John$_{acc}$

is clearly ungrammatical.

(161) a. *Noga swędzi Jana.*
Leg$_{nom}$ itches John$_{acc}$.
'John has an itchy leg.'

b. * *Jan jest swędzony przez nogę.*
John$_{nom}$ is itched by leg$_{acc}$.
'John has an itchy leg.'

Again, this contrasts with the usual behaviour of Polish transitive verbs:

(162) a. *Jan lubi Marię.*
John$_{nom}$ likes Mary$_{acc}$.
'John likes Mary.'

b. *Maria jest lubiona przez Jana.*
Mary$_{nom}$ is liked by John$_{acc}$.
'Mary is liked by John.'

Note that this contrast does not seem to be justified by any difference in case markings of the arguments of *boleć* and *swędzić* on the one hand, and *lubić* on the other.[63]

The foregoing observations lead us to the conclusion that the phenomena of passivization is (to a large extent) independent of the syntactic case values of verb's arguments, and — as such — outwith the scope of this paper.[64]

4. Unified Account of Agreement and Case

In this final section, we will shortly describe how the principles introduced above interact and allow for a correct account of agreement feature and case assignment in Polish.

In order to focus our attention we will consider the following example:

(163) *Moich dwóch przyjaciół dało Marii dużo pięknych kwiatów.*
my$_{acc}$ two$_{acc}$ friends$_{gen}$ gave$_{3rd,neut}$ Maria$_{ins}$ lots beautiful$_{gen,nmh1}$ flowers$_{gen,nmh1}$
'My two friends gave Maria a lot of beautiful flowers.'

[63] Actually, verbs like *boleć* or *swędzić* are interesting in one more respect: they cannot be nominalized. This may be caused by the fact that the first argument of these verbs is not an agent. Thus, we could add the condition of agentivity to the prerequisites of nominalization lexical rule and, perhaps, passivization lexical rule. Such an amendment, however, would not explain the passivization behaviour of instrumental objects described above (cf. examples (157)–(159)).

[64] The reader interested in semantic account of passivization in Polish is referred to [Hol91].

We begin with the SUBCAT list of the verb *dawać* ('give'). The occurrence of the accusative subject *moich dwóch przyjaciół* suggests that the subject of *dawać* should be specified as NP[*str*]. Further, the phrase *dużo pięknych kwiatów* indicates that the accusative argument of *dawać* constitutes a structural environment (cp. 3.4.2.). Finally, taking the instrumental phrase *Marii* into account leads to the following SUBCAT list:[65]

(164) dawać: ⟨ NP[*str*], NP[*str*], NP[*lins*] ⟩

The correctness of this list is confirmed by the following example:

(165)

Piotr	*dał*	*Marii*	*kwiaty.*
Piotr_{nom}	$\text{gave}_{3rd,masc}$	Maria_{ins}	flowers_{acc}

'Piotr gave Maria flowers.'

In (165), the structural subject is saturated by the nominative phrase *Piotr*, and the structural object by the accusative phrase *kwiaty* ('flowers').

Let us consider the case assignment in phrases *moich dwóch przyjaciół* and *dużo pięknych kwiatów*. The genitive case of the NP *kwiatów* is easily explained by the syntactic requirements of *dużo* as specified in (144), p. 83. The lexical entry in (144) requires also that *dużo* be assigned structural case which squares with SUBCAT specification in (164).

The phrase *dużo pięknych kwiatów* is also affected by the Case Principle as it constitutes a structural object of a verb. Thus, its structural case is specified as *sacc* according to the Case Principle on page 88. The correct form of the adjective *pięknych* is readily explained by the attributive adjective agreement pattern in (49), p. 33, which predictes the observed case and gender concord.

The phrase *moich dwóch przyjaciół* contains the numeral *dwóch*. According to our approach, *dwóch* is a noun specified as NUMERAL + which heads the considered phrase. As can be seen from the subcat list in (164), the phrase *moich dwóch przyjaciół* occupies a position to which structural case is assigned. Thus, *dwóch* receives here structural case. Bearing structural case, it requires a genitive noun phrase as a complement (cf. (127), p. 73), which is realized by the genetive NP *przyjaciół*. The phrase is also subject to the Case Principle as a structural subject of a verb. Since its head is specified as NUMERAL +, it receives structural accusative case. The case concord between the possessive pronoun *moich* and the phrase *dwoóch przyjaciół* is predicted by the lexical entry for the pronoun *mój* as specified in (58) on page 37.

[65] Recall that order of complements on the SUBCAT list corresponds to functional hierarchy, not surface word order.

Finally, the phrase *moich dwóch przyjaciół* constitutes a nonnominative subject of the verb *dawać* in (163). According to the Verb Agreement Principle (87), p. 52, the verb assumes the 3rd person, neuter form *dało*. Thus, we provided explanation for all agreement feature and case values in (163).

As far as example (165) is concerned, the Case Principle accounts for the nominative case of the subject *Piotr* which is a noun phrase specified as NUMERAL – and for the accusative case of the structural object *kwiaty*. The Verb Agreement Principle takes care of the correct verbal form: as the subject is a nominative NP, the indexical agreement between the verb and the subject obtains, which results in the correct 3rd person, *masc* form of the verb.

References

[Bab86] Leonard H. Babby. The locus of case assignment and the direction of percolation: Case theory and Russian. In Brecht and Levine [BL86], pages 170–219.

[BF88] Michael Barlow and Charles A. Ferguson, editors. *Agreement in Natural Language: Approaches, Theories, Descriptions*, Stanford, 1988. CSLI.

[BKS71] Danuta Buttler, Halina Kurkowska, and Halina Satkiewicz. *Kultura języka polskiego: Zagadnienia poprawności gramatycznej.* Państwowe Wydawnictwo Naukowe, Warszawa, 1971.

[BL86] Richard D. Brecht and James S. Levine, editors. *Case in Slavic.* Slavica Publishers, Inc., 1986.

[Car92] Bob Carpenter. *The Logic of Typed Feature Structures.* Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 1992.

[Cho86a] Noam Chomsky. *Knowledge of Language.* Praeger, New York, 1986.

[Cho86b] Noam Chomsky. *Barriers.* MIT Press, Cambridge, Mass., 1986.

[Com86] Bernard Comrie. On delimiting cases. In Brecht and Levine [BL86], pages 86–106.

[Cor78] Greville Corbett. Numerous squishes and squishy numerals in Slavonic. In Bernard Comrie, editor, *Classification of Grammatical Categories*, pages 43–73. Linguistic Research, Inc., Edmonton, 1978.

[Cow92] Elisabeth A. Cowper. *Concise Introduction to Syntactic Theory: The Government-Binding Approach.* The University of Chicago Press, 1992.

[DG92] Walter Daelmans and Gerald Gazdar, editors. *Computational Linguistics 18(2–3): Special issue on inheritance I & II*, 1992.

[DJ89] David Dowty and Pauline Jacobson. Agreement as a semantic phenomenon. In *Proceedings of the Fifth Eastern States Conference on Linguistics*, pages 95–108, Columbus, 1989. The Ohio State University Press.

[Dor80] Witold Doroszewski, editor. *Słownik Poprawnej Polszczyzny*, Warszawa, 1980. Państwowe Wydawnictwo Naukowe.

[FB88] Charles A. Ferguson and Michael Barlow. Introduction. In Barlow and Ferguson [BF88], pages 1–22.

[Fra86] Steven Franks. Case and the structure of NP. In Brecht and Levine [BL86], pages 220–243.

[Fra90] Steven Franks. Case, configuration and argumenthood: Reflections on second dative. *Russian Linguistics*, 14:231–254, 1990.

[Fra94] Steven Franks. Parametric properties of numeral phrases in Slavic. *Natural Language and Linguistic Theory*, 1994.

[Hae91] Liliane Haegeman. *Introduction to Government and Binding Theory*. Blackwell Textbooks in Linguistics. Blackwell, Oxford, 1991.

[Hai85] Hubert Haider. The case in German. In J. Toman, editor, *Studies on German Grammar*. Foris, Dodrecht, 1985.

[HM94] Wolfgang Heinz and Johannes Matiasek. Argument structure and case assignment in German. In Nerbonne et al. [NNP94], pages 199–236.

[Hol91] Axel Holvoet. *Transitivity and Clause Structure in Polish. A Study in Case Marking*, volume 95 of *Prace Slawistyczne*. Slawistyczny Ośrodek Wydawniczy PAN, Warszawa, 1991.

[Jak71] Roman O. Jakobson. Morfologičeskie nabljudenija nad slavjanskim skloneniem. In *Selected Writings II*. Mouton, The Hague, 1958/71.

[Katng] Andreas Kathol. Agreement and the syntax-morphology interface in HPSG. volume 45 of *OSU Working Papers in Linguistics*. forthcoming.

[Kle86] Zenon Klemensiewicz. *Podstawowe wiadomości z gramatyki języka polskiego*. Państwowe Wydawnictwo Naukowe, Warszawa, thirteenth edition, 1986.

[Kur60] Jerzy Kuryłowicz. Le problème du classement des cas. *Biuletyn Polskiego Towarzystwa Językoznawczego*, 9:20–43, 1949/60.

[Mań56] W. Mańczak. Ile jest rodzajów w polskim? *Język Polski*, XXXVI(2):116–121, 1956.

[Mel86] Igor Mel'čuk. Toward a definition of case. In Brecht and Levine [BL86], pages 35–85.

[MK94] Małgorzata Marciniak and Anna Kupść. Wprowadzenie do HPSG: Teoria i praktyczne zastosowanie. Technical Report 750, Institute of Computer Science, Polish Academy of Sciences, September 1994.

[Net94] Klaus Netter. Towards a theory of functional heads. In Nerbonne et al. [NNP94], pages 297–340.

[NNP94] John Nerbonne, Klaus Netter, and Carl Pollard, editors. *German in Head-Driven Phrase Structure Grammar*, Stanford, CA, 1994. CSLI.

[PG87] Derek Proudian and David Goddeau. Constituent coordination in HPSG. Technical Report CSLI-87-97, Hewlett-Packard Laboratories and CSLI, May 1987.

[Pol89] Jean-Yves Pollock. Verb movement, universal grammar and the structure of IP. *Linguistic Inquiry*, 20:365–424, 1989.

[Pol93] Kazimierz Polański, editor. *Encyklopedia językoznawstwa ogólnego*, Wrocław, 1993. Ossolineum.

[Pol94] Carl Pollard. Toward a unified account of passive in German. In Nerbonne et al. [NNP94], pages 273–296.

[PS87] Carl Pollard and Ivan A. Sag. *Information-Based Syntax and Semantics, Volume 1: Fundamentals.* CSLI, 1987.

[PS94] Carl Pollard and Ivan A. Sag. *Head-driven Phrase Structure Grammar.* Chicago University Press, 1994.

[SŚ85] Zygmunt Saloni and Marek Świdziński. *Składnia współczesnego języka polskiego.* Państwowe Wydawnictwo Naukowe, 1985.

[Str93] Janusz Strutyński. *Zarys gramatyki polskiej: Morfologia.* Uniwersytet Jagielloński, Instytut Filologii Polskiej, Kraków, 1993.

[Świ92] Marek Świdziński. *Gramatyka formalna języka polskiego*, volume 349 of *Rozprawy Uniwersytetu Warszawskiego.* Wydawnictwa Uniwersytetu Warszawskiego, Warszawa, 1992.

[Tim86] Alan Timberlake. Hierarchies in the genitive of negation. In Brecht and Levine [BL86], pages 338–360.

[Wil90] Ewa Willim. On case-marking in Polish. *Papers and Studies in Contrastive Linguistics*, XXV:203–220, 1990.

[Zab89] Marglit Zabludowski. Quantificational noun phrases in Polish. Unpublished manuscript, September 1989.