8

A Hierarchy of Polish Genders

Adam Przepiórkowski
Polish Academy of Sciences

Contents

1   Introduction

Major generative linguistic frameworks, including Principles & Parameters (P&P; Chomsky 1981, 1995), Lexical-Functional Grammar (LFG; Bresnan 1982, 2000) and Head-driven Phrase Structure Grammar (HPSG; Pollard and Sag 1987, 1994), make crucial use of feature matrices. LFG and HPSG, which formalise their feature logics (cf., e.g., Kaplan and Bresnan 1982, Kaplan 1995 for LFG and King 1989, 1994, Carpenter 1992, Richter 2000 for HPSG), explicitly allow features to have complex values which themselves are feature matrices. HPSG requires further that such feature structures (FSs) be typed, where each type specifies the repertoire of features that a FS of that type bears, as well as types of values of those features. All types are ordered into a multiple inheritence hierarchy, i.e., a partial order, where subtypes inherit the information introduced by their (possibly multiple) supertypes, perhaps adding new information.

The aim of this paper is to show the usefulness, if not the necessity, of the formal apparatus of HPSG feature structures and types to model the morphosyntactic category of gender in Polish: I propose a multiple inheritance hierarchy for the type *gender* which reflects, in a non-reduntand way, various gender syncretisms and gender agreement patterns in Polish, including the troublesome agreement patterns in numeral phrases.

2   Features and types in HPSG

In HPSG, ϕ-features are represented as a Feature Structure (FS) of type *index*. All FSs of type *index* have exactly three attributes, GENDER, NUMBER and PERSON, with values being FSs of type *gender*, *number* and *person*, respectively. This information is represented graphically in (1).

(1) ϕ-features:

$$\begin{bmatrix} index \\ \text{GENDER } gender \\ \text{NUMBER } number \\ \text{PERSON } person \end{bmatrix}$$

More generally, all linguistic objects, be it ϕ-features, cases or utterances, are represented as FSs of specific types. To give another example, all linguistic signs (words, phrases and sentences) are represented by FSs of type *sign*, whose two features, PHON and SYNSEM, correspond to the sign's form and its syntactico-semantic content, (2).

(2) $$\begin{bmatrix} sign \\ \text{PHON } phon \\ \text{SYNSEM } synsem \end{bmatrix}$$

In the following two subsections, I will briefly introduce HPSG type hierarchies, crucial of the account of Polish gender presented in the remainder of this paper.

2.1   Type hierarchies

HPSG types are ordered into an inheritance hierarchy, i.e., a partial order, where subtypes inherit the information introduced by their supertypes, perhaps adding new information.

For example, the type *sign* has two subtypes, *word* and *phrase*, the latter introducing two additional attributes, whose values represent the tree-configurational relation of immediate dominance:

(3) $$\begin{bmatrix} sign \\ \text{PHON } phon \\ \text{SYNSEM } synsem \end{bmatrix}$$

word $$\begin{bmatrix} phrase \\ \text{HD-DTR } sign \\ \text{NONHD-DTRS } list(sign) \end{bmatrix}$$

Since *phrase* inherits the information specific for its supertype, *sign*, each FS of type *phrase* is specified to have exactly four attributes:

(4)
$$\begin{bmatrix} phrase \\ \text{PHON } phon \\ \text{SYNSEM } synsem \\ \text{HD-DTR } sign \\ \text{NONHD-DTRS } list(sign) \end{bmatrix}$$

Two more trivial pieces of the HPSG type hierarchy concern the values of the attributes NUMBER and PERSON in (1) above.

(5)   *number*
           ⁀
        sg   *pl*

(6)   *person*

      1st   *2nd*   *3rd*

According to (5)–(6), the type *number* has two subtypes, *sg* and *pl*, while the type *person* has three subtypes, *1st*, *2nd* and *3rd*. Since a feature structure of a certain type must be of exactly one maximally specific subtype of this type, each feature structure of type *number* is either of type *sg*, or of type *pl*, so the value of the attribute GENDER is a FS of exactly one of these two maximally specific types. FSs of these types (as well as FSs of type *number* and its subtypes) are atomic: they are trivial FSs which do not actually have any features.

2.2 Multiple inheritance

The HPSG type inheritance is a multiple type inheritance: a type may be an immediate subtype of two or more different types. In order to give a linguistically justified example, let's first look closer at the type *synsem*, the type of values of the attribute SYNSEM introduced in (3) above.[1]

(7)
$$\begin{bmatrix} synsem \\ \text{CAT } head \\ \text{CONT } content \\ \text{CONX } context \end{bmatrix}$$

The three attributes of *synsem* represent the sign's morphosyntactic category, its semantics and its pragmatics, respectively. The values of the CAT attribute are largely language-specific; for Polish, we may assume the existence of *nominal* elements, which have CASE, and *verbal* elements, which have ASPECT:

---

[1]The type hierarchy presented here is slightly simplified with respect to the standard HPSG type hierarchy.

(8)

$$\text{head}$$

$$\begin{bmatrix} \textit{nominal} \\ \text{CASE } \textit{case} \\ \dots \end{bmatrix} \qquad \begin{bmatrix} \textit{verbal} \\ \text{ASPECT } \textit{aspect} \\ \dots \end{bmatrix}$$
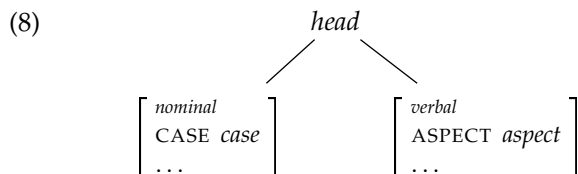
There are, however, morphosyntactic categories which are 'mixed' in the sense of having both the nominal and the verbal properties, e.g., in Polish, the *-nie/-cie* verbal nouns, or gerunds. In HPSG they are modelled as subtypes of both *nominal* and *verbal*, inheriting all properties of both types, i.e., simultaneously specified for CASE and ASPECT:

(9)

$$\text{head}$$

$$\begin{bmatrix} \textit{nominal} \\ \text{CASE } \textit{case} \\ \dots \end{bmatrix} \qquad \begin{bmatrix} \textit{verbal} \\ \text{ASPECT } \textit{aspect} \\ \dots \end{bmatrix}$$

$$\dots \qquad \textit{gerund} \qquad \dots$$
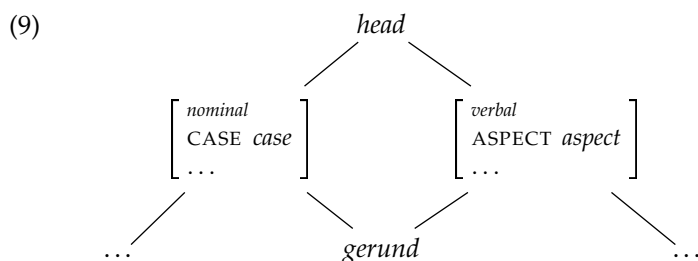
The main theoretical thesis of this paper is that all this machinery made available in HPSG is very useful, if not necessary, to provide a satisfactory account of Polish gender.

3   Genders in Polish

Let us assume the repertoire of five genders in Polish argued for by Mańczak 1956: virile (*m1*; e.g., *mężczyzna* 'man', *student* 'student'), so-called masculine-animate (*m2*; e.g., *koń* 'horse', *babsztyl* 'broad, woman', *walc* 'waltz'), masculine-inanimate (*m3*; e.g., *stół* 'table'), neuter (*neut*; e.g., *okno* 'window', *dziecko* 'child', *dziewczę* 'girl') and feminine (*fem*; e.g., *kobieta* 'woman', *ciota* 'faggot', *szafa* 'wardrobe'). The full paradigm of the adjective *fajny* 'nice, cool' is presented in the table in (10) on the following page.

Note that the distinction between, on the one hand, *m1* and *m2*, and, on the other hand, *m3* is needed in order to, e.g., distinguish between singular accusative forms of adjectives modifying nouns such as *mężczyzna* 'man', *babsztyl* 'broad, woman' or *walc* 'waltz' on the one hand, and *okno* 'window' or *dziecko* 'child' on the other hand. Similarly, the distinction between *m1* on the one hand, and *m2* and *m3* on the other surfaces in the plural nominative and accusative. Note also that, given the above examples of nouns in the three masculine genders, it is incorrect to claim that the distinction between *m1*, *m2* and *m3* is just a semantic animacy distinction (although some correlation between the three masculine genders and semantic animacy is clear).

(10)

		m1	m2	m3	neut	fem
sg	**nom**	fajny	fajny	fajny	fajne	fajna
	acc	fajnego	fajnego	fajny	fajne	fajną
	gen	fajnego	fajnego	fajnego	fajnego	fajnej
	dat	fajnemu	fajnemu	fajnemu	fajnemu	fajnej
	ins	fajnym	fajnym	fajnym	fajnym	fajną
	loc	fajnym	fajnym	fajnym	fajnym	fajnej
pl	**nom**	fajni	fajne	fajne	fajne	fajne
	acc	fajnych	fajne	fajne	fajne	fajne
	gen	fajnych	fajnych	fajnych	fajnych	fajnych
	dat	fajnym	fajnym	fajnym	fajnym	fajnym
	ins	fajnymi	fajnymi	fajnymi	fajnymi	fajnymi
	loc	fajnych	fajnych	fajnych	fajnych	fajnych

4   Polish gender in HPSG

Given the 5 genders in Polish, we could adopt the following type hierarchy for the type *gender*, similar to the hierarchies (5) and (6) above:

(11)   *gender*

          m1   *m2*   *m3*   *neut*   *fem*
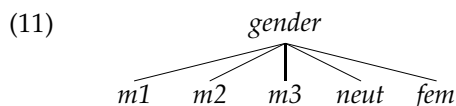
However, this *gender* hierarchy does not take into account various gender syncretisms in Polish, cf. §4.1 below, nor does it allow for an easy account of the intriguing gender agreement facts in Polish numeral phrases, discussed in §4.2.

4.1   Gender syncretisms

The table (10) illustrates a large number of systematic syncretisms in the adjectival paradigm, including the following gender syncretisms:

- for all genders: the same plural forms in gen, dat, ins, loc;
- for *m1*, *m2*, *m3*, *neut*: the same singular forms in gen, dat, ins, loc;
- for *m1*, *m2*, *m3*: the same singular nom forms;
- *m1* and *m2*: the same singular acc forms;
- *m2*, *m3*, *neut*, *fem*: the same plural forms in nom and acc.

In order to minimise the redundancy in the lexicon of forms, those systematic syncretisms should be reified in the grammar. In HPSG, an obvious way of doing so is to postulate a more structured type hierarchy for *gender*, with new types corresponding to the sets of syncretic genders mentioned above. The simplest such type hierarchy is presented below.[2]

---

[2]Note that a type corresponding to the first bullet above, i.e., the type *gender*, encompassing all gender values, is already present in the original type hierarchy.

(12)
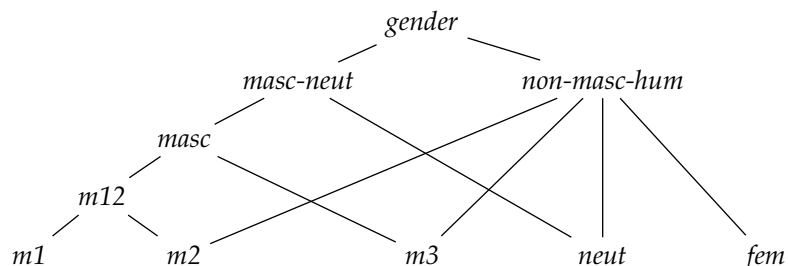
Note that the types *masc* and *non-masc-hum* are gender values not only of systematically syncretic adjectival forms, but also of some systematically syncretic verbal forms: *masc* is the gender value of past verbal forms in the singular such as *przyszedł* 'came', *non-masc-hum* is the type of past verbal forms in the plural such as *przyszły* 'came'.

In some versions of HPSG, e.g., as formalised in Carpenter 1992, there is an additional technical requirement that type hierarchies be *bounded complete*, i.e., for every pair of types which have a common subtype, there must be a unique most general common subtype of these two types. In the hierarchy in (12), this condition is not satisfied, e.g., for the types *masc-neut* and *non-masc-hum*: they have common subtypes *m2*, *m3* and *neut*, and yet they lack a unique common subtype being a supertype of these three types. The hierarchy presented below is a minimal modification of (12) satisfying the requirement of bounded completeness. For example, the types *masc-neut* and *non-masc-hum* have a number of common subtypes (*m2*, *m3*, *neut*, *m23* and *ms23neut*), but one of them, *m23neut*, is most general: it is a supertype of all the other common subtypes of *masc-neut* and *non-masc-hum*.

(13)
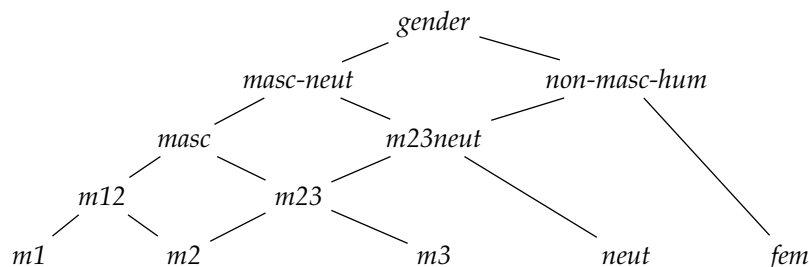
The two new types, *m23* and *m23neut*, although introduced here for technical reasons, turn out to have a certain empirical and theoretical importance which will be discussed in the next subsection.

4.2   Gender agreement in Polish Numeral Phrases

Let us consider gender agreement in phrases involving the numeral *dwa* 'two':

(14)   a.   *dwaj* mężczyźni stali
              two men$_{M1}$ stood
        b.   *dwa* konie   / stoły   / okna   stały
              two horses$_{M2}$ / tables$_{M3}$ / windows$_{NEUT}$ stood

    c.   *dwie* kobiety   stały
        two   women_FEM stood

Given that different forms of the lexeme *dwa* combine with nouns of different gender, it makes sense to say that the lexeme *dwa* inflects for gender. The form *dwa*, as in (14b), combines with singular *m2*, *m3* and *neut* nouns, so — in the lexicon — the form *dwa* should be specified as having the 'technical' gender value *m23neut*.

As noted by Saloni 1976, gender agreement in such numeral phrases is a more complex matter than the example above would suggest. Consider (15) below:

(15)   a.   *dwoje / *dwa* dzieci   / kurcząt   / oczu   / studentów
         two   children_NEUT / chicks_NEUT / eyes_NEUT / students_NEUT
    b.   *dwa / *dwoje* okna   / pola   / euro
         two   windows_NEUT / fields_NEUT / euro_NEUT

On the basis of such facts, Saloni 1976 proposes to split the neuter gender into *n1*, for *dwoje* and nouns such as those in (15a), and *n2*, for *dwa* and nouns in (15b).

Moreover, Saloni 1976 introduces three separate *plurale tantum* genders, for nouns which do not occur in the singular form. First, *plurale tantum* nouns which occur with *m1* verb forms, as in (16a) below, are distinguished from *plurale tantum* nouns occurring with *non-masc-hum* verb forms, as in (16b).

(16)   a.   wujostwo   / państwo / rodzice   stali
         uncle and aunt_P1 / people_P1 / parents_P1 stood
    b.   skrzypce / drzwi   / okulary   / urodziny   stały
         violin_P-1 / door_P-1 / glasses_P-1 / birthday_P-1 stood

Second, the latter class of nouns is further split into those nouns that can be modified by a numeral, as in (17a), and those that cannot, as in (17b). The three new genders are called *p1*, *p2* and *p3*, as indicated in (16a) and (17).

(17)   a.   dwoje skrzypiec / drzwi
         two   violins_P2 / doors_P2
    b.   *dwoje okularów / urodzin
         two   glasses_P3 / birthdays_P3

The need for these three *plurale tantum* genders has recently been questioned by, independently, Przepiórkowski *et al.* 2002 and Woliński 2001. The arguments presented below follow Przepiórkowski *et al.* 2002.

First of all, it turns out that there is no reason to distinguish between *m1* and *p1*. Saloni 1976 makes such a distinction solely on the basis of the fact that *m1* nouns have singular forms, while *p1* nouns — do not. Although Woliński 2001 and Przepiórkowski *et al.* 2002 argue that the fact that some lexemes lack singular forms is not a sufficient criterion for delimiting genders, Woliński 2001 still maintains the distinction between *m1* and *p1*, namely, on the basis of the observation that *p1* nouns combine with collective numeral forms, while *m1* nouns combine with ordinary nu-

meral forms:[3]

(18) a. dwoje wujostwa   / państwa / rodziców
           two   uncle and aunt$_{P1}$ / people$_{P1}$ / parents$_{P1}$
      b. *dwoje chłopców / mężczyzn / facetów
           two   boys$_{M1}$ / men$_{M1}$ / guys$_{M1}$
      c. dwaj chłopcy / mężczyźni / faceci   / studenci   / uczniowie...
           two   boys$_{M1}$ / men$_{M1}$ / guys$_{M1}$ / students$_{M1}$ / pupils$_{M1}$

Note, however, that also nouns that are conspicuously *m1* in the sense of Saloni 1976
and Woliński 2001 can be combined with collective numerals, namely, when they refer
to groups of people of mixed sex:

(19) dwoje studentów / uczniów / adeptów
      two   students / pupils / adepts

Since the *p1* nouns in (18a) refer to groups of mixed sex by virtue of their lexical
meanings, any account of (19), involving uncontroversial *m1* nouns, will also account
for (18a) with the relevant nouns analysed as having the *m1* gender. I conclude that
the contrast in (18) is not a sufficient reason to postulate a separate *plurale tantum* hu-
man masculine gender such as *p1*.

Second, once the fact that some nouns only have plural forms is deemed irrelevant
for the delimitation of genders, there is no argument for postulating separate *p2* and
p3 genders, either: even though the specific gender of such *plurale tantum* nouns can-
not be empirically determined above establishing that they are not *m1* (relevant tests
involve singular forms), it does not follow that such nouns have a gender different
from those already proposed by Mańczak 1956.

Finally, given that the way nouns combine with collective or ordinary numeral
forms is partially governed by pragmatics, i.e., whether a given plural noun such as
studenci 'students' refers in a given context to a group of mixed or same sex, it is not
clear that the connectivity with numerals should be used as a test for delimiting the
members of the morphosyntactic category of gender at all.

In what follows, I present an HPSG analysis of Polish gender which accounts for
the facts discussed by Saloni 1976 and Woliński 2001, which does not, however, pro-
liferate the number of genders beyond those proposed by Mańczak 1956, and which
takes into account the partially pragmatic nature of connectivity with collective nu-
merals.

4.3 Gender and numerals in HPSG

In order to account for the correlation between gender and connectivity with numer-
als, let us introduce an attribute to the now no longer atomic type GENDER:

---

[3]Note that collective numeral forms such as *dwoje* in (18a) always occur with genitive nouns, while forms
such as *dwaj*, *dwa* and *dwie* in (14)–(15) above and in (18c) below agree in case with the nouns they modify.
Much more can be said about the complexity of Polish numeral system; cf., e.g., Przepiórkowski 1999 and
references therein, as well as Rappaport's article in this volume.

(20) $\begin{bmatrix} gender \\ \text{NUM } num \end{bmatrix}$

The values of NUM are of type *num*, which has three subtypes: *main*, for nouns which may combine with ordinary numerals, *coll*ective, for nouns which may combine with collective numerals, and *non-accom*modating, for nouns which do not combine with numerals at all (cf. Saloni's gender *p3* discussed above):

(21)   *num*

main   *coll*   *non-accom*

Considering ordinary (non-*plurale tantum*) nouns first, we observe that nouns in *m2*, *m3* and *fem* always combine with main numerals, i.e., for those genders the value of NUM is always *main*. This is represented in the following modified gender hierarchy:

(22)

$\begin{bmatrix} gender \\ \text{NUM } num \end{bmatrix}$

masc-neut   *non-masc-hum*

masc   *m23neut*

m12 $\begin{bmatrix} m23 \\ \text{NUM } main \end{bmatrix}$

m1   *m2*   *m3*   *neut* $\begin{bmatrix} fem \\ \text{NUM } main \end{bmatrix}$

In case of *m2*, *m3* and *fem* nouns, their lexical entries only need to contain the information about their basic gender; the value of their NUM attribute will follow from the hierarchy above and will always be *main*. The situation is different in case of neuter nouns. As illustrated in (15) above, some such nouns may occur with collective numerals such as *dwoje*, and other neuter nouns may only occur with ordinary numeral forms, such as *dwa*. For our purposes, it makes sense to assume that this information is given in the lexicon, so that, e.g., gender specifications relevant neuter nouns look as follows:

(23)   a.   *okno* 'window': $\begin{bmatrix} neut \\ \text{NUM } main \end{bmatrix}$

        b.   *dziecko* 'child': $\begin{bmatrix} neut \\ \text{NUM } coll \end{bmatrix}$

Finally, in case of *m1* nouns, we have noted that their connectivity with numerals is governed by their pragmatics, i.e., roughly, whether they refer in a given context to a group of mixed or same sex. In HPSG the relevant principle may be schematically

formalised as follows:

$$(24) \quad \begin{bmatrix} \text{CAT} & \begin{bmatrix} noun \\ \text{GENDER} & \begin{bmatrix} m1 \\ \text{NUM} & \boxed{0} \end{bmatrix} \end{bmatrix} \\ \text{CONT} & \boxed{1} \\ \text{CONX} & \boxed{2} \end{bmatrix} \rightarrow \begin{array}{l} (\texttt{refers-to-mixed-sex}(\boxed{1},\boxed{2}) \rightarrow \boxed{0} = coll \wedge \\ \texttt{refers-to-same-sex}(\boxed{1},\boxed{2}) \rightarrow \boxed{0} = main) \end{array}$$

In order to fully state this principle, relations `refers-to-mixed-sex` and `refers-to-same-sex` would have to be formally defined and possible values of the semantic attribute CONT and the pragmatic attribute CONX would have to be carefully introduced, which is beyond the scope of this article. Suffice it to say that HPSG provides mechanisms making it possible to define such relations (see Richter 2000) and formalise a theory of semantics and pragmatics (see, e.g., Pollard and Sag 1994, Pollard and Yoo 1998, Przepiórkowski 1998, Richter and Sailer 1997, Vallduví and Engdahl 1996, Green 2000, Ginzburg and Sag 2000).

With such a principle in the grammar, lexical entries of *m1* nouns do not need to contain information about their NUM values: when such nouns refer, in a given context, to a mixed sex entity, the value of their NUM is *coll*, by virtue of the principle above; otherwise it is *main*. Of course, the lexical semantics of the vast majority of *m1* nouns is incompatible with reference to mixed sex, and such nouns will always combine with *main* numerals only.

Before discussing the place of *plurale tantum* nouns in this setup, let us consider gender agreement in numeral phrases. I assume the following lexically specified values of GENDER for different forms of the numeral *dwa* 'two':

$$(25) \quad \text{a.} \quad dwoje: \begin{bmatrix} gender \\ \text{NUM} & coll \end{bmatrix}$$

$$\text{b.} \quad dwaj: \begin{bmatrix} m1 \\ \text{NUM} & main \end{bmatrix}$$

$$\text{c.} \quad dwa: \begin{bmatrix} m23neut \\ \text{NUM} & main \end{bmatrix}$$

$$\text{d.} \quad dwie: \begin{bmatrix} fem \\ \text{NUM} & main \end{bmatrix}$$

In case of the phrase *dwie kobiety* 'two women', lexically specified GENDER values of *dwie*, $\begin{bmatrix} fem \\ \text{NUM} & main \end{bmatrix}$, and *kobiety*, *fem*, are compatible, and the combined value is $\begin{bmatrix} fem \\ \text{NUM} & main \end{bmatrix}$, which also satisfies the requirement in (22) that the NUM value of *fem* nouns be *main*.

$$(26) \quad \textit{dwie kobiety} \text{ 'two women'}: \begin{bmatrix} fem \\ \text{NUM} & main \end{bmatrix} + fem = \begin{bmatrix} fem \\ \text{NUM} & main \end{bmatrix}$$

The feminine noun *kobiety* cannot be modified by any other numeral form of *dwa*: e.g., in case of *dwaj*, the type *m1* is incompatible with the type *fem* (those two types do not

have any common subtypes), while in case of *dwoje*, the type of the NUM value of the numeral, *coll*, is incompatible with the type of the NUM value of a feminine noun as specified in (22), *main*:

(27)   a.   *dwaj kobiety*: $\begin{bmatrix} m1 \\ \text{NUM } main \end{bmatrix} + fem = {}^*$

       b.   *dwoje kobiet*: $\begin{bmatrix} gender \\ \text{NUM } coll \end{bmatrix} + fem = \begin{bmatrix} fem \\ \text{NUM } coll \end{bmatrix} = {}^*$ (cf. (22))

Similarly, for the phrase *dwa konie* 'two horses', gender agreement consists in the simultaneous satisfaction of both lexical requirements, i.e., $\begin{bmatrix} m23neut \\ \text{NUM } main \end{bmatrix}$ for *dwa* and *m2* for *konie*:

(28)   *dwa konie* 'two horses': $\begin{bmatrix} m23neut \\ \text{NUM } main \end{bmatrix} + m2 = \begin{bmatrix} m2 \\ \text{NUM } main \end{bmatrix}$

Examples of successful and failed gender agreement involving neuter nouns are given below.

(29)   a.   *dwa okna* 'two windows': $\begin{bmatrix} m23neut \\ \text{NUM } main \end{bmatrix} + \begin{bmatrix} neut \\ \text{NUM } main \end{bmatrix} = \begin{bmatrix} neut \\ \text{NUM } main \end{bmatrix}$

       b.   *dwa dzieci* 'two children': $\begin{bmatrix} m23neut \\ \text{NUM } main \end{bmatrix} + \begin{bmatrix} neut \\ \text{NUM } coll \end{bmatrix} = {}^*$

       c.   *dwoje okien* 'two windows': $\begin{bmatrix} gender \\ \text{NUM } coll \end{bmatrix} + \begin{bmatrix} neut \\ \text{NUM } main \end{bmatrix} = {}^*$

       d.   *dwoje dzieci* 'two children': $\begin{bmatrix} gender \\ \text{NUM } coll \end{bmatrix} + \begin{bmatrix} neut \\ \text{NUM } coll \end{bmatrix} = \begin{bmatrix} neut \\ \text{NUM } coll \end{bmatrix}$

Nouns lexically specified as having the *m1* gender may, in principle, combine with two different forms:

(30)   a.   *dwaj studenci* 'two students: $\begin{bmatrix} m1 \\ \text{NUM } main \end{bmatrix} + m1 = \begin{bmatrix} m1 \\ \text{NUM } main \end{bmatrix}$

       b.   *dwa studenci* 'two students: $\begin{bmatrix} m23neut \\ \text{NUM } main \end{bmatrix} + m1 = {}^*$

       c.   *dwie studenci* 'two students: $\begin{bmatrix} fem \\ \text{NUM } main \end{bmatrix} + m1 = {}^*$

       d.   *dwoje studentów* 'two students: $\begin{bmatrix} gender \\ \text{NUM } coll \end{bmatrix} + m1 = \begin{bmatrix} m1 \\ \text{NUM } coll \end{bmatrix}$

The numeral appropriate in a given context will be determined by principle (24).

  How do *plurale tantum* nouns fit in this setup? First of all, Saloni's *p1* nouns such as *wujostwo* or *rodzice* in (16a) above, agree with unambiguously *m1* verbal forms such as *stali*, so those nouns must in fact be *m1*. As discussed above, the fact that they combine with collective numerals follows from principle (24).

  Second, *p2* and *p3* nouns agree with *non-masc-hum* verbs, as in (16b) above. Further, *p2* nouns occur with collective numeral forms, so their NUM value is *coll*, while *p3*

nouns cannot be combined with any numerals, so their NUM value is *non-accom*:

(31)   a.   so called *p2* nouns: $\begin{bmatrix} \textit{non-masc-hum} \\ \text{NUM } \textit{coll} \end{bmatrix}$

       b.   so called *p3* nouns: $\begin{bmatrix} \textit{non-masc-hum} \\ \text{NUM } \textit{non-accom} \end{bmatrix}$

Now, according to the gender type hierarchy in (22), the only *non-masc-hum* gender whose NUM value may be *coll* or *non-accom* is *neut*: in case of the other *non-masc-hum* genders, i.e., *m2*, *m3* and *fem*, their NUM value is *main*. That means that, in fact, *p2* and *p3* genders are neuter genders differing in NUM values:

(32)   a.   *p2*: $\begin{bmatrix} \textit{neut} \\ \text{NUM } \textit{coll} \end{bmatrix}$

       b.   *p3*: $\begin{bmatrix} \textit{neut} \\ \text{NUM } \textit{non-accom} \end{bmatrix}$

I consider this to be an intriguing theorem of the HPSG approach to Polish gender presented above.

5   Conclusion

I hope to have provided a satisfactory account of Polish gender, encompassing all advantages of the structuralist accounts of Mańczak 1956, Saloni 1976 and Woliński 2001, but avoiding various pitfalls and increasing the perspicuity and compactness of the resulting analysis. In doing that I have been trying to demonstrate the usefulness, if not the necessity, of the formal and theoretical machinery provided by HPSG:

- Multiple inheritance hierarchy has been crucial for handling systematic syncretisms of adjectival and verbal forms: instead of multiplying lexical entries for phonetically and orthographically identical forms differing only in their gender values, various sets of genders are linguistically reified as types in the HPSG signature.

- The fact that all feature values are (perhaps atomic, i.e., feature-less) typed feature structures themselves made it easy to reconsider values of GENDER as nonatomic FSs, with the attribute NUM responsible for the connectivity with numerals. This way we were able to dissociate gender from number, and to provide a clean representation of interdependence between gender and numeral connectivity.

- As HPSG is a theory encompassing all levels of grammatical representation, it makes it possible to state dependencies between morphosyntax, on one hand, and semantics and pragmatics, on the other, in a perspicuous way. This advantage is illustrated by principle (24).

References

Bresnan, J., editor (1982). *The Mental Representation of Grammatical Relations*. MIT Press Series on Cognitive Theory and Mental Representation. The MIT Press, Cambridge, MA.

Bresnan, J. (2000). *Lexical-Functional Syntax*. Blackwell Textbooks in Linguistics. Blackwell.

Carpenter, B. (1992). *The Logic of Typed Feature Structures*. Cambridge University Press, Cambridge.

Chomsky, N. (1981). *Lectures on Government and Binding*. Foris, Dordrecht.

Chomsky, N. (1995). *The Minimalist Program*, volume 28 of *Current Studies in Linguistics*. The MIT Press, Cambridge, MA.

Czuba, K. (1997). Towards minimal linguistic description: A hierarchical gender structure for Polish. In U. Junghanns and G. Zybatow, editors, *Formale Slavistik*, volume 7 of *Leipziger Schriften zur Kultur-, Literatur-, Sprach- und Übersetzungswissenschaft*, pages 347–356, Frankfurt am Main. Vervuert Verlag.

Czuba, K. and Przepiórkowski, A. (1995). Agreement and case assignment in Polish: An attempt at a unified account. IPI PAN Research Report 783, Institute of Computer Science, Polish Academy of Sciences.

Ginzburg, J. and Sag, I. A. (2000). *Interrogative Investigations: The Form, Meaning and Use of English Interrogatives*. CSLI Publications, Stanford, CA.

Green, G. M. (2000). The nature of pragmatic information. In R. Cann, C. Grover, and P. Miller, editors, *Grammatical Interfaces in HPSG*, pages 113–135. CSLI Publications, Stanford, CA.

Kaplan, R. M. (1995). The formal architecture of Lexical-Functional Grammar. In M. Dalrymple, R. M. Kaplan, J. T. Maxwell III, and A. Zaenen, editors, *Formal Issues in Lexical-Functional Grammar*, number 47 in CSLI Lecture Notes, pages 7–27. CSLI Publications, Stanford, CA.

Kaplan, R. M. and Bresnan, J. (1982). Lexical-Functional Grammar: A formal system for grammatical representation. In Bresnan (1982), pages 173–281.

King, P. J. (1989). *A Logical Formalism for Head-driven Phrase Structure Grammar*. Ph. D. dissertation, University of Manchester.

King, P. J. (1994). An expanded logical formalism for Head-driven Phrase Structure Grammar. Arbeitspapiere des Sonderforschungsbereichs 340 Bericht Nr. 59, Seminar für Sprachwissenschaft, Universität Tübingen.

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

Pollard, C. and Sag, I. A. (1987). *Information-Based Syntax and Semantics, Volume 1: Fundamentals*. Number 13 in CSLI Lecture Notes. CSLI Publications, Stanford, CA.

Pollard, C. and Sag, I. A. (1994). *Head-driven Phrase Structure Grammar*. Chicago University Press / CSLI Publications, Chicago, IL.

Pollard, C. and Yoo, E. J. (1998). A unified theory of scope for quantifiers and *wh*-phrases. *Journal of Linguistics*, **34**, 415–445.

Przepiórkowski, A. (1998). 'A Unified Theory of Scope' revisited: Quantifier retrieval without spurious ambiguities. In G. Bouma, G.-J. M. Kruijff, and R. T. Oehrle, edi-

tors, *Proceedings of the Joint Conference on Formal Grammar, Head-Driven Phrase Structure Grammar, and Categorial Grammar*, pages 185–195, Saarbrücken.

Przepiórkowski, A. (1999). *Case Assignment and the Complement-Adjunct Dichotomy: A Non-Configurational Constraint-Based Approach*. Ph. D. dissertation, Universität Tübingen, Germany.

Przepiórkowski, A., Kupść, A., Marciniak, M., and Mykowiecka, A. (2002). *Formalny opis języka polskiego: Teoria i implementacja*. Akademicka Oficyna Wydawnicza EXIT, Warsaw.

Richter, F. (2000). *A Mathematical Formalism for Linguistic Theories with an Application in Head-Driven Phrase Structure Grammar*. Ph. D. dissertation, Universität Tübingen.

Richter, F. and Sailer, M. (1997). Underspecified semantics in HPSG. In H. Bunt, L. Kievit, R. Muskens, and M. Verlinden, editors, *Proceedings of the* Second International Workshop on Computational Semantics, 8–10 January 1997, Tilburg, pages 234–246.

Saloni, Z. (1976). Kategoria rodzaju we współczesnym języku polskim. In *Kategorie gramatyczne grup imiennych we współczesnym języku polskim*, pages 41–75. Ossolineum, Wrocław.

Vallduví, E. and Engdahl, E. (1996). The linguistic realization of information packaging. *Linguistics*, **34**, 459–519.

Woliński, M. (2001). Rodzajów w polszczyźnie jest osiem. In W. Gruszczyński, U. Andrejewicz, M. Bańko, and D. Kopcińska, editors, *Nie bez znaczenia... Prace ofiarowane Profesorowi Zygmuntowi Saloniemu z okazji jubileuszu 15000 dni pracy naukowej*, pages 303–305. Wydawnictwo Uniwersytetu Białostockiego, Białystok.

Adam Przepiórkowski
Institute of Computer Science
Polish Academy of Sciences
ul. Ordona 21
01-237 Warszawa
Poland
adamp@ipipan.waw.pl