8

A Hierarchy of Polish Genders

Adam Przepiórkowski
Polish Academy of Sciences

Contents

1 Introduction ........................................... 109
2 Features and types in HPSG .............................. 110
  2.1 Type hierarchies .................................. 110
  2.2 Multiple inheritance ............................... 111
3 Genders in Polish ...................................... 112
4 Polish gender in HPSG ................................. 113
  4.1 Gender syncretisms ............................... 113
  4.2 Gender agreement in Polish Numeral Phrases ...... 114
4.3 Gender and numerals in HPSG ..................... 116
5 Conclusion ............................................ 120

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 inheritance hierarchy, i.e., a partial order, where subtypes inherit the information introduced by their (possibly multiple) supertypes, perhaps adding new information.

This article presents a slight revision of the approach to Polish gender put forward in Przepiórkowski et al. 2002. It differs from Czuba and Przepiórkowski 1995, Czuba 1997, an earlier analysis of Polish gender, in almost every respect apart from the general proposal to treat various gender syncretisms via a complex inheritance hierarchy. I am grateful for comments to the audiences at GLiP-5 and at the Seminar für Sprachwissenschaft, Universität Tübingen.
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-redundant 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, $\phi$-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) \[
\begin{array}{c}
\text{index} \\
\text{GENDER gender} \\
\text{NUMBER number} \\
\text{PERSON person}
\end{array}
\]

More generally, all linguistic objects, be it $\phi$-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) \[
\begin{array}{c}
\text{sign} \\
\text{PHON phon} \\
\text{SYNSEM synsem}
\end{array}
\]

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) \[
\begin{array}{c}
\text{sign} \\
\text{PHON phon} \\
\text{SYNSEM synsem}
\end{array}
\]

\[
\begin{array}{c}
\text{word} \\
\text{phrase} \\
\text{ HD-DTR sign} \\
\text{ NONHD-DTRS last(sign)}
\end{array}
\]
Since *phrase* inherits the information specific for its supertype, *sign*, each FS of type *phrase* is specified to have exactly four attributes:

\[
\begin{array}{c}
\text{phrase} \\
\text{PHON} \ phantom \\
\text{SYNSEM} \ synsem \\
\text{HD-DTR} \ sign \\
\text{NONHD-DTRS list(sign)}
\end{array}
\]

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

\[
\begin{array}{c}
\text{number} \\
\text{sg} \quad \text{pl}
\end{array}
\]

\[
\begin{array}{c}
\text{person} \\
1st \quad 2nd \quad 3rd
\end{array}
\]

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.\(^1\)

\[
\begin{array}{c}
\text{synsem} \\
\text{CAT} \ head \\
\text{CONT} \ content \\
\text{CONX} \ context
\end{array}
\]

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.
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:

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 \((m_1); \text{e.g., }\) mężczyzna ‘man’, student ‘student’), so-called masculine-animate \((m_2); \text{e.g., }\) koń ‘horse’, babsztyl ‘broad, woman’, walc ‘waltz’), masculine-inanimate \((m_3); \text{e.g., }\) stół ‘table’), neuter \text{ (neut; \text{e.g., }\) okno ‘window’, dziecko ‘child’, dziewczę ‘girl’)} and feminine \((fem; \text{e.g., }\) 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, \(m_1\) and \(m_2\), and, on the other hand, \(m_3\) 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 \(m_1\) on the one hand, and \(m_2\) and \(m_3\) 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 \(m_1\), \(m_2\) and \(m_3\) is just a semantic animacy distinction (although some correlation between the three masculine genders and semantic animacy is clear).
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:

\[
\begin{array}{c|c|c|c|c}
\text{sg} & \text{m1} & \text{m2} & \text{m3} & \text{neut} & \text{fem} \\
\text{nom} & \text{fajny} & \text{fajny} & \text{fajny} & \text{fajne} & \text{fajna} \\
\text{acc} & \text{fajnego} & \text{fajnego} & \text{fajnego} & \text{fajne} & \text{fajną} \\
\text{gen} & \text{fajnego} & \text{fajnego} & \text{fajnego} & \text{fajnego} & \text{fajnej} \\
\text{dat} & \text{fajnemu} & \text{fajnemu} & \text{fajnemu} & \text{fajnemu} & \text{fajnej} \\
\text{ins} & \text{fajnym} & \text{fajnym} & \text{fajnym} & \text{fajnym} & \text{fajną} \\
\text{loc} & \text{fajnym} & \text{fajnym} & \text{fajnym} & \text{fajnym} & \text{fajnej} \\
\text{pl} & \text{fajni} & \text{fajne} & \text{fajne} & \text{fajne} & \text{fajne} \\
\text{nom} & \text{fajnych} & \text{fajne} & \text{fajne} & \text{fajne} & \text{fajne} \\
\text{acc} & \text{fajnych} & \text{fajnych} & \text{fajnych} & \text{fajnych} & \text{fajnych} \\
\text{gen} & \text{fajnych} & \text{fajnych} & \text{fajnych} & \text{fajnych} & \text{fajnych} \\
\text{dat} & \text{fajnym} & \text{fajnym} & \text{fajnym} & \text{fajnym} & \text{fajnym} \\
\text{ins} & \text{fajnimi} & \text{fajnimi} & \text{fajnimi} & \text{fajnimi} & \text{fajnimi} \\
\text{loc} & \text{fajnich} & \text{fajnych} & \text{fajnych} & \text{fajnych} & \text{fajnych} \\
\end{array}
\]

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.
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 *przyszed* 'came', *non-masc-hum* is the type of past verbal forms in the plural such as *przyszy* '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* and *neut*), 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*.

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

   two men*$_{M1}$* st*ood

b. *dwa konie / stoli / okna / stali*

   two horses*$_{M2}$* / tables*$_{M3}$* / windows*$_{NEUT}$* stood

   two horses*$_{M2}$* / tables*$_{M3}$* / windows*$_{NEUT}$* 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 / kurczat / oczu / studentów
two children<sub>NEUT</sub> / chicks<sub>NEUT</sub> / eyes<sub>NEUT</sub> / students<sub>NEUT</sub>
b. *dwa* / *dwoje* okna / pola / euro
two windows<sub>NEUT</sub> / fields<sub>NEUT</sub> / euro<sub>NEUT</sub>

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<sub>P1</sub> / people<sub>P1</sub> / parents<sub>P1</sub> stood
b. skrzypce / drzwi / okulary / urodziny stali
violin<sub>P1</sub> / door<sub>P1</sub> / glasses<sub>P1</sub> / birthday<sub>P1</sub> 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<sub>P2</sub> / doors<sub>P2</sub>

b. *dwoje okularów* / *urodzin*
two glasses<sub>P3</sub> / birthdays<sub>P3</sub>

The need for these three *plurale tantum* genders has recently been questioned by, independently, Przepiórkowski et al. 2002 and Wolinski 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 Wolinski 2001 and Przepiórkowski et al. 2002 argue that the fact that some lexemes lack singular forms is not a sufficient criterion for delimiting genders, Wolinski 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:

\[(18)\]

\begin{enumerate}
\item \textit{dwoje wujostwa / państwa / rodziców}
\two \uncle \text{ and} \aunt_{\text{p1}} / \text{people}_{\text{p1}} / \text{parents}_{\text{p1}}
\item \textit{*dwoje chłopców / mężczyźni / facetów}
\two \text{\textbf{boys}}_{\text{m1}} / \text{men}_{\text{m1}} / \text{guys}_{\text{m1}}
\item \textit{dwaj chłopcy / mężczyźni / faceci / studenci / uczniowie…}
\two \text{\textbf{boys}}_{\text{m1}} / \text{men}_{\text{m1}} / \text{guys}_{\text{m1}} / \text{students}_{\text{m1}} / \text{pupils}_{\text{m1}}
\end{enumerate}

Note, however, that also nouns that are conspicuously \textit{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:

\[(19)\]
\[\text{dwoje studentów / uczniów / adeptów}
\two \text{\textbf{students}} / \text{pupils} / \text{adepts}\]

Since the \textit{p1} nouns in (18a) refer to groups of mixed sex by virtue of their lexical meanings, any account of (19), involving uncontroversial \textit{m1} nouns, will also account for (18a) with the relevant nouns analysed as having the \textit{m1} gender. I conclude that the contrast in (18) is not a sufficient reason to postulate a separate \textit{plurale tantum} human masculine gender such as \textit{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 \textit{p2} and \textit{p3} genders, either: even though the specific gender of such \textit{plurale tantum} nouns cannot be empirically determined above establishing that they are not \textit{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 \textit{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, proliferate the number of genders beyond those proposed by Mańczak 1956, and which takes into account the partially pragmatic nature of connectivity with collective numerals.

4.3 Gender and numerals in HPSG

In order to account for the correlation between gender and connectivity with numerals, let us introduce an attribute to the now no longer atomic type \textit{GENDER}:

\footnote{Note that collective numeral forms such as \textit{dwoje} in (18a) always occur with genitive nouns, while forms such as \textit{dwa}, \textit{dwa} and \textit{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; \textit{e.g.}, Przepiórkowski 1999 and references therein, as well as Rappaport’s article in this volume.}
(20) \[ \text{gender} \\
\text{NUM num} \]

The values of \text{NUM} are of type \text{num}, which has three subtypes: \text{main}, for nouns which may combine with ordinary numerals, \text{collective}, for nouns which may combine with collective numerals, and \text{non-accommodating}, for nouns which do not combine with numerals at all (cf. Saloni’s gender \text{p3} discussed above):

(21) \[
\begin{array}{c}\text{num} \\
\text{main} \quad \text{coll} \quad \text{non-accom} \end{array}
\]

Considering ordinary (non-plurale tantum) nouns first, we observe that nouns in \text{m1}, \text{m2} and \text{fem} always combine with main numerals, i.e., for those genders the value of \text{NUM} is always \text{main}. This is represented in the following modified gender hierarchy:

(22) \[
\begin{array}{c}\text{masc-neut} \\
\text{masc} \quad \text{m23neut} \\
\text{m12} \quad \text{m23} \\
\text{m1} \quad \text{m2} \quad \text{m3} \quad \text{neut} \end{array}
\]

In case of \text{m2}, \text{m3} and \text{fem} nouns, their lexical entries only need to contain the information about their basic gender; the value of their \text{NUM} attribute will follow from the hierarchy above and will always be \text{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 \text{dwoje}, and other neuter nouns may only occur with ordinary numeral forms, such as \text{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. \text{okno} ‘window’: \[ \text{neut} \\
\text{NUM main} \]

b. \text{dziecko} ‘child’: \[ \text{neut} \\
\text{NUM coll} \]

Finally, in case of \text{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) \[
\begin{array}{c}
\text{CAT} \\
\text{GENDER} \\
\text{CON} \\
\text{CONX}
\end{array}
\begin{array}{c}
\text{noun} \\
\text{m1} \\
1 \\
2
\end{array}
\] \rightarrow \begin{cases}
\text{refers-to-mixed-sex} \quad \text{if } 1 = \text{coll} \\
\text{refers-to-same-sex} \quad \text{if } 1 = \text{main}
\end{cases}

In order to fully state this principle, relations \textit{refers-to-mixed-sex} and \textit{refers-to-same-sex} would have to be formally defined and possible values of the semantic attribute \text{CON} and the pragmatic attribute \text{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 \textit{m1} nouns do not need to contain information about their \text{NUM} values: when such nouns refer, in a given context, to a mixed sex entity, the value of their \text{NUM} is \text{coll}, by virtue of the principle above; otherwise it is \text{main}. Of course, the lexical semantics of the vast majority of \textit{m1} nouns is incompatible with reference to mixed sex, and such nouns will always combine with \text{main} numerals only.

Before discussing the place of \textit{plurale tantum} nouns in this setup, let us consider gender agreement in numeral phrases. I assume the following lexically specified values of \text{GENDER} for different forms of the numeral \textit{dwa} ‘two’:

(25) a. \textit{dwoje}:
\[
\begin{array}{c}
\text{gender} \\
\text{NUM} \\
\text{coll}
\end{array}
\]

b. \textit{dwaj}:
\[
\begin{array}{c}
m1 \\
\text{NUM} \\
\text{main}
\end{array}
\]

c. \textit{dwa}:
\[
\begin{array}{c}
m2/mneut \\
\text{NUM} \\
\text{main}
\end{array}
\]

d. \textit{dwie}:
\[
\begin{array}{c}
fem \\
\text{NUM} \\
\text{main}
\end{array}
\]

In case of the phrase \textit{dwie kobiety} ‘two women’, lexically specified \text{GENDER} values of \textit{dwie}, \text{fem} \text{NUM} \text{main}, and \textit{kobiety}, \text{fem}, are compatible, and the combined value is \text{fem} \text{NUM} \text{main}, which also satisfies the requirement in (22) that the \text{NUM} value of \text{fem} nouns be \text{main}.

(26) \textit{dwie kobiety} ‘two women’:
\[
\text{fem} \text{NUM} \text{main} + \text{fem} = \text{fem} \text{NUM} \text{main}
\]

The feminine noun \textit{kobiety} cannot be modified by any other numeral form of \textit{dwa}: e.g., in case of \textit{dwaj}, the type \textit{m1} is incompatible with the type \text{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:

\begin{align*}
\text{dwoj kobiety: } &\begin{bmatrix} m1 \\ \text{NUM main} \end{bmatrix} + \text{fem} = * \\
\text{dwoje kobiet: } &\begin{bmatrix} \text{gender} \\ \text{NUM coll} \end{bmatrix} + \text{fem} = \begin{bmatrix} \text{fem} \\ \text{NUM coll} \end{bmatrix} = * \quad \text{(cf. (22))}
\end{align*}

Similarly, for the phrase dwa konie ‘two horses’, gender agreement consists in the simultaneous satisfaction of both lexical requirements, i.e., \( \begin{bmatrix} m23\text{neut} \\ \text{NUM main} \end{bmatrix} \) for dwa and m2 for konie:

\begin{align*}
\text{dwa konie ‘two horses’: } &\begin{bmatrix} m23\text{neut} \\ \text{NUM main} \end{bmatrix} + m2 = \begin{bmatrix} m2 \\ \text{NUM main} \end{bmatrix}
\end{align*}

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

\begin{align*}
\text{dwa okna ‘two windows’: } &\begin{bmatrix} m23\text{neut} \\ \text{NUM main} \end{bmatrix} + \text{neut} = \begin{bmatrix} \text{neut} \\ \text{NUM main} \end{bmatrix} \\
\text{dwa dzieci ‘two children’: } &\begin{bmatrix} m23\text{neut} \\ \text{NUM main} \end{bmatrix} + \text{neut} = \begin{bmatrix} \text{neut} \\ \text{NUM coll} \end{bmatrix} = * \\
\text{dwie studenci ‘two students’: } &\begin{bmatrix} \text{fem} \\ \text{NUM main} \end{bmatrix} + \text{m1} = * \\
\text{dwoje dzieci ‘two children’: } &\begin{bmatrix} \text{gender} \\ \text{NUM coll} \end{bmatrix} + \text{neut} = \begin{bmatrix} \text{neut} \\ \text{NUM coll} \end{bmatrix}
\end{align*}

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

\begin{align*}
\text{dwoj studenti ‘two students’: } &\begin{bmatrix} m1 \\ \text{NUM main} \end{bmatrix} + \text{m1} = \begin{bmatrix} m1 \\ \text{NUM main} \end{bmatrix} \\
\text{dwa studenti ‘two students’: } &\begin{bmatrix} m23\text{neut} \\ \text{NUM main} \end{bmatrix} + \text{m1} = * \\
\text{dwa studenti ‘two students’: } &\begin{bmatrix} \text{fem} \\ \text{NUM main} \end{bmatrix} + \text{m1} = * \\
\text{dwoje studentów ‘two students’: } &\begin{bmatrix} \text{gender} \\ \text{NUM coll} \end{bmatrix} + \text{m1} = \begin{bmatrix} m1 \\ \text{NUM coll} \end{bmatrix}
\end{align*}

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 \( p_2 \) nouns: \[
    \begin{array}{c}
    \text{non-masc-hum} \\
    \text{NUM} \\
    \text{coll}
    \end{array}
    \]
    b. so called \( p_3 \) nouns: \[
    \begin{array}{c}
    \text{non-masc-hum} \\
    \text{NUM} \\
    \text{non-accom}
    \end{array}
    \]

Now, according to the gender type hierarchy in (22), the only non-masc-hum gender whose NUM value may be \textit{coll} or non-accom is neut: in case of the other non-masc-hum genders, i.e., \( m_2 \), \( m_3 \) and \( f \), their NUM value is \textit{main}. That means that, in fact, \( p_2 \) and \( p_3 \) genders are neuter genders differing in NUM values:

(32)  
    a. \( p_2 \): \[
    \begin{array}{c}
    \text{neut} \\
    \text{NUM} \\
    \text{coll}
    \end{array}
    \]
    b. \( p_3 \): \[
    \begin{array}{c}
    \text{neut} \\
    \text{NUM} \\
    \text{non-accom}
    \end{array}
    \]

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 Wolinski 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 non-atomic 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


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