

Defining ‘subgender’: Virile and devirilized nouns in Polish[☆]

Dunstan Brown*

Linguistic and International Studies, University of Surrey, Guildford, Surrey GU2 5XH, UK

Received 9 October 1997; accepted 20 November 1997

Abstract

As a contribution to morphological typology, we analyse dependencies between grammatical categories within a formal framework, namely Network Morphology. The dependencies are expressed by Category Dependency Constraints of the framework, and they determine the dependency of case on number, of gender on number, and of gender on case. Within this area a particularly interesting challenge is the notion of ‘subgender’, which is an additional gender distinction within a minimal subset of the paradigm (Corbett, 1991: 163). In certain instances it may not be clear whether one is dealing with a main gender distinction or one involving subgender, as in the difficult case of Polish masculine-personal nouns. Progress with regard to this question can be made by applying Category Dependency Constraints. We must recognise the ‘structured gender’ **masc person** as one of the values of the main gender system, and also three different values for the subgender of animacy, inanimate, animate and person. The **person** subgender fuses with **masc** to create the new structured gender **masc person** when it spreads to the nominative case, and from there to agreement targets which do not realise case. Category Dependency Constraints predict that subgender could only arise in languages with morphological case, because otherwise the important minimal subset criterion

[☆] The research reported here was supported by the Economic and Social Research Council under grant number R000236063. This support is gratefully acknowledged. Earlier versions of this paper were presented at the annual meeting of the Deutsche Gesellschaft für Sprachwissenschaft, Freiburg, February 28–March 1 1996, and at the annual meeting of the British Association for Slavonic and East European Studies, Cambridge, March 31, 1996. I would like to thank participants in the morphology session at both conferences for comments and suggestions, in particular Nigel Gotteri, Albert Ortmann, Tobias Scheer, Terence Wade and Daniel Weiss. I am grateful to Tanja Anstatt, Marek Gensler, Katarzyna Jaszczolt, Adam Przepiórkowski, Anna Siewierska, Jacek Słodkiewicz and Grażyna Zajaczkowska for help with informant work in different ways. Adam Przepiórkowski also kindly gave me some useful references. Greville Corbett, Norman Fraser, Andrew Hippisley and Stephen Hutchings are to be thanked for reading various drafts. I would also like to acknowledge the help of Paul Marriott on dependencies, Krzysztof Czuba and Jadwiga Zieniukowa. Of course, any misapprehensions or mistakes are my responsibility. Attendance at the conference in Freiburg was in part made possible by generous support from the British Council under the Academic Research Collaboration Programme, which has enabled co-operation with Dieter Wunderlich’s SFB 282 research group in Düsseldorf. I am very grateful for this support.

* Phone: +44 1483 25 9957; Fax: +44 1483 302605; E-mail: d.brown@surrey.ac.uk

would not be fulfilled. The examples of ‘devirilized’ nouns show that our approach is justified, as they may take either the exceptional case default for the main gender, or, less often, the exceptional case default for subgender. This article is therefore a contribution to the investigation of universals in grammatical categories. Moreover, since it is based on a formal analysis (in Appendix) which has been checked computationally, this demonstrates that the theory outlined here makes the correct predictions.

1. Introduction

Semantics, morphology or phonology may determine gender in a natural language, and there is always a semantic basis to any gender system (Corbett, 1991: 7–70). This article considers Corbett’s (1991: 163) formulation of ‘subgender’ as applied to Polish and Russian, and attempts a more formal definition within the Network Morphology framework. It is claimed that Corbett’s distinction between ‘subgender’ in the Slavonic languages and the ‘combined gender systems’ of the Mba languages (Corbett, 1991: 187), where the former is restricted to a limited number and case combination and the latter is much less restricted, is accounted for by ‘Category Dependency Constraints’ of the framework. In particular, these constraints determine the dominance of number relative to case, of number relative to gender, and of case relative to gender. These constraints are local, as they apply only to two categories in relation to each other and may be limited to a certain syntactic class, such as adjectives, but they are sufficient to determine the ordering of *feature* specifications containing two or more *features*. The constraints force on us the claim that subgender could only arise in languages with morphological case. Another claim is that a subgender may evolve into a full-blown gender which has a structured *value*, as appears to be happening with Polish. Crucially, it is claimed that the personal subgender creates a new structured gender when it spreads to the nominative case and from there to agreement targets which do not realize case. The explicit analysis is represented in the Appendix and has been checked with the aid of a computer.

2. Category Dependency Constraints

Category Dependency Constraints are imposed by the Network Morphology framework on morphosyntactic categories. More specifically, it is claimed that the presence of *features* from a category A may influence the presence of *features* from a category B, but category B may not influence category A. These Category Dependency Constraints (CDCs) enable the construction of more sophisticated typological generalizations from less sophisticated ones. A CDC is a statement about the conditioning relation of two categories. In (1), which does not apply to nouns for reasons to be explained, it is stated that the category of number is in a conditioning relation to the category of gender.

- (1) GENDER | NUMBER

The interpretation of (1) should be, ‘Where a language has NUMBER and GENDER, a particular number may influence or condition the number of genders available, but a particular gender may never influence the number of numbers available’. We should note here that the dependencies outlined are not about the order of affixes per se. In principle, it would be possible for a gender affix to occur inside of a number affix on agreement targets. The important question for linguistic typology as we see it is whether morphology realises particular distinctions for a category and what categories condition loss of distinction within another category. Our approach should therefore not be confused with hierarchies of functional categories put forward by, among others, Wunderlich and Fabri (1995: 246–247), where claims about the order of affixes are made.

As an illustration of what we mean, we may take modern Russian, where there are three main genders in the singular and none in the plural, and the verbs have gender agreement in the past.¹

- (2) komnat-a byl-a
 room-SG.NOM was-SG.FEM
 ‘the room was’
- (3) mal’čik byl
 boy-SG.NOM was-SG.MASC
 ‘the boy was’
- (4) okn-o byl-o
 window-SG.NOM was-SG.NEUT
 ‘the window was’
- (5) komnat-y byl-i
 room-PL.NOM was-PL
 ‘the rooms were’
- (6) mal’čik-i byl-i
 boy-PL.NOM was-PL
 ‘the boys were’
- (7) okn-a byl-i
 window-PL.NOM was-PL
 ‘the windows were’

Network Morphology treats the morphology which realizes the number and case *features* of the nouns as not specifying a morphosyntactic *feature* for gender, and therefore noun morphology is not subject to CDC (1). By definition morphosyntactic *features* must vary for the category to which they belong in order to be legitimate. Gender is seen here as inherent to the noun in question. In contrast to this, it does alter the form of the agreeing item, in this case the predicate. Furthermore, we can see that the plural *feature* conditions or triggers loss of gender distinction.² Although

¹ These Russian examples are in transliteration rather than a phonological transcription.

² We take the idea of triggering or conditioning features from Carstairs (1984). Given the CDC in (1) we rule out certain of the conditioning features he has. He has examples of gender features conditioning

the choice of triggering *feature* may differ from language to language the CDC in (1) forces on us the typological claim that a number *feature* may influence gender distinction, but never the other way round.

The next CDC which will concern us in this article is the one which defines the relation between case and number. This is given in (8).

(8) CASE | NUMBER

Where case and number occur together, a number *feature* may determine the number of case distinctions, but not the other way round. In Russian, for example, the marginal second locative and second genitive cases occur in the singular only, but not the plural. In German nouns, singular and plural are distinguished and case less so, but the plural conditions marking of dative case with the *-n* formative.³

For the three categories which are important for the definition of subgender the other CDC to be stated is for the relative importance of case to gender in (9).

(9) GENDER | CASE

In both Russian and Polish adjectives, masculine and neuter genders are syncretic throughout the oblique cases of the singular. Apparent counterexamples to (9) either involve a conflation of the concept of noun class with gender, or an identification of the inherent semantics of a noun with that category. In both (1) and (9) GENDER is used as a category which shows up on agreement targets, such as adjectives and verbs. It is, of course, determined by the noun. Certain noun classes may show more case distinctions. For instance, class II nouns in both Polish and Russian do not distinguish the dative and prepositional (locative) cases formally, and nouns with female referents are assigned to this class. A number of languages could be cited with regard to the differentiation of case for particular nouns on the basis of their reference. For instance, Yessan-Mayo, a non-Ndu language, spoken in the East Sepik District of Papua New Guinea (Foreman, 1974: 10), uses a dative case marker to distinguish animate 'undergoers', whereas it does not for inanimates (Foley, 1986: 101). We claim that phenomena of this type do not involve a particular morphosyntactic *feature* of the GENDER category conditioning the presence of certain CASE *features*, but is the direct result of the noun semantics.

number syncretism, which we claim cannot be possible. Apparent counterexamples to the CDC in (1) are usually of three types. The first is where there is a feature of another category which is in fact the conditioning feature. The second involve 'takeovers' or 'referrals' (Zwicky, 1985). The third is where one set of features for one word class are confused with another. Gender is a good example of this. To claim that gender features for nouns and adjectives are the same thing is merely a play on words. Adjective gender is not inherent and changes according to agreement controller. Noun gender is inherent and does not change. Hence, counterexamples to (1) on the basis of noun morphology are due to differences in inflectional class, which may determine, but is not exactly the same thing as, gender.

³ Related work by Bleiching et al. (1996) and Cahill and Gazdar (1997) addresses issues of this kind for German. Bleiching, Drexel and Gibbon discuss noun syncretism, and Cahill and Gazdar, in the first of a series of papers which give a complete explicit account of inflection in standard German, use DATR to formulate their analysis of adjectives, determiners and pronouns.

As will become apparent, use of the CDCs (8) and (9) is important for our definition of the animate subgender in Polish and Russian.

Where a language has morphological marking of number, case and gender, we may combine the CDCs (8) and (9) to yield (10).

(10) GENDER | CASE | NUMBER

Note that it would not be possible to determine (10) from (1) and (9) alone, as the order of NUMBER in relation to CASE would be underdetermined.⁴ In our analysis of both Polish and Russian it will be argued that number *features* may influence case *features* and that both number and case have an effect on gender distinction.

Network Morphology is a declarative framework and CDCs determine whether certain of the basic entities of a Network Morphology theory are well formed. The basic entities in question are *paths*. *Paths* are ordered *attribute* structures. *Attributes* are either morphosyntactic *features*, which belong to categories such as CASE or NUMBER, or *hierarchy identifiers* which identify a particular level of linguistic structure. An example of a *hierarchy identifier* is the *attribute* 'mor', which identifies the morphological level of linguistic structure in question. Precursing our use of the DATR formalism (Evans and Gazdar, 1996) to express Network Morphology generalisations, *paths* are enclosed in angle brackets. A convention, the *Hierarchy Identifier Convention*, constrains the order of *attributes* in a *path* by requiring that a *hierarchy identifier*, such as the *attribute* 'mor', occur before any *features*. In (11) the *path* is illegal because the *hierarchy identifier* does not come first. In (12) the *path* is legal, because the *hierarchy identifier* comes first. In fact, it should be noted that representing the CDCs (1), (8), (9) and (10) as DATR *paths* will involve the opposite ordering from the 'l' notation. Hence, the first, or conditioned, category with the 'l' notation will appear further to the right in a DATR *path*.

(11) *<pl mor acc>

(12) <mor pl acc>

In (11) the category of number is not included within morphology. Although it is possible that number will not be expressed morphologically, it does not make sense to use such a *feature* in a *path*, if it is not expressed at that level of linguistic structure. Adopting *hierarchy identifiers* and requiring that they occur at the beginning of each *path* means that we must always be explicit about the categories which are encoded at that level of linguistic structure.

So far we have illustrated how CDCs partially determine the order of *features* within a *path*. *Features* have to have a realization. This is one of the roles of a Net-

⁴ In fact, we could argue that the presence of certain CDCs could be a matter of parameters, whereby NUMBER | CASE would be a possibility for languages of a certain type. This might be required for Koryak, which has a singular, dual, plural distinction in the absolutive case, but no such distinction in the other cases for nouns with inanimate referents (Žukova, 1972: 95, 123). An interesting point to note is that the number differentiation occurs for the absolutive, an 'unmarked' or 'core' case. We would still wish to claim that some CDCs, such as (1), are universal.

work Morphology *fact*. It pairs a *path* with a *value*, it pairs a *path* with another *path*, or a combination of both. An example *fact* is given in (13).

(13) <mor sg loc> == "<stem>" _e

The *fact* in (13) is expressed in the DATR formalism developed by Evans and Gazdar (see Evans and Gazdar, 1989a; Evans and Gazdar, 1989b; Evans and Gazdar, 1996; Keller, 1995). Use of the DATR formalism does not oblige one to adopt the order of *attributes* in (13). The Network Morphology framework requires that the order of *attributes* in the *path* in (13) is determined by the *Hierarchy Identifier Convention* and the CDC (8). This *path* is paired with another *path* which contains solely a *hierarchy identifier* 'stem' and the *value* '_e'. Informally, we may say that the double quotes, or *indirection markers*, surrounding the *path* on the right-hand side state that the *value* with which this *path* must also be paired will depend on the lexical item in question. This *value* is then concatenated with the ending '_e'. Hence (13) states that morphological realization of the singular locative is a concatenation of the stem of the particular lexical item in question plus the ending *-e*.

3. The organisation of morphological knowledge

So far we have outlined how CDCs and the *Hierarchy Identifier Convention* determine the order of *attributes* in the *paths* which occur in *facts*. More importantly, we need to distinguish particular morphological classes from each other in order to state which class of words particular *facts* apply to. For instance, we would wish to say that the *fact* in (13) generally holds for Polish nouns. In other words we need to stipulate a location for particular *facts*.

Network Morphology treats language as a *network* of linguistic information. *Facts* of the type we have introduced are stored at locations in that *network*. These locations are called *nodes*. *Facts* are located at particular *nodes*, and the connectivity of the *network* determines which *facts* are available at which *node*. This is not a new view of language. Flickinger (1987), for example, treats lexical knowledge in terms of inheritance hierarchies in which information may be passed down from one *node* to another. Such an inheritance-based approach is used in Word Grammar (Hudson, 1990; Fraser and Hudson, 1992), and in a number of recent natural language processing computer systems (summarized in Daelemans et al., 1992). Furthermore, this approach can also be compared with the semantic nets approach for representing real world knowledge, which goes back at least as far as Quillian (1968) and is discussed in detail in Fahlman (1979).

From the above it would be wrong to assume that there are no distinctions between the general understanding of the *network* of morphological knowledge and the *nodes* where individual *facts* are located. A distinction is made between the *network* of morphological knowledge as a whole and the *hierarchies* which are part of that *network*. Furthermore, different types of *hierarchy* are recognised as part of the knowledge of morphology.

The *lexemic hierarchy* is the central *hierarchy* of the *network* and gives information regarding the location of other sources of information, which will often be *nodes* from other *hierarchies*. The *lexemic hierarchy* looks something like Fig. 1.

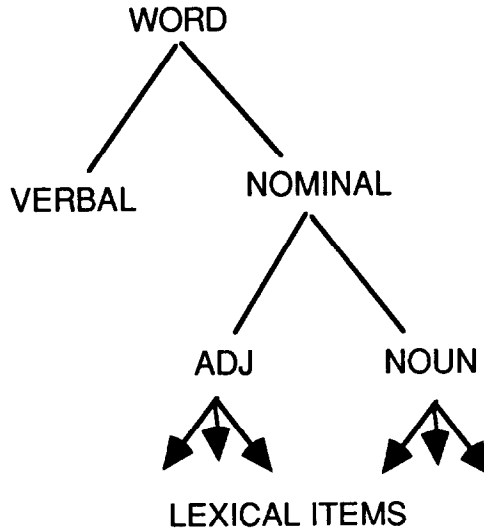


Fig. 1. Part of the lexemic hierarchy

Lexical items inherit from the ADJ and NOUN *nodes*, but we also allow for varying degrees of idiosyncrasy. Default generalizations may be made about sources of morphological information, for instance, but they may be overridden by a lexical entry. To illustrate what kind of *facts* are found in the *lexemic hierarchy* we take the *node* NOUN in its formal representation in DATR in (14).

```

(14) NOUN:
    <> == NOMINAL                                %[74]
    <syn cat> == n                                %[76]
    <syn gender> == GENDER:< "<syn animacy>" >    %[78]
    <syn animacy> == "<sem animacy>"                %[79]
    <sem animacy> == ANIMACY:< "<sem sex>" >        %[80]
    ...
  
```

The *fact* [74] at NOUN states that any *fact* not specified at NOUN will be inherited from the *node* NOMINAL. This is what is called a *hierarchy relation*, as it specifies another *node* as a potential source for all information not specified locally. The next *fact* which has been given in (14) is [76]. We have left out [75] which can be found in the Appendix. *Fact* [76] begins with a *path* which starts with the *hierarchy identifier* 'syn' for syntactic information. It states the syntactic category of nouns. The *path* in *fact* [80] contains the *hierarchy identifier* for semantics 'sem'. The *node* ANIMACY is

referred to. This is an example of a *network relation*, as a *node* is referred to for a particular type of *fact*, in this case one about the semantics of animacy. This *network relation* also involves the evaluation of another *fact* about the biological sex of the referent of the lexical item in question in order to determine which *fact* to refer to at the *node* ANIMACY.

Note that the *node* in (14) contains *facts* which have different *hierarchy identifiers* in their left-hand *paths*. One *path* contains a 'syn' *hierarchy identifier*, the other a 'sem' *hierarchy identifier*. There are other *hierarchies* in Network Morphology which are marked by one *hierarchy identifier* only. This is the reason why the *lexemic hierarchy* is named in the way that it is. It is central to the *network* of morphological knowledge and combines different levels of linguistic structure. Lexical entries are terminal *nodes* at the bottom of the *lexemic hierarchy*. In contrast to other morphological frameworks it is assumed that lexical entries are lexemes in the sense of Matthews (1972: 161), namely an abstraction over a paradigm of fully inflected forms and their grammatical representations.⁵ This is motivated by the fact that it would probably be impossible to deal with the morphology of highly 'polysynthetic' languages if one assumed that each element of the paradigm could be listed as a separate lexical entry.

Network Morphology does not assume that the whole of lexical structure may be represented by one large *hierarchy*. Instead there is an assumption that there may be a series of parallel *hierarchies*. This means that there need not be a direct relation between the generalizations for morphology as a whole and, say, generalizations about the syntactic properties of lexical items.

It is the function of the *morphological hierarchy* to provide the inflected forms of a particular lexical item. At the NOMINAL *node* in Fig. 1 it is specified for Polish and other fusional languages, such as Russian, that inflectional class must be assigned in order to determine the inflection of the lexical item. There are two ways in which inflectional class is assigned. By evaluating semantic and morphonological information stipulated or inherited by a particular lexical item, or by stipulating declension class for a particular lexical item.

An example of a lexical item is the Polish noun *Żona* 'wife' which inherits from the NOUN *node* in the *lexemic hierarchy*, which in turn inherits from the NOMINAL *node*. The lexical item *Żona* is represented in (15).⁶

(15) Zzona:

<> == NOUN	%[250]
<gloss> == wife	%[251]

⁵ Readers should be careful not to assume that lexeme and lexical entry are synonyms. In our use of the term here it would not necessarily be true to say that a lexical entry was a lexeme, if it was a fully inflected form corresponding to other entries which were fully inflected. A framework which posits different lexical entries for *eat* and *eats* could not refer to these as lexemes, if it assumed the definition here. Network Morphology assumes one lexical entry for these which is the lexeme EAT. It is one of the roles of morphology to provide the appropriate forms of the paradigm.

⁶ In the DATR representation the lexical entry and its root are given in a lower ASCII phonological transcription, partly explained in the fragment in the Appendix. Here double z is used to transcribe <z>.

<root> == zzo	%[252]
<root final> == N	%[253]
<sem sex> == female	%[254]
<sem animacy> == person.	%[255]

Note that there is no information about declension class given here. This is because the semantics of the lexeme, namely that it has a female referent, determine its declension class membership. This is what is known as a semantically based interdependency between the *lexemic hierarchy* and the *morphological hierarchy*.

In distinguishing a *lexemic hierarchy* and a *morphological hierarchy* Network Morphology assumes that there is a difference between lexemes as signs and the morphology which realises particular morphosyntactic *features*. This obviously makes it distinct from so-called morpheme-based approaches which assume that affixes have the same status as lexical entries as do uninflected stems. Network Morphology is also distinguished from other lexeme-based approaches in stating that lexemes themselves are also part of a larger structure.

The relationship between the *lexemic* and the *morphological hierarchy* is as in Fig. 2, where the dashed lines indicate *network relations* between *hierarchies*.

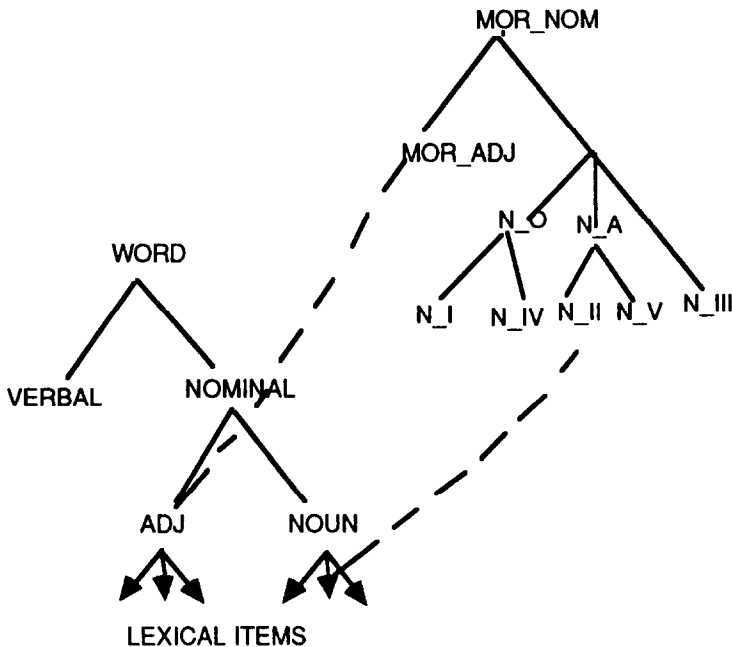


Fig. 2. Combined lexemic and morphological hierarchies for Polish

Fig. 2 is an idealisation in that all lexical entries will be addressed to the *morphological hierarchy* in order to determine the realisation of morphosyntactic words.

Note that we have implicit in this system at least two possible types of principle: principles which determine how information is ordered within a *hierarchy*, and principles which determine what kind of connections between *hierarchies* are possible.

One of the principles which determine how information is ordered within *hierarchies* is the *Overextended Ancestor Prohibition* (OAP). This basically states that an ancestor *node* cannot contain among its *facts* one which specifies a *path* which is more specific than a *path* at a *node* below it. In the definition in (16) this special kind of subsumption is called *path extension*. The *hierarchy relation* at the lower *node* is itself excluded from this requirement, as it contains an empty *path* which therefore subsumes any other *path*, and its presence determines whether the OAP should apply.

(16) *The Overextended Ancestor Prohibition*

If *node* A is in a *hierarchy relation* with B, no *path* (left-hand side) at B may extend the categories of a *path* (left-hand side) at A.

The OAP interacts with CDCs in an interesting way and helps us to make claims about the likelihood of morphology being shared or not. As we have argued, nouns do not specify the category of gender in their *paths*, because they do not change gender within their paradigm, whereas adjectives, as agreement targets, do. This means that certain adjectival *facts* will contain *paths* that specify gender, whereas noun *paths* will not. In (17), for instance, we give the *paths* associated with the realization of the singular nominative masculine of the Polish adjective *nowy* 'new' and the singular nominative of the Polish noun *hotel* 'hotel', which is masculine in gender.⁷

- (17) <mor sg nom masc> = nov _i.
 <mor sg nom> = hotel.

To specify the noun for gender would mean that we would expect it to vary in gender throughout its paradigm, as it can do for number and case. Generally, the OAP means that we should expect those *facts* which do not specify gender to be more likely candidates for sharing between nouns and adjectives. In Russian, for example, nouns share with adjectives the dative, instrumental and locative endings in the plural (Corbett and Fraser, 1993).⁸ If we compare in (18) the realizations of the plural instrumental of the Russian adjective *novyj* 'new' and the Russian noun *gostinica* 'hotel', which has feminine gender, we see that their *paths* are equal in specificity. (The Russian example is given in transcription.)

- (18) <mor pl inst> = nov _i _m'i.
 <mor pl inst> = gost'in'ic _a _m'i.

⁷ Here and elsewhere, the Polish examples cited are written in the standard orthography. Examples from the DATR fragment and derivable theorems are given in a phonological transcription.

⁸ Although their theme vowels differ, of course.

The *OAP* prohibits the realization of the singular nominative masculine of Polish adjectives from being a nominal default because of the presence of the gender *feature*. In (18) it allows the realization of the plural instrumental to be a nominal default, because of the lack of a gender *feature*. This is related to CDC (1), where plural number triggers, in the sense of Carstairs (1984), loss of gender distinction. Combining the CDC (1) with the *OAP* makes us claim that there is a relationship between the morphology shared by nouns and adjectives and the trigger *features* of Carstairs (1984).⁹

As we shall see, subgender introduces a second gender category, and we must ask whether this is also subject to CDCs (1) and (9), and how one accounts for the syncretism which both nouns and adjectives share for subgender. We show that CDCs (1) and (9) also apply to the dependency between subgender and number, as well as subgender and case.

4. Subgender

‘Subgender’ is a term used in Corbett (1991: 163) for agreement classes which “control minimally different sets of agreements”. The examples of subgender that Corbett gives are from the Slavonic languages. He contrasts subgender with ‘multiple gender systems’, exemplified by languages of the Mba group (Ubangian branch of Niger-Kordofanian). The difference is that ‘multiple gender systems’ do not occur in a minimally different set of agreements. In the Mba language itself certain facultative personal pronouns distinguish male human from other animates in the singular. If there is no case marking, this means that the distinction is not minimal, as it appears throughout a significant proportion of agreement.

- (19) mósóngò (6í) úmá.
 European ANIM 1.one
 ‘one European woman’
 (Pasch, 1986: 174)

⁹ Occasionally we do find instances where nouns and adjectives share a realization that must be specified for gender for the adjective. In Polish, for example, we find for adjectives that the singular instrumental feminine ending is *-q*, and this is the same as the singular instrumental ending for some of the declension classes which assign feminine gender. These are accounted for by orthogonal *network relations* between adjectival and noun declension *nodes*. The Network Morphology framework limits such *relations* in two ways. First, the number of allowable *network relations* between *nodes* in the same *hierarchy* is determined by the topological distance between them, defined in terms of the number of *nodes* between them when one travels via the lowest common ancestor. Second, the *network relations* in question can never involve a *fact* which contains a left-hand *path* which is extended by the *path* it refers to at another *node*. If adjectives and nouns share the same realization for a particular number, case and gender combination, it must be the adjective which borrows it from the noun, and never the other way round. This is because nouns do not specify gender in their paradigms, as it is inherent. If nouns referred to an adjective paradigm which specified, say, feminine gender, then they would be referring to an extension of their own paradigm. The only other way for adjectives and nouns to share agreement morphology is for it to be underspecified for gender, as we have seen in our discussion of (18) and the *Overextended Ancestor Prohibition*.

Class 1 agreement on the numeral is marked here by a lack of a prefix. The use of the pronouns is facultative, but they cover a significant proportion of the agreement set, as they occur throughout a particular number. They therefore do not meet the minimality requirement for subgender.

Another characteristic, although not necessarily a defining one, of ‘subgenders’ is ‘dependent target gender’ (Corbett, 1991: 164), whereby a number and case combination which contains a subgender is realised by syncretism, rather than by an independent marker. In Russian, for example, the accusative of masculine animate nouns in the singular and animate nouns in the plural is syncretic with the genitive case, whereas masculine inanimates, neuter and feminine nouns and inanimate nouns in the plural have the same form as the nominative. Hence, the animate subgender in Russian, dealt with in Corbett and Fraser (1993), is not realised by an independent form (Table 1).

Table 1
The animate subgender in Russian

	<i>student</i> ‘student’ (animate)	<i>zakon</i> ‘law’	<i>komnata</i> ‘room’	<i>ženščina</i> ‘woman’ (animate)	<i>vino</i> ‘wine’	<i>čudovišče</i> ‘monster’ (animate)
SG						
NOM	student	zakon	komnat-a	ženščin-a	vin-o	čudovišč-e
ACC	student-a	zakon	komnat-u	ženščin-u	vin-o	čudovišč-e
GEN	student-a	zakon-a	komnat-y	ženščin-y	vin-a	čudovišč-a
PL						
NOM	student-y	zakon-y	komnat-y	ženščin-y	vin-a	čudovišč-a
ACC	student-ov	zakon-y	komnat-y	ženščin	vin-a	čudovišč
GEN	student-ov	zakon-ov	komnat	ženščin	vin	čudovišč

Furthermore, and in virtue of its definition as a subgender, animacy affects agreement targets in the same way. In (20) are given the accusative forms (transliterated) of a Russian plural animate noun and a plural inanimate noun which have the same gender in the singular.

- (20) star-ye komnat-y
 old-PL.ACC.INAN room-PL.ACC
 ‘old rooms’ (acc=nom)
- star-yx ženščin
 old-PL.ACC.ANIM woman-PL.ACC
 ‘old women’ (acc=gen)

The subgender is differentiated only in the accusative and is syncretic with the forms of another case, either the nominative or genitive. In the plural in Russian it takes over as the sole gender distinction, because the main genders converge. The CDC (1)

which, where a language has both, orders number before gender, predicts that gender convergence may be determined by number. The CDC is still more liberal than Greenberg's (1963: 112) universal number 37, which states that there never will be more genders in non-singular numbers than in the singular, as the CDC does not state which *feature* should trigger convergence, only that a *feature* of that *category* can do so. We see this liberalism as an advantage, as we claim that plurality does condition a new gender distinction for Polish. Generally, plural number conditions loss of gender distinction. In either instance, it is number which conditions the presence or absence of particular gender features, in line with CDC (1). In addition, accusative case, in line with CDC (9), conditions a newer subgender distinction. It should be noted here that accusative case exists as a separate formal distinction in both Russian and Polish for nouns of the second declension.

Another point to note about the animate subgender in Russian is that it has no formal marking on agreement targets external to the noun phrase. There is, for example, no marking of animacy on the past tense of verbs, even though these are marked for gender.

In Fraser and Corbett's (1995) Network Morphology account of gender and animacy assignment in Russian the distinction between gender and subgender is made in terms of different *path: value* pairings. Every noun lexeme will be assigned a *value* for the *path* <syn gender> which is either masculine, feminine or neuter. In addition, it will be assigned a *value* for the *path* <syn animacy>. Evaluation of the theory presented by Fraser and Corbett (1995) yields, among other things, information about the gender and subgender of nouns, as in (21) where the words for 'law' and '(female) teacher' are given.

- (21) Zakon:<syn gender> = masc.
 Zakon:<syn animacy> = inanimate.

Učitel'n'ica:<syn gender> = fem.
 Učitel'n'ica:<syn animacy> = animate.

The subgender of animacy has a separate *path* with its own *values*, either animate or inanimate. This has the effect that animacy and gender can be semi-autonomous. It would be conceivable, however, that we could devise a theory in which the information about animacy was incorporated directly into the information about gender. Such a theory would accept the notion of a 'structured gender', that is a gender with a structured *value*. Conceivably, (21) would end up as (22).

- (22) Zakon:<syn gender> = masc inanimate.
 Učitel'n'ica:<syn gender> = fem animate.

Treating animacy and gender as one structure, as in (22), would be undesirable for Russian, because it would require one to state that the plural accusative of feminine and neuter animate are realized in the same way as the masculine animate. But whether something is masculine, feminine, or neuter is irrelevant in the plural.

Therefore (21) is better because it can easily be used to capture the fact that the animate subgender is the only gender distinction in the plural.

4.1. *Subgender in Polish*

Zieniukowa (1981: 22), who attributes the concept of subgender to Vaillant (1958), rejects it as applicable to Polish and states that it has not been generally taken up in the Polish grammatical tradition. This might explain why there is such controversy surrounding the number of genders posited for Polish. According to Corbett (1983a: 83) Klemensiewicz (1965: 51) said there are three, Mańczak (1956) claimed there are five, Brooks and Nalibow (1970: 137) six, and Wertz (1977) seven. Furthermore, Saloni (1976) says that there are nine. Czuba and Przepiórkowski (1995) give a fairly comprehensive account of Polish gender using sorted *feature* structures. Significantly, they propose a sort which combines both number and gender information. Although this provides for a non-redundant account of Polish, it would not fit with our general typological approach which tries to account for the relationship between particular morphosyntactic categories. Significantly, the sort gender first divides into singular and plural in this account (Czuba and Przepiórkowski, 1995: 28, 31), which suggests that the CDC (1) is well founded for Polish.

Genders are also established on the basis of evidence from numerals, but we believe that this is problematic for at least three reasons. First, as numerals may or may not be treated as heads, their status as regards controller and target gender needs to be determined. Second, as it has been shown that there is a tendency for numerals to become more noun-like as the numerical value that they indicate increases (Corbett, 1978: 368; Hurford, 1987: 187–197), this dimension must be taken into account. Third, some of the numeral types are often considered to be archaic. Any theory of Polish gender will have to tackle this at some point, but we believe that the best initial step is to consider the facts independent of numerals.

Polish divides up the masculine gender in the singular according to whether the nouns in question have an animate or an inanimate referent. This is, of course, similar to what happens in Russian. However, in the plural the situation is somewhat different.

Polish distinguishes masculine-personal nouns. The question is whether this could usefully be understood in terms of subgender (Corbett, 1991: 161–168), as it is only manifested in the forms of the nominative and accusative on nouns and agreeing attributives. Furthermore, as masculine-personal nouns are basically a subset of masculine nouns, is 'masculine-personal' a gender in its own right, or a combination of the 'masculine' gender and a subgender 'person'? However, for agreement targets there is a non-syncretic realization of the plural masculine-personal. For adjectives the plural nominative masculine-personal has its own realization. The masculine-personal has its own separate realization on past tense verbs with any combination of plural number and either first, second or third person. The 'dependent target gender' characteristic of subgender is not met by masculine-personal in Polish.

The forms of the adjective *nowy* ‘new’ (an agreement target) are given in the nominative and accusative case in Table 2.

Table 2

The nominative and accusative forms of *nowy* ‘new’

	MASC-PERS.	MASC. ANIM	MASC. INAN	FEM.	NEUTER
SG NOM	now-y	now-y	now-y	now-a	now-e
PL NOM	now-i	now-e	now-e	now-e	now-e
SG ACC	now-ego (as gen sg)	now-ego (as gen sg)	now-y	now-ą	now-e
PL ACC	now-ych (as gen pl)	now-e	now-e	now-e	now-e

For agreement targets in each number the accusative has the same form as the nominative for any inanimate noun, except for the singular accusative of class II nouns. Agreement with animate masculine nouns in the accusative singular requires an accusative singular form syncretic with the genitive singular. In the plural, only nouns with male human referents must have an accusative form syncretic with the genitive. Note also that the form of the masculine-personal in the nominative plural is different from those of all other (sub)genders, which would again show that ‘dependent target gender’ cannot be a defining characteristic, if we wish to claim that we are dealing with a subgender.¹⁰

Examples of agreement patterns to be found in Polish are given in (23)–(32).

- (23) Widz-ę t-ego chłopc-a
 see-SG.1 this-SG.ACC.MASC_PERS boy-SG.ACC
 ‘I see this boy.’
- (24) Widz-ę t-ego ps-a
 see-SG.1 this-SG.ACC.MASC.ANIM dog-SG.ACC
 ‘I see this dog.’
- (25) Widzę now-y pokój
 see-SG.1 new-SG.ACC.MASC.INAN room-SG.ACC
 ‘I see the new room.’
- (26) Widzę t-ą dziewczynk-ę.
 see-SG.1 this-SG.ACC.FEM girl-SG.ACC
 ‘I see this girl.’
- (27) Widzę now-e miast-o.
 see-SG.1 new-SG.ACC.NEUT town-SG.ACC
 ‘I see the new town.’
- (28) Widz-ę t-ych chłopc-ów
 see-SG.1 this-PL.ACC.MASC_PERS boy-PL.ACC
 ‘I see these boys.’

¹⁰ For consistency we try to use the term ‘male human’ for the semantic category and ‘masculine personal’ for the grammatical category.

- (29) Widz-ę t-e ps-y
 see-SG.1 this-PL.ACC.MASC.ANIM dog-PL.ACC
 'I see these dogs.'
- (30) Widzę now-e pokoj-e
 see-SG.1 new-SG.ACC.MASC.INAN room-PL.ACC
 'I see the new rooms.'
- (31) Widzę t-e dziewczynk-i
 see-SG.1 this-PL.ACC.FEM girl.PL.ACC
 'I see these girls.'
- (32) Widzę now-e miast-a
 see-SG.1 new-PL.ACC.NEUT town.PL.ACC
 'I see the new towns.'

In the examples (23)–(32) the noun inflection has not been treated as realising gender. As mentioned in relation to CDC (1), this is because gender is inherent to the noun and is either determined by its semantics or its membership of an inflectional class (formal gender). As already explained, it does not make sense to talk of noun and adjective gender in the same way, as adjectives vary with regard to gender inflection, whereas nouns do not. Examples (23) and (24) demonstrate that animate masculine nouns require the same form as the genitive in the singular, both for themselves and also for agreeing adjectives. The examples (25) and (27) show that nouns which are inanimate masculine or neuter gender will have a form syncretic with the nominative in the singular, again both for themselves and for the accompanying adjective (see Table 2). The feminine noun and accompanying adjective in (26) have a separate realization for the singular accusative. From the point of view of examples (23)–(27) Polish appears to be just like Russian, having an animate subgender. The sentences (28)–(32) show a different pattern of behaviour from Russian. Here we see that only masculine-personal nouns, as example (28), have accusative-genitive syncretism. In contrast, for all other nouns, including masculine nouns which do not have male human referents as in (29), the plural accusative is syncretic with the plural nominative.

Given the information about agreement patterns we have so far, we might consider that Polish has a gender system like that in Fig. 3. As masculine-personal has its own realization in the plural nominative (see Table 2), we should distinguish two target genders in the plural (masculine-personal and non-masculine-personal), as well as the three in the singular (masculine, feminine and neuter). If we consider Polish in terms of controller gender (Corbett, 1991: 151), then there are at least five possible routes from singular to plural in Fig. 3.

Fig. 3 is also a simplification, as the case dimension is not taken into account. The subgender of animacy (person, animate, inanimate) is manifested in the accusative case in both the singular and plural, but the two numbers in Polish treat animates in different ways. Hence, in order to account for the possible patterns, we must make reference to at least four categories: number, case, gender, and animacy. The fact that we must make reference to number and case in order to account for the gender patterns is predicted by CDCs (1) and (9).

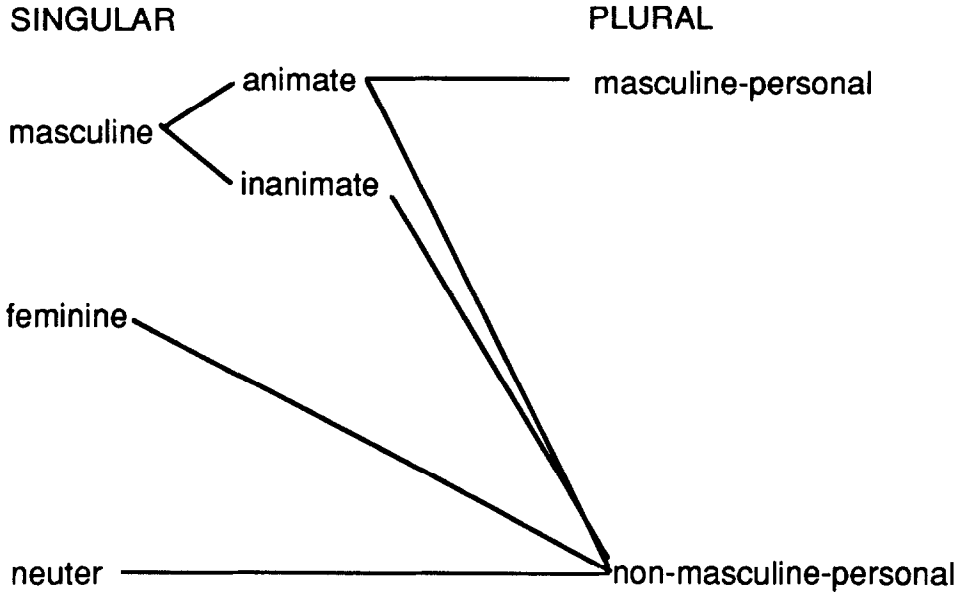


Fig. 3. A partial model of Polish gender.

Animacy is both a semantic and syntactic category, because there are examples where certain items may be taken to be animate, even though they are not in the real world. By default, however, syntactic animacy is determined by semantic animacy. We claim for Polish that there are three possible *values* for animacy: person, animate and inanimate. Persons are generally sex-differentiable, but so are some animals. Take, for instance, these examples from Saloni (1976: 49).

- (33) *Ta krowa jest bykiem.
 that.SG.NOM.FEM cow.SG.NOM is bull.SG.INST
 'That cow is a bull.'
- (34) *Ten byk jest krową
 that.SG.NOM.MASC bull.SG.NOM is cow.SG.INST
 'That bull is a cow.'

Sex differentiation is clearly associated with humans and certain animals, and we shall claim that it is the core of the Polish system. Weiss (1993) points out that there is an additional dimension of non-adulthood which assigns neuter gender in nouns such as *niemowlę* 'baby'. He also notes that there is a tendency for the male-female opposition to suppress this dimension (Weiss, 1993: 86).

Polish has five declension classes: masculine o-stems (class I), neuter o-stems (class IV), a-stems (class II), ja-stems (class V), and i-stems (class III). They are given here in Tables 3–5. The forms have been given in a phonological transcription

which corresponds to that of the DATR representation. The declensions are originally taken from Rothstein (1993: 698–701).¹¹

Table 3
Polish masculine and neuter o-stems

	I (Masculine o-stems) grud 'castle'		IV (Neuter o-stems) másto 'city'	
NOM	grud	grod-i	m'ast-o	m'ast-a
VOC	grodz'-e	grod-i	m'ast-o	m'ast-a
ACC	grud	grod-i	m'ast-o	m'ast-a
GEN	grod-u	grod-uv	m'ast-a	m'ast
DAT	grod-ov'í	grod-om	m'ast-u	m'ast-om
INST	grod-em	grod-am'í	m'ast-em	m'ast-am'í
LOC	grodz'-e	grod-ax	m'es'c'-e	m'ast-ax

Table 4
Polish *a-* and *ja-*stems

	II (Feminine a-stems) zzona 'wife'		V (Feminine ja-stems) bogini 'goddess'	
NOM	zzon-a	zzon-i	bogin'-i	bogin'-e
VOC	zzon-o	zzon-i	bogin'-i	bogin'-e
ACC	zzon-e	zzon-i	bogin'-e	bogin'-e
GEN	zzon-i	zzon	bogin'-i	bogin'
DAT	zzon'-e	zzon-om	bogin'-i	bogin'-om
INST	zzon-a	zzon-am'í	bogin'-a	bogin'-am'í
LOC	zzon'-e	zzon-ax	bogin'-i	bogin'-ax

Table 5
Polish *i-*stems

	III (Feminine i-stems) kos'c' 'bone'		
NOM		kos'c'	kos'c'-i
VOC		kos'c'	kos'c'-i
ACC		kos'c'	kos'c'-i
GEN		kos'c'-i	kos'c'
DAT		kos'c'-i	kos'c'-om
INST		kos'c'a	kos'c'm'í
LOC		kos'c'í	kos'c'-ax

¹¹ The examples are given in the transcription used in the DATR fragment. For class III the usual plural ending is *-am*. The example with the noun *kość* 'bone' is irregular in that it omits the theme vowel *-a*.

When a noun cannot be assigned gender on semantic grounds, it is assigned gender according to the declension class to which it belongs. The morphological assignment rules are similar to those of Russian as given by Corbett (1991: 36).

- (35) (i) nouns of class I are assigned masculine gender;
 (ii) nouns of class IV are assigned neuter gender;
 (iii) nouns of the other three classes are assigned feminine gender.

As with Russian the subgender of animacy is assigned on the basis of semantics. This means that this information need not be provided by declensional classes. The rules in (35) are not sufficient to provide us with the appropriate gender distinctions in the plural. If masculine-personal is a full gender distinction in its own right, we must be able to assign it to the appropriate nouns, and we also need to explain how it naturally follows that it will require the same forms as the ordinary masculine on agreement targets in the singular. This is done by claiming that masculine-personal is a 'structured gender'. However, we must motivate this claim. We shall first argue that we need three *values* for subgender: person, animate, inanimate.

Corbett (1991: 285) quotes an example from Doroszewski (1962: 237) in which a feminine and masculine animate noun are conjoined, but where the form of the verb is the plural masculine-personal. A similar example can be found in Zieniukowa (1981: 81), who states that feminine person nouns and masculine animates are often conjoined with the *-li* (masculine-personal) form of the verb.

- (36) Basia i piesek bawili się w kuchni.
 Basia and dog.DIM.SG.NOM played.PL.MASC_PERS REFL in kitchen.SG.LOC
 'Basia and the little dog were playing in the kitchen.'

The example in (36) indicates that we need to take into account the subgender characteristics of the conjuncts. Neither of the nouns involved would have masculine-personal agreement if they occurred as plurals outside of a conjunct. However, the fact that they are both animate nouns, and one denoting a female person, appears to have an effect on the choice of gender agreement found on the verb. In other words, although animacy usually plays no role in determining the morphology and agreement for a proper noun such as Basia, it still has to be part of the characterization for a singular noun (Corbett, 1991: 286). This should fall out from the assignment of 'person' as a *value* for the subgender of animacy. Gender resolution rules apply to examples such as (36) and must make reference to the subgender *value* in order to determine whether to select the masculine-personal gender for agreement with the conjunct.

We shall show that our Network Morphology based approach is able to provide the *features* required for gender resolution in such conjuncts and that (36) partially justifies our approach. Furthermore we shall tie this in with the question of 'devirilized' nouns, by showing that a *value* 'person' for the subgender of animacy together with a structured gender 'masc person' can account for differences between agreement behaviour in the nominative and accusative cases.

The next set of examples are used to back up our claim that Polish has both a sub-gender and a structured gender, and it is related to the agreement hierarchy (Corbett, 1983b). Fig. 4 is taken from Corbett (1991: 226), where moving from left to right we expect an increasing likelihood of semantic agreement.

The Agreement Hierarchy

attributive < predicate < relative pronoun < personal pronoun

Fig. 4. The Agreement Hierarchy (Corbett, 1991: 226).

Corbett (1991: 234) considers derogatory nouns in Polish where the plural nominative form is non-masculine-personal. Here we find that the attributive adjective, the verb (predicate) and relative pronoun take non-masculine-personal agreement, whereas the personal pronoun takes semantic agreement on the basis that the noun denotes male human beings. Example (37), taken from Corbett (1991: 234) shows the modifier ‘this’ and the relative pronoun taking syntactic agreement. Example (38), again taken from Corbett (1991: 234)¹², shows modifier and verb taking syntactic agreement and the personal pronoun agreeing with the noun on the basis of its semantics.

- (37) te łajdaki które mnie
 those.NON_MASC_PERS wretches who.NON_MASC_PERS me
 oszukaty ...
 cheated.NON_MASC_PERS
 ‘those wretches who cheated me ...’
- (38) te łajdaki zepsuty mi radio
 those.NON_MASC_PERS wretches damaged.NON_MASC_PERS me.DAT radio
 do reszty!
 to rest
 Oni już ci kiedyś zepsuli telewizor.
 They.MASC_PERS already you.DAT some.time damaged television
 ‘Those wretches have ruined my radio!’
 ‘They have already damaged your television.’

Examples such as (37) and (38) indicate that we need to have both semantic and syntactic information for agreement purposes. It is clear that there is a difference between masculine-personal and masculine as controller genders, and, as we have just shown, it is also not always the case that masculine-personal agreement and the semantics of being a male human being overlap, although masculine-personal gender is assigned by default to male humans. In fact, we claim that nouns of this ‘devirilized’ type, which are termed ‘depreciative’ by Saloni (1988: 161, 166), are assigned masculine (syntactic) gender, rather than masculine-personal (syntactic) gender.

¹² The example is based on the judgements of Ewa Jaworska.

It is still necessary to assume that ‘person’ is also a *value* for the animacy sub-gender, however. This is because the group of so-called ‘devirilized’ nouns, in which we may include *tajdaki*, may have non-masculine-personal agreement in the plural nominative, but still have accusative-genitive syncretism in the plural, a phenomenon usually ascribed solely to masculine-personal nouns. Examples are taken from Wertz (1977: 59). Note the genitive-accusative syncretism for both *mnisi* ‘monks’ and *karty* ‘dwarfs’, but the non-masculine-personal agreements and forms for *karty* in the nominative plural in Table 6.

Table 6
Masculine-personal and ‘devirilized’ nouns

	MASCULINE-PERSONAL		‘DEVIRILISED’	
PL. NOM	duż-i	mnis-i	duż-e	kart-y
PL. ACC	duż-ych	mnich-ów	duż-ych	kart-ów
PL. GEN	duż-ych	mnich-ów	duż-ych	kart-ów
	‘large	monks’	‘large’	dwarfs’

The situation is slightly more complicated than this. Huntley (1980: 193) cites an example from Polish with the noun *chtopy* ‘fellow’ (non-masculine-personal) where one may also get nominative-accusative syncretism. There is also dialectal variation, such that the special forms for the plural nominative are optional for male persons in certain western dialects and the so-called devirilized type of ending does not have pejorative meaning (Zagórski, 1967: 24–7. Map 8, cited by Huntley, 1980: 193).

In fact, there may be variation from speaker to speaker as regards which nouns may be ‘devirilized’, and there may be variation as to whether nominative-accusative syncretism may accompany the use of the non-masculine-personal form for the nominative plural. However, there are constraints on this variation. In examples 39a–f we have given the majority verdict of the informants consulted as regards acceptability.¹³

- (39) a. Ci głupi łąjdacy poszli do pracy.
 that.PL.NOM.MP stupid.PL.NOM.MP scoundrel.PL.NOM go.PAST.PL.MP to work
 ‘Those stupid scoundrels went to work.’
 b. Te głupie łąjdaki poszły do
 that.PL.NOM.NMP stupid.PL.NOM.NMP scoundrel.PL.NOM go.PAST.PL.NMP to
 pracy.
 work
 ‘Those stupid scoundrels went to work.’

¹³ There were five informants. They were asked to rate each sentence on a scale of [1]–[4], where [1] meant ‘I might say that’, [2] ‘I know people that might say that’, [3] ‘I’m not sure about that’, and [4] ‘No-one would ever say that’. Some informants remarked that the ‘devirilized’ forms were ‘humorous’ or ‘dialect’, in which case they marked them as [2]. I have treated sentences rated [1] or [2] as acceptable.

- c. Widzę tych głupich łądaków.
 see.1ST.SG that.PL.ACC.MP stupid.PL.ACC.MP scoundrel.PL.ACC
 'I see those stupid scoundrels.'
- d. *Widzę te głupie łądaki.
 see.1ST.SG that.PL.ACC.NMP stupid.PL.ACC.NMP scoundrel.PL.ACC
 'I see those stupid scoundrels.'
- e. Widzę tego głupiego łądaka.
 see.1ST.SG that.SG.ACC.MASC.ANIM stupid.SG.ACC.MASC.ANIM
 scoundrel.SG.ACC
 'I see that stupid scoundrel.'
- f. *Widzę ten głupi łądak.
 see.1ST.SG that.SG.ACC.MASC.INAN stupid.SG.ACC.MASC.INAN
 scoundrel.SG.ACC
 'I see that stupid scoundrel.'

These judgements are essentially the same as those of Wertz' (1977: 59) informant in Table 6. In (39a) the expected masculine-personal form is used for the noun and agreeing adjective and determiner. In (39b) the 'devirilized' non-masculine-personal form is used for the noun and agreeing adjective and determiner. In (39c) the accusative object of the verb is syncretic with the genitive on both noun and adjective and determiner. In (39d), which is unacceptable for the majority of informants, the noun, together with agreeing adjective and determiner, has the same form as the nominative (non-masculine-personal). In (39e) the accusative singular has the same form as the genitive singular on both noun, adjective and determiner (as would be expected for any animate noun). In (39f) we see that syncretism between the accusative and nominative is not possible (as would be expected for any animate noun).

One informant found (39d) acceptable, but said of (39c) that it is 'more common'. On the basis of this evidence one might assume that there are two 'dialects': one in which the form of the nominative plural may be determined by different means than the form of the accusative plural (the majority); another in which the form of the nominative plural is always determined by the same means as the form of the accusative plural (the minority). However, there is an additional consideration. The informants were also presented with the nouns *fanatyk* 'fanatic', *heretyk* 'heretic', *kleryk* 'cleric', *krytyk* 'critic', *medyk* 'medic', *muzyk* 'musician'. Three of the four informants whose ratings correspond to those in (39) rated all of these nouns in the same way as (39), ruling out the nominative-accusative form of the accusative plural. The other informant's judgement depended on the noun in question. The informant who found (39d) acceptable for the noun *łądak* 'scoundrel', however, accepted only two of the nouns, *heretyk* 'heretic' and *muzyk* 'musician', as possible for 'devirilization' (i.e. as (39b)). Furthermore, this informant rejected the plural nominative-accusative syncretism for these nouns. In other words, for the nouns *heretyk* 'heretic' and *muzyk* 'musician' the judgements of the informant who accepted (39d) for the

noun *tajdak* ‘scoundrel’ pattern in the same way as the judgements of the majority of informants for *tajdak* ‘scoundrel’, who rejected (39d). So judgements about which nouns ‘devirilization’ may apply to vary from speaker to speaker. Furthermore, it appears that for the same speaker some ‘devirilized’ nouns may be allowed nominative-accusative syncretism in the plural together with the ‘devirilized’ form in the plural nominative, whereas other nouns may only have the ‘devirilized’ form in the plural nominative (the patterning in (39)). There are essentially, therefore, two variants of ‘devirilized’ nouns: the most common one in which the loss of the special plural nominative form is still accompanied by genitive-accusative syncretism in the plural (variant A); the less common one in which loss of the special plural nominative form is accompanied by nominative-accusative syncretism (variant B).

In (40) we summarize the generalizations which appear to follow from these informant judgements. When we talk of a particular noun, the generalization also applies to its agreement targets, such as adjectives and determiners.

- (40) a. Optional (‘devirilized’) plural nominative-accusative syncretism implies an optional ‘devirilized’ form of the plural nominative.
 b. An optional ‘devirilized’ form of the plural nominative does not imply optional (‘devirilized’) plural nominative-accusative syncretism.
 c. For ‘devirilized’ nouns singular accusative-nominative syncretism is not allowed.

There are four *values* for gender: masculine, masculine-personal, feminine and neuter. If the determination of the singular accusative form is a matter of both gender and subgender, and the plural nominative a matter of gender, then ‘devirilization’ of the plural nominative involves changing (syntactic) gender from the structured masculine-personal to the basic masculine. The subgender of animacy has three *values*: person, animate, and inanimate. A noun which has its gender changed from masculine-personal to masculine may still have the *value* person for its subgender. This accounts for (40c), as changing gender does not necessarily mean a change in subgender. (40a) and (40b) can be accounted for, if the accusative case in the plural is seen as behaving in the same way as the accusative case in the singular. That is, the form of the accusative case for both numbers is determined by gender and subgender. The form of the plural nominative is determined solely according to gender. Hence, if a ‘devirilized’ noun changes its (syntactic) gender from masculine-personal to masculine, only the form of the plural nominative will change (as in (39)). Finally, a more subtle account is required for the patterning shown by the choices of our informant who accepted (39d) for *tajdak* ‘scoundrel’ but not for the two other nouns for which they accepted the non-masculine-personal form of the plural nominative. The nouns with a change in the plural nominative only are obviously to be treated in exactly the same way as *tajdak* for the other informants. The difference with *tajdak* in the case of the informant who accepted (39d) is that the subgender of animacy also appears to be affected. We shall claim for nouns of this type that both (syntactic) gender and (syntactic) subgender have been altered. However, in our account the change in gender from masculine-personal to masculine follows as an

automatic consequence of changing the subgender from person to animate (which will have no effect on the form of the singular accusative). In Table 7 we give the *values* for gender and for subgender.

Table 7

Values for gender and subgender

GENDER	SUBGENDER
masc	person
masc person	animate
fem	inanimate
neuter	

Our analysis ties in with more general facts about Polish and other Slavonic languages. First, full gender distinction, as opposed to subgender distinction, is found in Polish nominals only in the nominative case in the singular and plural, as allowed by our Category Dependency Constraints (1) and (8). For both Polish and Russian, as well as other Slavonic languages, in the singular oblique cases of adjectives the distinction between masculine and neuter gender is lost. Second, in Russian we find that the accusative case triggers distinction of subgender that is not found elsewhere. This is in line with the Category Dependency Constraint (9) between case and gender and corresponds to the claim that subgender is limited to a small subset of the paradigm. This triggering of subgender distinction in the accusative case of the singular is also clear for Polish. Third, the fact that masculine-personal in Polish is marked by its own unique realization in the plural nominative means that it can be generalised to other agreement targets. That masculine-personal in the plural nominative does not follow the ‘dependent target gender’ criterion does not in itself argue against it as a subgender. Most crucially, however, we find that there is a separate realization on the past tense of verbs in the plural. This is illustrated below where examples (41a–c) give the non-masculine-personal forms, and (41 d–f) the masculine-personal.¹⁴

- (41) a. *dziękował-y-śmy*
 thank.PAST-PL.NON_MASC_PERS-1ST.PL
 ‘We (not male human) thanked.’
 b. *dziękował-y-ście*
 thank.PAST-PL.NON_MASC_PERS-1ST.PL
 ‘You (not male human) thanked.’

¹⁴ Together with the writing system the glosses simplify matters somewhat, as the masculine personal feature is not really realised by <i> but by the consonant mutation in the stem. Linearly, therefore, the realization of masculine personal comes before that of plural number. It should be reiterated that our CDCs are about the presence of a distinction in one category triggered or conditioned by a feature of another category. The masculine-personal distinction is triggered by the plural, even if it is realized linearly before it.

- c. *dziękował-y*
thank.PAST-PL.NON_MASC_PERS
'They (not male human) thanked.'
- d. *dziękował-i-śmy*
thank.PAST-PL.MASC_PERS-1ST.PL
'We (male human) thanked.'
- e. *dziękował-i-ście*
thank.PAST-PL.MASC_PERS-2ND.PL
'You (male human) thanked.'
- f. *dziękował-i*
thank.PAST-PL.MASC_PERS
'They (male human) thanked.'

Normally, only nouns in the nominative case control verbal agreement. As we have seen in Table 6 the plural nominative and plural accusative may differ in terms of their behaviour as regards the masculine-personal gender. We claim that the difference in behaviour is to be expected, because masculine-personal is a fully blown structured gender which is found in the plural nominative. The accusative-genitive syncretism in the plural in Table 6 is to be accounted for, because there is a person subgender which is triggered by the accusative case in the plural. Furthermore, the accusative case in Polish is determined by reference to gender and subgender. The difference between a 'devirilized' noun and a run-of-the-mill masculine-personal noun, as shown in Table 6, is that the former has its syntactic gender lexically specified as masculine, whereas the latter has it assigned by a default rule, as we shall show. The accusative-genitive syncretism comes about because the subgender *feature* 'person' combines with the gender *feature* 'masculine' in the devirilized examples. In contrast, the ordinary masculine-personal nouns also have the accusative-genitive syncretism because their structured gender *features* are identical with the gender and subgender combination of the devirilized nouns.

5. Polish gender in Network Morphology

Polish adjectives obey CDCs (1), (8) and (9). This is illustrated in (42). In the *fact* [15] we see that gender is not specified for the realization of the singular genitive *-ego*. This is because neutralization of the masculine and neuter gender is conditioned by singular number and the oblique cases, in accordance with CDCs (1) and (9). Only for feminine realizations is gender specified in the oblique cases of the singular.

(42) MOR_ADJ:

<> == MOR_NOMINAL	%[10]
<mor sg nom masc> == "<stem>" _i	%[11]
<mor sg nom fem> == N_II	%[12]
<mor sg nom neuter> == "<stem>" _e	%[13]

```

<mor sg acc fem> == "<stem>" _o&                                %[14]
<mor sg gen> == "<stem>" _ego                                     %[15]
...

```

The singular and plural accusative are specified at the *node* MOR_NOMINAL, as illustrated in (43).

```

(43) MOR_NOMINAL:
    <mor case> == nom                                             %[3]
    <mor case masc person> == gen                                 %[4]
    <mor sg acc> ==
    ACCUSATIVE:< sg "<syn gender>" "<syn animacy>" >           %[5]
    <mor pl acc> ==
    ACCUSATIVE:< pl "<syn gender>" "<syn animacy>" >           %[6]
    ...

```

We shall return to *facts* [3] and [4] later on. They are both used in our treatment of the accusative. *Fact* [3] basically states that the default morphological case is nominative. That is, if you require a realization of the accusative case, use the form of the nominative, if no other is specified. *Fact* [4] states that the morphological case for nouns which have masculine personal gender is the genitive. So the genitive should be used to realise the accusative if the noun is masculine personal gender, or masculine gender and person subgender. If a noun is accusative-nominative syncretic, or accusative-genitive syncretic, then so will be the adjective which agrees with it. *Facts* [5] and [6] are therefore located at the MOR_NOMINAL *node* from which MOR_NOUN, the *node* for noun morphology, inherits.

The accusative singular or plural depends on information regarding syntactic animacy and gender. This last piece of information enables us to account for nouns of the *tajdaki* type we have seen earlier. The hierarchy for nominals looks like Fig. 5.

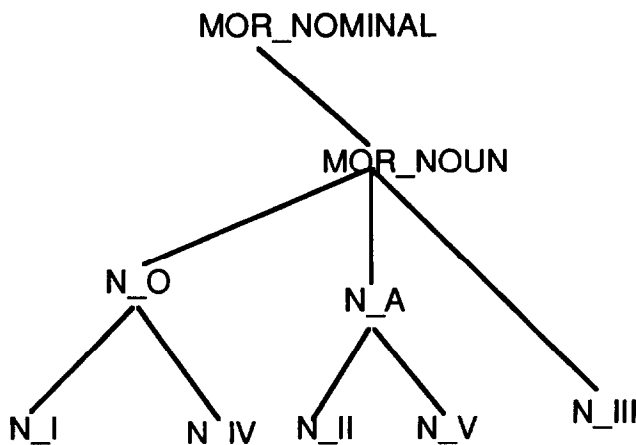


Fig. 5. Partial morphological hierarchy of Polish.

The noun declensions in Fig. 5 specify only three ‘morphological’ genders, namely masculine, feminine and neuter. If there were only one gender per inflectional class and one inflectional class per gender, then we would expect the number of genders to be the same as the number of inflectional classes, namely five. However, as there are three declensions which all share the same gender, the number of genders assigned morphologically is only three.

We have to ask how gender is assigned over and above the morphological assignment system. If there are only three possible *values* for gender assigned morphologically, how then do we account for the data in (23)–(32) where we introduced the different agreement forms for masculine-personals as opposed to other nouns?

As an approximation it is claimed that Polish has two elements to the semantic assignment of its main genders. Animacy is a peripheral element, and biological sex is the core which cannot be overridden. ‘Devirilization’ involves either circumventing the peripheral animacy part of main gender assignment (variant A), or taking the exceptional case default for the assignment of syntactic animacy (variant B). In fact, in each instance we resort to the idea of the exceptional case defaults argued for by Fraser and Corbett (1997) for Arapesh. The important point is that one is not permitted to override just any information and the system proposed sees gender assignment according to biological sex as inviolable. In sum, variant A ‘devirilized’ nouns are changed from masculine-personal gender to masculine gender, because the animacy part is circumvented. However, they still have person subgender and therefore still have genitive-accusative syncretism. The less common variant B ‘devirilized’ nouns alter their subgender from person to animate. This means they lose the special form in the plural nominative and have nominative-accusative syncretism, but keep genitive-accusative syncretism in the singular because they are animate. These exceptional case defaults are the only options open for any noun.

Polish assigns the *values* *animate*, *person* or *inanimate* for the subgender of animacy on the basis of sex differentiation. If a noun is sex-differentiable, then it will be animate. This is stated by *fact* [81] in (44). The reason why [81] takes the form that it does is that any evaluation of the sex of the referent, will by default yield a *value* ‘undifferentiated’ which the noun in question inherits by default. Hence, any noun which is not sex-differentiable will match with *fact* [83], which states that it will be inanimate. By default only nouns with male referents are assigned the *value* ‘person’, as in *fact* [82]. The effect is that nouns with female referents are animate, from [81], and nouns with male referents are persons, from [82]. An important consequence of this is that the default *value* (that is, fully unspecified) for animacy is ‘animate’, and ‘animate’ is the exceptional case default, which nouns can refer to in the last resort. Note that ‘animate’ is the unspecified option. We claim that the unspecified option is what is open to a lexical entry in the last resort.

(44) ANIMACY:

<> == animate	%[81]
<male> == person	%[82]
<undifferentiated> == inanimate.	%[83]

Non-sex-differentiable animals stipulate their semantic animacy in the lexical entry. The lexical entries of nouns with female humans as referents must stipulate both their biological sex and that they are persons. We argued that subgender information about being a person is also required for nouns with female persons as referents on the basis of examples such as (36) where this information may be used for agreement with a conjunct.

Information about syntactic animacy is inherited from the NOUN *node* (45) in the lexemic hierarchy. Fact [79] states that syntactic animacy is by default the same as semantic animacy. Semantic animacy is determined by evaluating biological sex and referring to the appropriate *paths* at the ANIMACY *node* in (44).

(45) NOUN:

<> == NOMINAL	%[74]
<declensional_class> == DECLENSION:< "<sem sex>" >	%[75]
<syn cat> == n	%[76]
<sem sex> == undifferentiated	%[77]
<syn gender> == GENDER:< "<syn animacy>" >	%[78]
<syn animacy> == "<sem animacy>"	%[79]
<sem animacy> == ANIMACY:< "<sem sex>" >.	%[80]

Fact [78] in (45) states that one must evaluate syntactic animacy and find *values* which match the evaluated *paths* at the *node* GENDER. It is this part of the main gender assignment system which could be circumvented by a 'devirilized' noun, which could instead refer directly to the GENDER *node*. The GENDER *node* is given in (46).

(46) GENDER:

<> == < "<sem sex>" >	%[94]
<undifferentiated> == "<mor gender>"	%[95]
<male> == masc	%[96]
<female> == fem	%[97]
<male person> == <male> person.	%[98]

What is interesting about the GENDER *node* is that there is no *path* which begins with any of the animacy *features* evaluated in (45), *fact* [78]. This means that the *path* containing an evaluated animacy *feature* matches with the empty *path* in [94]. Here it is stated that one must evaluate biological sex of the referent to determine gender. As the evaluated animacy *feature* is an extension of the empty *path* in [94], so it will be an extension of the *path* containing the evaluated *features* for biological sex. This means that any noun with male human referent will evaluate to [98] and be assigned the structured gender 'masc person.' Otherwise, any noun that has a male referent will be assigned masculine gender by [96], and any noun that has a female referent will be assigned feminine gender by [97]. All other nouns will be assigned gender according to their inflectional class, which is what is stated in [95]. It should also be noted that [94] will be the only option for any noun which specifies syntactic gender lexically, as we propose that the last resort for any lexical item can only be the unspecified option.

In (47) we see the lexical entry for an animate masculine noun.¹⁵

(47) *Bik*:

<> == NOUN	%[186]
<gloss> == bull	%[187]
<root> == bi	%[188]
<root final> == K	%[189]
<sem sex> == male	%[190]
<sem animacy> == animate.	%[191]

The noun in (47) *byk* ‘bull’ is assigned animate subgender, because it is semantically animate, and *fact* [79] in (45) says that syntactic animacy is by default the same as semantic animacy. Gender is assigned to (47) initially by evaluating its syntactic animacy, as required by [78]. The evaluated *path* then matches with one at the GENDER *node*. This will be the empty *path* [94] in (46) which requires an evaluation of biological sex, which entails that *byk* is assigned masculine gender.

In (48) we see the lexical entry for a masculine personal noun.

(48) *Student*:

<> == NOUN	%[228]
<gloss> == student	%[229]
<root> == studen	%[230]
<root final> == T	%[231]
<sem sex> == male.	%[232]

The noun in (48) *student* ‘student’ is assigned the *value* person for semantic animacy by default. This follows from *fact* [82] in (44), which states that males are by default persons. From this it also follows that *student* will have the *value* person for the animacy subgender, because of [79] in (45). Again, the noun is assigned syntactic gender by evaluating its syntactic animacy, as stated at [78] in (45). As no *path* at the *node* GENDER begins with the *feature* person, syntactic gender is assigned by the least specific *path* [94] in (46). As the person *feature* extends the empty *path* on the left-hand side of [94], so it will extend the evaluated *path* on the right-hand side of [94]. This means that syntactic gender for *student* will be assigned by the *fact* which contains the correct *feature* for biological sex and its extension for animacy, namely [98] in (46). In this way *student* is assigned the structured syntactic gender **masc person**. In sum, the noun *student* is assigned a structured gender masculine-personal and a subgender person.

On the basis of the interdependency *node* in (49) the correct forms for the singular and plural accusative of *byk* and *student* are determined. Note that this *node* is referenced by *facts* [5] and [6] at the MOR_NOMINAL *node* in the morphological hierarchy, given in (43).

¹⁵ The examples are given in a lower ASCII phonological transcription.

(49) ACCUSATIVE:

<sg> == "<mor sg "<mor case>" >"	%[91]
<pl> == "<mor pl "<mor case>" >"	%[92]
<sg masc animate> == <sg masc person>.	%[93]

In *facts* [91] and [92] at (49) it is stated that the accusative of either number depends on a particular 'morphological case' for its form. We have seen these morphological cases earlier in (43). Hence, if something has either masculine personal gender (**masc person**), or masculine gender (**masc**) and person subgender (**person**), its accusative will be realised by the form of the genitive case, as stated by *fact* [4] in (43). This is because the *values* for syntactic gender and animacy evaluated by [5] and [6] in (43) will extend the *paths* of [91] and [92] in (49). Accordingly, they will also extend the evaluated *path* <mor case> in [91] and [92]. The *values* for the extensions are provided by [3] and [4] in (43). If the noun is anything other than one with masculine personal gender, **masc person**, then it will match with [3], and [91] states that the realization of accusative will be the form of the nominative case for that number. If the noun is masculine personal, then it will match with [4] in (43), and [91] states that the realization of accusative will be the form of the genitive case for that number. In addition to this, [93] states that masculine animates will follow masculine person nouns in the singular, which accounts for examples such as (24). That is, they will have the genitive case form, as we have just shown. Note that in all instances the *feature* specifications in (49) obey the CDCs (1), (8) and (9).

In (50) and (51) we give the derivable forms for the nominative, accusative and genitive singular and plural of the nouns *student* and *byk*, together with agreeing adjective, and contrast them with those for an inanimate masculine noun in (52).¹⁶

(50) Student:

<gloss> = student
<syn sg nom / adj noun> = [no _v _i [studen _t]]
<syn sg acc / adj noun> = [no _v _ego [studen _t _a]]
<syn sg gen / adj noun> = [no _v _ego [studen _t _a]]
<syn pl nom / adj noun> = [no _v ' _i [studen _t _s ' _i]]
<syn pl acc / adj noun> = [no _v _i _x [studen _t _uv]]
<syn pl gen / adj noun> = [no _v _i _x [studen _t _uv]]

(51) Bik:

<gloss> = bull
<syn sg nom / adj noun> = [no _v _i [bi _k]]
<syn sg acc / adj noun> = [no _v _ego [bi _k _a]]

¹⁶ Note that our fragment also includes consonantal and vocalic alternations to be found in noun stems. The consonantal alternations are important, as they play a role in the realization of masculine-personal gender. We have generated the noun phrases in (48) by the use of a 'dummy syntax' noun phrase *node* in the DATR fragment. We do not make any theoretical claims on the basis of this part of the fragment, other than that there must be some rule of syntax which makes use of animacy features in the generation of accusative noun phrases. The main aim is to demonstrate that morphology can provide the appropriate forms for syntax.

<syn sg gen / adj noun> = [no _v _ego [bi _k _a]]
 <syn pl nom / adj noun> = [no _v _e [bi _k _i]]
 <syn pl acc / adj noun> = [no _v _e [bi _k _i]]
 <syn pl gen / adj noun> = [no _v _i _x [bi _k _uv]]

(52) Grod:

<gloss> = castle
 <syn sg nom / adj noun> = [no _v _i [gr _u _d]]
 <syn sg acc / adj noun> = [no _v _i [gr _u _d]]
 <syn sg gen / adj noun> = [no _v _ego [gr _o _d _u]]
 <syn pl nom / adj noun> = [no _v _e [gr _o _d _i]]
 <syn pl acc / adj noun> = [no _v _e [gr _o _d _i]]
 <syn pl gen / adj noun> = [no _v _i _x [gr _o _d _uv]]

The forms in (50)–(52) demonstrate that we are able to capture the syntcretisms of the kind that we have illustrated in (23)–(32). It now remains to show that the more marginal ‘devirilized’ cases can also be accommodated.

In order to account for examples such as (36), where a feminine noun denoting a female human being and an animate masculine noun are conjoined, we must demonstrate that the syntactic *features* are available to do this. According to Corbett (1991: 286) the following generalizations may account for gender resolution of this kind:

- (53) (i) masculine-personal conjunct in subject, predicate has masculine-personal agreement
 (ii) (optional) subject includes *features* masculine and personal, whether semantic or syntactic, predicate may have masculine-personal agreement
 (iii) (optional) subject includes a masculine animate conjunct, the predicate may have masculine-personal agreement
 (iv) otherwise the predicate will have non-masculine-personal agreement
 (Corbett, 1991: 286)

Clause (ii) of (53) states that the *features* can be either semantic or syntactic. If we consider a feminine noun such as *žona* ‘wife’ in (15), it should be noted that *fact* [255] specifies that the noun has a person referent. This is achieved by saying that the *value* of its semantic animacy is **person**. As there is also a statement [79] in (45) that syntactic animacy is the same as semantic animacy, the noun *žona* ‘wife’ has a *value* for the subgender of animacy, namely **person**, even though this is rarely used. It also means that we could modify (ii) such that reference only need be made to syntactic *features*.

Not only is animacy and person information required for conjunct resolution, but it is also required to account for ‘devirilized’ nouns. These are specified as such lexically. However, as we have stated, they are only allowed to refer to the GENDER *node* for the unspecified option, namely the evaluation of biological sex [94] in (46). We thereby capture the generalization that even devirilized nouns obey semantic assignment of gender, because they are assigned masculine gender as males. Only the peripheral animacy assignment is overridden. Masculine nouns such as

mężczyzna ‘man’, which belong to declension class II that normally assigns feminine gender, also fit into this account, because they could in theory only take the unspecified option if gender assignment is to be overridden. This would mean that in the worst case such nouns would be assigned masculine rather than masculine personal gender. They cannot, however, be assigned the gender of their inflectional class.

For the lexical entry for the variant A ‘devirilized’ noun *tajdak* ‘scoundrel’ in (54), *fact* [238] specifies as a last resort that syntactic gender of the noun is whatever the default is at the *node* GENDER. Recall that the default states that the sex of the referent must be evaluated, which means that it will be assigned masculine gender, as opposed to usual masculine personal where syntactic animacy is taken into account.

(54) *Wajdak*_1:

<> == NOUN	%[233]
<gloss> == scoundrel	%[234]
<root> == wajda	%[235]
<root final> == K	%[236]
<sem sex> == male	%[237]
<syn gender> == GENDER.	%[238]

It should be noted that (54) will still be assigned a *value* person for syntactic animacy, it is just that this *value* is not incorporated into the main gender.

As there may be variation in the ‘devirilized’ status of nouns such as *tajdak*, elimination of the last line of the entry would make it behave exactly as any other masculine-personal noun. As this noun will still also be assigned the subgender ‘person’, it will still have genitive-accusative syncretism in the plural.

The result of *fact* [238] in (54) is that agreeing elements such as adjectives or verbs will have masculine agreement rather than masculine-personal agreement. For example, the nominative plural of adjectives is determined by using this information. An adjectival *node* inherits from the *node* MOR_NOMINAL and, among other things, provides information about the agreement form of the adjective. The relevant information in the adjectival *node* for nominative plural agreement is as in (55).

(55) MOR_ADJ:

<> == MOR_NOMINAL	%[10]
<mor pl nom> == PL_NOM2:<>	%[23]
...	

The *node* PL_NOM2, referred to by *fact* [23], makes use of syntactic gender only (not subgender) to determine the form of the plural nominative. Masculine-personal adjectives will have stem mutation plus the ending *-i*, otherwise the adjective will have the ending *-e*. Any noun which does not have masculine-personal gender, including the ‘devirilized’ nouns, will take an adjective ending in *-e*. The correct syncretism is obtained for the accusative plural, with both masculine-personal and devirilized nouns having accusative-genitive syncretism.

All that is required to account for the less common ‘devirilization’ in which nominative-accusative syncretism in the plural accompanies loss of the special plural nominative form and agreements, is for a lexical specification that the noun in question takes the (unspecified) default from the *node* ANIMACY, as in (56).

(56) Wajdak_2:

<> == NOUN	%[239]
<gloss> == scoundrel	%[240]
<root> == wajda	%[241]
<root final> == K	%[242]
<sem sex> == male	%[243]
<syn animacy> == ANIMACY.	%[244]

In (56) the variant B ‘devirilized’ noun *tajdak* ‘scoundrel’ can only retrieve the default for animacy as a last resort, namely animate. This variant, which is not accepted by the majority of speakers, can only arise by this means. Most importantly, we also capture generalisation (40a), which states that nominative-accusative syncretism may only arise, if there is the non-masculine-personal form for the plural nominative. As syntactic animacy for this noun is animate, this means that the noun will be assigned masculine gender, which therefore entails loss of the special form in the plural nominative. As it is neither masculine-personal, nor does it have person subgender, it will not have genitive-accusative syncretism in the plural. It will still have genitive-accusative syncretism in the singular, however, because it is an animate noun. We therefore obey all of the generalisations in (40).

6. Conclusion

The claim of this article is that Polish does have subgender, as other Slavonic languages such as Russian. Subgender in Polish is more differentiated from that of Russian, as it distinguishes between animates, persons and inanimates. However, it is also claimed that Polish has a main gender of masculine-personal. This structured gender is what is found in the plural nominative. The fact that accusative-genitive syncretism and the masculine-personal realization in the plural nominative need not always go together, as is true of the ‘devirilized’ nouns, shows that these two phenomena are not to be accounted for in exactly the same way. Accusative-genitive syncretism occurs for masculine-personals (gender) nouns or for masculine nouns which have person subgender. In the default instance any noun for a male human will belong to the masculine-personal gender. ‘Devirilized’ nouns obtain the exceptional case default for syntactic gender (gender assignment according to sex of referent only), when they would otherwise be assigned masculine-personal gender by default. This means that they do not have the plural nominative form, but do have the syncretism, because they still have the *value* ‘person’ assigned for the subgender of animacy, and in the accusative case both gender and subgender are evaluated. Masculine-personal nouns have the accusative-genitive syncretism, because the sec-

ond element of the structured gender ‘masculine-personal’ is the same as the sub-gender *value*.

Other important claims arise from the theory outlined here. Subgender appears to be intimately connected with case, because the presence of case *features* mean that it occurs only in a minimal subset of the paradigm. As the verbal forms indicate, the minimal subset criterion rules out masculine-personal as a subgender for Polish when the verbal system is taken into account. Our formal approach claims that masculine-personal gender could arise from the fusion of the two autonomous gender and subgender *features*. The crucial point is that in Polish, unlike Russian for example, the subgender *feature* spreads to the nominative case. Once this occurs it is free to spread to other agreement targets that do not realise case, such as the past tense of verbs in Polish. At this point we are no longer dealing with a gender and subgender, but a new gender, masculine-personal. This probably also means that the ‘dependent target gender’ criterion is more important than Corbett claims, because we would not expect the nominative case to refer to other cases for its realization, as we know this to be the primary case distinction (Blake 1994). Hence, any distinction in the singular or plural nominative must involve its own realization. If it involves its own realization, this means that the distinction is morphologized and that it can spread to other agreement targets. The creation of a structured gender should therefore be compared with the ‘multiple gender’ systems found in languages such as those of the Mba group, where what we have in the structured gender is something equivalent to a fusional variant of a ‘multiple gender’ system. It can be seen that our analysis not only has the advantage of being explicit by virtue of its computational modelling, but also makes specific predictions about subgenders.

Appendix

%
%
% File: polish15.dtr %
% Purpose: polish nouns %
% Author: Dunstan Brown, March 26, 1996 %
% Email: d.brown@surrey.ac.uk %
% Address: LIS, University of Surrey, Guildford GU2 5XH %
% Documentation: %
% Related Files: %
% Version: 1.13 (September 22, 1997) %
%
% %

% Example theorems from this fragment can be viewed at
% <http://www.surrey.ac.uk/LIS/SMG/theorems.html>

% This fragment covers the nominal morphology of Polish. It includes the
% vocalic and consonantal alternations which occur in the nominal morphol-

% ogy. A lower ASCII transcription has been used, with following correspondences between it and the orthography (where C stands for any consonant in the transcription or the orthography):

% Transcription	Writing
% Ci	Cy
% C'	Ci or Ó
% ts	c
% ts'	ci or ć
% w	†
% zz	ž

% Where the morphophonemics are concerned it should be noted that we have written *ć* as /ts'/, as this alternates with dental /t/ and the alternation is one of softening and affrication. Elsewhere, in roots for example, we have not transcribed *ć* as /ts'/, but as /c'/, although it would be trivial to change this. The same is also true of the alternation between /k/ and /ts/ in /wajdatsi/. The alternant with /k/ is represented as /ts/, whereas the written <c> in *noc* is transcribed as /c/, rather than /ts/. Again, it would be trivial to alter this, but we have kept the roots similar to the writing system for ease of recognition. The other transcription of the alternations is to make the true underlying processes modelled by our morphological section more perspicuous.

%
 %
 % MORPHOLOGICAL HIERARCHY
 %
 %

MOR_NOMINAL:

<> == "<stem closed>" % [1]
 <mor hardness> == MOR_HARDNESS:< "<stem final type>" > % [2]
 <mor case> == nom % [3]
 <mor case masc person> == gen % [4]
 <mor sg acc> ==
 ACCUSATIVE:< sg "<syn gender>" "<syn animacy>" > % [5]
 <mor pl acc> ==
 ACCUSATIVE:< pl "<syn gender>" "<syn animacy>" > % [6]
 <mor pl inst> == "<stem>" "<mor vowel>" _m'i % [7]
 <mor pl loc> == "<stem>" "<mor vowel>" _x % [8]
 <mor pl gen> == "<mor pl loc>". % [9]

MOR_ADJ:

<> == MOR_NOMINAL % [10]

<mor sg nom masc> == "<stem>" _i	%[11]
<mor sg nom fem> == N_ll	%[12]
<mor sg nom neuter> == "<stem>" _e	%[13]
<mor sg acc fem> == "<stem>" _o&	%[14]
<mor sg gen> == "<stem>" _ego	%[15]
<mor sg gen fem> == "<stem>" _ej	%[16]
<mor sg dat> == "<stem>" _emu	%[17]
<mor sg dat fem> == "<mor sg gen fem>"	%[18]
<mor sg inst> == "<stem>" _im	%[19]
<mor sg inst fem> == N_ll	%[20]
<mor sg loc> == "<mor sg inst>"	%[21]
<mor sg loc fem> == "<mor sg dat fem>"	%[22]
<mor pl nom> == PL_NOM2:<>	%[23]
<mor pl dat> == "<stem>" _im	%[24]
<mor vowel> == _i.	%[25]

MOR_NOUN:

<> == MOR_NOMINAL	%[26]
<mor sg voc> == "<mor sg loc>"	%[27]
<mor sg gen> == "<stem>" _i	%[28]
<mor sg dat> == "<mor sg loc>"	%[29]
<mor sg loc> == "<stem i>" _e	%[30]
<mor pl voc> == "<mor pl nom>"	%[31]
<mor pl nom> == PL_NOM1:< "<mor hardness>" "<syn gender>" >	%[32]
<mor pl gen> == MGP:< "<mor hardness>" >	%[33]
<mor pl dat> == "<stem>" _om	%[34]
<mor vowel> == _a.	%[35]

N_O:

<> == MOR_NOUN	%[36]
<mor sg gen> == GENITIVE:< "<syn gender>" "<syn animacy>" >	%[37]
<mor sg inst> == "<stem>" _em	%[38]
<mor sg loc> == SG_LOC:< "<stem final type>" >.	%[39]

N_I:

<> == N_O	%[40]
<mor pl gen> == "<stem>" _uv	%[41]
<mor sg dat> == "<stem>" _ov'i	%[42]
<mor gender> == masc.	%[43]

N_IV:

<> == N_O	%[44]
<mor sg nom> == "<stem>" _o	%[45]
<mor sg voc> == "<mor sg nom>"	%[46]
<mor sg dat> == "<stem>" _u	%[47]

<mor pl nom> == "<stem>" _a % [48]
 <mor gender> == neuter. % [49]

N_A:

<> == MOR_NOUN % [50]
 <mor sg acc> == "<stem>" _e& % [51]
 <mor sg inst> == "<stem>" _o& % [52]
 <mor gender> == fem. % [53]

N_II:

<> == N_A % [54]
 <mor sg nom> == "<stem>" _a % [55]
 <mor sg voc> == "<stem>" _o. % [56]

N_III:

<> == MOR_NOUN % [57]
 <mor hardness> == soft % [58]
 <mor sg loc> == "<mor sg gen>" % [59]
 <mor sg inst> == N_II % [60]
 <mor gender> == fem. % [61]

N_V:

<> == N_A % [62]
 <mor hardness> == soft % [63]
 <mor sg nom> == "<mor sg gen>" % [64]
 <mor sg loc> == N_III % [65]
 <mor pl gen> == "<stem>". % [66]

%
 % %
 % LEXEMIC HIERARCHY %
 % %
 %

NOMINAL:

<> == undefined % [67]
 <syn> == SYNTAX % [68]
 <mor> == "<declensional_class>" % [69]
 <stem final type> == "<root final type>" % [70]
 <stem> == "<root>" "<root final shape>". % [71]

ADJ:

<> == NOMINAL % [72]
 <declensional_class> == MOR_ADJ:<mor>. % [73]

```

<> == NOMINAL % [74]
<declensional_class> == DECLENSION:< "<sem sex>" > % [75]
<syn cat> == n % [76]
<sem sex> == undifferentiated % [77]
<syn gender> == GENDER:< "<syn animacy>" > % [78]
<syn animacy> == "<sem animacy>" % [79]
<sem animacy> == ANIMACY:< "<sem sex>" >. % [80]

```

[illegible]

```
<> == animate %[81]
<male> == person %[82]
<undifferentiated> == inanimate. %[83]
```

```
<male> == N_I:<mor> % [84]
<female> == N_II:<mor>. % [85]
```

```

<> == "<stem>"_i % [86]
<soft> == PL_NOM2 % [87]
<hard masc person> == PL_NOM2:<masc person>. % [88]

```

```
<> == "<stem>"_e % [89]
<masc person> == "<stem i>"_i. % [90]
```

```
<sg> == "<mor sg "<mor case>" ">" % [91]
<pl> == "<mor pl "<mor case>" ">" % [92]
<sg masc animate> == <sg masc person>. % [93]
```

```
<> == < "sem sex" > %[94]
<undifferentiated> == "<mor gender>" %[95]
<male> == masc %[96]
<female> == fem %[97]
<male person> == <male> person. %[98]
```

GENITIVE:

<> == "<stem>" _a % [99]
 <masc inanimate> == N_IV:<mor sg dat>. % [100]

SG_LOC:

<> == MOR_NOUN:<mor sg loc> % [101]
 <velar> == N_IV:<mor sg dat>. % [102]

MGP:

<> == "<stem>" _i % [103]
 <hard> == "<mor>". % [104]

MOR_HARDNESS:

<> == hard % [105]
 <soft consonant> == soft. % [106]

%
 % %
 % MORPHONOLOGY %
 % %
 %

CONS:

<> == % [107]
 <root final type> == hard consonant % [108]
 <root final grade i> == _' . % [109]

STOP:

<> == CONS % [110]
 <root final shape> ==
 "<root final stop>"
 "<root final affric>" "<root final grade>". % [111]

FRIC:

<> == CONS % [112]
 <root final shape> ==
 "<root final fric>" "<root final grade>". % [113]

I_II_STOP:

<> == STOP % [114]
 <root final grade i i> == . % [115]

I_II_FRIC:

<> == FRIC % [116]

<root final grade i i> == _z.	%[117]
FRIC_STOP:	
<> == CONS	%[118]
<root final shape> == FRIC_STOP.	%[119]
I_II_FRIC_STOP:	
<> == FRIC_STOP	%[120]
<root final grade i i> == I_II_FRIC.	%[121]
T:	
<> == I_II_STOP	%[122]
<root final stop> == _t	%[123]
<root final affric i> == S:<root final fric i>.	%[124]
D:	
<> == I_II_STOP	%[125]
<root final stop> == _d	%[126]
<root final affric i> == Z:<root final fric i>.	%[127]
S:	
<> == I_II_FRIC	%[128]
<root final fric> == _s.	%[129]
Z:	
<> == I_II_FRIC	%[130]
<root final fric> == _z.	%[131]
ST:	
<> == I_II_FRIC_STOP	%[132]
<root final stop> == T	%[133]
<root final affric> == T	%[134]
<root final fric> == s.	%[135]
ZD:	
<> == I_II_FRIC_STOP	%[136]
<root final stop> == D	%[137]
<root final affric> == D	%[138]
<root final fric> == z.	%[139]
P:	
<> == STOP	%[140]
<root final stop> == _p.	%[141]

B:	
<> == STOP	%[142]
<root final stop> == _b.	%[143]
F:	
<> == FRIC	%[144]
<root final fric> == _f.	%[145]
V:	
<> == FRIC	%[146]
<root final fric> == _v.	%[147]
M:	
<> == STOP	%[148]
<root final stop> == _m.	%[149]
N:	
<> == STOP	%[150]
<root final stop> == _n.	%[151]
W:	
<> == CONS	%[152]
<root final> == _w	%[153]
<root final shape i> == _l.	%[154]
R:	
<> == CONS	%[155]
<root final> == _r	%[156]
<root final shape i> == _zz.	%[157]
X:	
<> == CONS	%[158]
<root final> == _x	%[159]
<root final type> == velar consonant	%[160]
<root final shape i> == _sz.	%[161]
K:	
<> == I__II_STOP	%[162]
<root final type> == velar consonant	%[163]
<root final stop> == _k	%[164]
<root final shape i> == _ts	%[165]
<root final shape i i> == _cz.	%[166]
G:	
<> == I__II_STOP	%[167]

<root final type> == velar consonant	%[168]
<root final stop> == _g	%[169]
<root final shape i> == _dz	%[170]
<root final shape i i> == _zz.	%[171]

SK:

<> == I__II_FRIC_STOP	%[172]
<root final type> == velar consonant	%[173]
<root final shape i> == _s K	%[174]
<root final shape i i> == x K.	%[175]

ZG:

<> == I__II_FRIC_STOP	%[176]
<root final type> == velar consonant	%[177]
<root final shape i> == _z G	%[178]
<root final shape i i> == _zzdzz.	%[179]

% %	%
%	%
%	VOWELS
%	%
% %	%

A_E:

<root vowel> == _a	%[180]
<root vowel i> == _e.	%[181]

O_U:

<root vowel> == _o	%[182]
<root vowel closed> == _u.	%[183]

E:

<root vowel> ==	%[184]
<root vowel closed> == _' _e.	%[185]

% %	%
%	%
%	LEXICAL ENTRIES
%	%
% %	%

Bik:

<> == NOUN	%[186]
<gloss> == bull	%[187]

<root> == bi	%[188]
<root final> == K	%[189]
<sem sex> == male	%[190]
<sem animacy> == animate.	%[191]

Bogin'i:

<> == NOUN	%[192]
<gloss> == goddess	%[193]
<declensional_class> == N_V:<mor>	%[194]
<root> == bogi	%[195]
<root final> == n'	%[196]
<sem sex> == female	%[197]
<sem animacy> == person.	%[198]

Grod:

<> == NOUN	%[199]
<gloss> == castle	%[200]
<declensional_class> == N_I:<mor>	%[201]
<root> == gr O_U:<root vowel>	%[202]
<root final> == D.	%[203]

Kos'c':

<> == NOUN	%[204]
<declensional_class> == N_III:<mor>	%[205]
<mor pl inst> == "<stem>" _m'i	%[206]
<mor pl nom> == PL_NOM1	%[207]
<gloss> == bone	%[208]
<root> == kos'	%[209]
<root final> == c'.	%[210]

M'asto:

<> == NOUN	%[211]
<gloss> == city	%[212]
<declensional_class> == N_IV:<mor>	%[213]
<root> == m' A_E:<root vowel>	%[214]
<root final> == ST.	%[215]

Noc:

<> == NOUN	%[216]
<gloss> == night	%[217]
<declensional_class> == N_III:<mor>	%[218]
<root> == no	%[219]
<root final> == c.	%[220]

P'es:

<> == NOUN	%[221]
<gloss> == dog	%[222]
<root> == p E:<root vowel>	%[223]
<root final> == S	%[224]
<mor sg dat> == N_IV %-u rather than -ov'i	%[225]
<sem sex> == male	%[226]
<sem animacy> == animate.	%[227]

Student:

<> == NOUN	%[228]
<gloss> == student	%[229]
<root> == studen	%[230]
<root final> == T	%[231]
<sem sex> == male.	%[232]

% Wajdak_1 takes the unspecified option for gender, which means that it
 % will be assigned masculine gender rather than masculine-personal. It has
 % gen-acc syncretism, but no special form for the plural nominative.

Wajdak_1:

<> == NOUN	%[233]
<gloss> == scoundrel	%[234]
<root> == wajda	%[235]
<root final> == K	%[236]
<sem sex> == male	%[237]
<syn gender> == GENDER.	%[238]

% Wajdak_2 takes the unspecified option for animacy, which means that it
 % will be assigned animate subgender rather than person. It has nom-acc
 % syncretism in the plural and no special form for the plural nominative. It
 % still has gen-acc syncretism in the singular, of course.

Wajdak_2:

<> == NOUN	%[239]
<gloss> == scoundrel	%[240]
<root> == wajda	%[241]
<root final> == K	%[242]
<sem sex> == male	%[243]
<syn animacy> == ANIMACY.	%[244]

% Wajdak_3 is a straightforward masculine-personal noun, with both gen-
 % acc syncretism and the special form for the plural nominative.

Wajdak_3:

<> == NOUN	%[245]
------------	--------

<gloss> == scoundrel % [246]
 <root> == wajda % [247]
 <root final> == K % [248]
 <sem sex> == male. % [249]

Zzona:

<> == NOUN % [250]
 <gloss> == wife % [251]
 <root> == zzo % [252]
 <root final> == N % [253]
 <sem sex> == female % [254]
 <sem animacy> == person. % [255]

%
 %
 % DUMMY SYNTAX %
 %
 %
 % These number and case variables are used for the 'dummy' syntax node
 % which generates noun phrases. No theoretical claims are made on the
 % basis of this node other than that accusative noun phrases require infor-
 % mation about the subgender of animacy.

vars \$number: sg pl.
 # vars \$case: nom voc acc gen dat inst loc.

SYNTAX:

<> == % [256]
 <syn form> == "<mor>" % [257]
 <syn head> == "<syn>"
 <syn \$number \$case / noun> ==
 [<syn form \$number \$case>] % [258]
 <syn \$number \$case / adj noun> ==
 ["Novi:<syn form \$number \$case <syn head gender> >"
 <syn \$number \$case / noun>] % [259]
 <syn \$number acc / adj noun> ==
 ["Novi:<syn form \$number acc <syn head gender> <syn head animacy> >"
 <syn \$number acc / noun>]. % [260]

References

- Blake, B.J., 1994. *Case*. Cambridge: Cambridge University Press.
 Bleiching, D., G. Drexel, and D. Gibbon, 1996. Ein Synkretismusmodell für die deutsche Morphologie.
 In: D. Gibbon (ed.), *Natural language processing and speech technology. Results of the 3rd KONVENS
 Conference*, Bielefeld, 237–248. Berlin: Mouton de Gruyter.

- Brooks, M.Z. and K.L. Nalibow, 1970. The gender of referentials in Polish. *International Journal of Slavic Linguistics and Poetics* 13, 136–42.
- Cahill, L. and G. Gazdar, 1997. The inflectional phonology of German adjectives, determiners, and pronouns. *Linguistics* 35, 211–245.
- Carstairs, A., 1984. Outlines of a constraint on syncretism. *Folia Linguistica* 18, 73–85.
- Corbett, G.G., 1978. Universals in the syntax of cardinal numerals. *Lingua* 46, 355–68.
- Corbett, G.G., 1983a. The number of genders in Polish. *Papers and Studies in Contrastive Linguistics* VXI, 83–89. UAM: Poznań.
- Corbett, G.G., 1983b. Hierarchies, targets and controllers: Agreement patterns in Slavic. Beckenham: Croom Helm.
- Corbett, G.G., 1991. *Gender*. Cambridge: Cambridge University Press.
- Corbett, G.G. and N.M. Fraser, 1993. Network Morphology: A DATR account of Russian nominal inflection. *Journal of Linguistics* 29, 113–42.
- Czuba, K. and A. Przepiórkowski, 1995. Agreement and case assignment in Polish: An attempt at a unified account. Report 783, Polish Academy of Sciences Institute of Computer Science.
- Daelemans, W., K. De Smedt and G. Gazdar, 1992. Inheritance in natural language processing. *Computational Linguistics* 18, 205–18.
- Doroszewski, W., 1962. *O kulturę słowa: Poradnik językowy*. Warsaw: PIW.
- Evans, R. and G. Gazdar, 1989a. Inference in DATR. *Proceedings of the 4th Conference of the European Chapter of the Association for Computational Linguistics*, 66–71. Manchester, England.
- Evans, R. and G. Gazdar, 1989b. The semantics of DATR. In: A.G. Cohn (ed.), *Proceedings of the Seventh Conference of the Society for the Study of Artificial Intelligence and Simulation of Behaviour*, 79–87. London: Pitman/Kaufmann.
- Evans, R. and G. Gazdar, 1996. DATR: A language for lexical knowledge representation. *Computational Linguistics* 22, 167–216.
- Fahlman, S.E., 1979. Representing and using real-world knowledge. In: P.H. Winston and R.H. Brown (eds.), *Artificial intelligence: An MIT perspective*, Volume 1, 451–470. Cambridge, MA: MIT Press.
- Flickinger, D.P., 1987. *Lexical rules in the hierarchical lexicon*. Unpublished doctoral dissertation, Stanford University.
- Foley, W.A., 1986. *The Papuan languages of New Guinea*. Cambridge: Cambridge University Press.
- Foreman, V.M., 1974. *Grammar of Yessan-Mayo*. Santa Ana, CA: Summer Institute of Linguistics. (= Language Data, Asian-Pacific Series 4)
- Fraser, N.M. and G.G. Corbett, 1995. Gender, animacy and declensional class assignment: A unified account for Russian. In: G. Booij and J. van Marle (eds.), *Yearbook of Morphology 1994*, 123–150. Dordrecht: Kluwer.
- Fraser, N.M. and G.G. Corbett, 1997. Defaults in Arapesh. *Lingua* 103, 25–57.
- Fraser, N.M. and R.A. Hudson, 1992. Inheritance in Word Grammar. *Computational Linguistics* 18, 133–158.
- Greenberg, J.H., 1963. Some universals of grammar with particular reference to the order of meaningful elements. In: J.H. Greenberg (ed.), *Universals of language*, 73–113. Cambridge, MA: MIT Press. [Paperback edition published 1966; page references to this edition]
- Hudson, R.A., 1990. *English Word Grammar*. Oxford: Blackwell.
- Huntley, D., 1980. The evolution of genitive-accusative animate and personal nouns in Slavic dialects. In: J. Fisiak (ed.), *Historical morphology*, 189–212. The Hague: Mouton.
- Hurford, J.R. 1987. *Language and number: The emergence of a cognitive system*. Oxford: Blackwell.
- Keller, B., 1995. DATR theories and DATR models. *Proceedings of the 33rd Annual Meeting of the Association for Computational Linguistics*. Cambridge, MA: A.C.L.
- Klemensiewicz, Z., 1965. *Podstawowe wiadomości z gramatyki języka polskiego*. Warsaw: PWN.
- Mańczak, W., 1956. Ile rodzajów jest w polskim? *Język polski* 36, 116–21.
- Matthews, P.H., 1972. *Inflectional morphology*. Cambridge: Cambridge University Press.
- Pasch, H., 1986. *Die Mba-Sprachen*. Hamburg: Buske. (= Sprache und Geschichte in Afrika 6)
- Quillian, M.R., 1968. Semantic memory. In: M. Minsky (ed.), *Semantic information processing*, 227–270. Cambridge, MA: MIT Press.

- Rothstein, R.A., 1993. Polish. In: B. Comrie and G. Corbett (eds.), *The Slavonic Languages*, 686–758. Routledge: London.
- Saloni, Z., 1976. Kategoria rodzaju we współczesnym języku polskim. In: *Kategorie gramatyczne grup imiennych w języku polskim*, 43–78. Wrocław: Prace Instytutu Języku Polskiego.
- Saloni, Z., 1988. O tzw. formach nieosobowych męskoosobowych we współczesnej polszczyźnie. *Biuletyn Polskiego Towarzystwa Językoznawczego* XLI, 155–166.
- Vaillant, A., 1958. *Grammaire comparée des langues slaves*. Lyon: I.A.C.
- Weiss, D., 1993. How many sexes are there? In: G. Hentschel and R. Laskowski (eds.), *Studies in Polish morphology and syntax*, 71–105. Munich: Otto. (= *Specimina Philologicae Slavicae* 99)
- Wertz, C.A., 1977. The number of genders in Polish. *Canadian Slavonic Papers* 19, 50–63.
- Wunderlich, D. and R. Fabri, 1995. Minimalist morphology: An approach to inflection. *Zeitschrift für Sprachwissenschaft* 14, 236–294.
- Zagórski, Z., 1967. Wewnętrzne tendencje rozwojowe gwar północnowielkopolskich. Poznań: Uniwersytet im. Adama Mickiewicza.
- Zieniukowa, J., 1981. Rodzaj męski osobowy we współczesnych językach zachodniostowiańskich. Wrocław: Zakład narodowy im. Ossolińskich Wydawnictwo Polskiej Akademii Nauk.
- Zwicky, A.M., 1985. How to describe inflection. *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*, 372–386. Berkeley, CA: Berkeley Linguistics Society.
- Žukova, A.N., 1972. *Grammatika korjaskogo jazyka*. Leningrad: Nauka.