1

Gender Resolution in Rumanian

Louisa Sadler

This paper offers a contribution to the treatment of agreement phenomena in LFG by providing an analysis of Rumanian nominal agreement, focussing on gender.

I take up two issues concerned with gender marking and agreement in Rumanian. The first of these is the apparent mismatch between the number of nominal controller genders (three), and the number of target genders (two): nouns appear to make more gender distinctions than the elements which agree with them. This phenomenon, which is not unique to Rumanian, has engendered a number of analyses and on the face of it is a challenge to approaches to agreement by token identity or co-specification. I show how this can be accommodated straightforwardly in LFG. Second, Rumanian is a language in which syntactic resolution of gender under coordination is limited to inanimate NPs: conjoined inanimates resolve to the feminine plural unless all of them are masculine but mixed sex animates resolve to the masculine (Farkas, 1990, Lumsden, 1992, Corbett, 1991, Farkas and Zec, 1995, Wechsler and Zlatić, 2003, Wechsler, to appear). It is therefore interesting in terms of understanding how syntactic and semantic resolution interact, an issue which arises in various forms in a substantial number of languages. I formulate an approach which combines the set-based approach to syntactic gender resolution of Dalrymple and Kaplan (2000) with a specification of semantic resolution. This paper started out in a very practical fashion, in that I needed to get to grips with the implementation of closed sets as values in the XLE, a grammar engineering platform for LFG (Crouch et al., 2006), and needed a domain. The analysis proposed here is implemented as an XLE grammar fragment, and as usual, the experience of writing a grammar fragment showed that the prob-

2 / LOUISA SADLER

lem had intricacies which were not at first apparent, and has ultimately helped clarify my thinking about the problem and in particular about the interaction of syntactic and semantic resolution. At several points in this paper I footnote minor divergences between the theoretical descriptions and the notation the XLE supports.

Section 1.1 reviews the data concerning the number of genders in Rumanian, showing the mismatch between the number of controller genders and target genders. This section also illustrates the agreement properties of coordinations of inanimate nouns. Section 1.2 looks at gender in animate nouns, presenting data concerning agreement patterns in the coordination of animate nouns, and also cases in which natural and grammatical gender diverge. I then turn to analyses of the data, starting with a brief review of the most comprehensive approach in the literature, that of Farkas (Farkas, 1990, Farkas and Zec, 1995), in section 1.3. Section 1.4 provides an LFG analysis of target and controller genders in Rumanian, proposing that targets underspecify the features of their controllers. Section 1.5 extends this approach to deal with gender resolution for coordinations of inanimate NPs, building on the proposals of Dalrymple and Kaplan (2000). I then discuss coordination of animate nouns in section 1.6 and formulate the agreement generalization for these coordinate structures. Section 1.7 shows how the agreement for all NP coordinations can be captured succinctly, and section 1.8 concludes with some additional data.

1.1 Three Nominal Genders

Rumanian nouns fall straightforwardly into three distinct gender classes when we consider their behaviour in construction with agreement targets such as adjectives, as illustrated in (1–6) below. In Rumanian, participles and predicate adjectives show predicate-argument agreement with the subject, and determiners and adjectives within NP agree with the head noun (head-modifier agreement), as shown in (1–4) for masculine and feminine nouns (examples from Farkas and Zec 1995, glosses slightly altered for consistency):

(1) un copac frumos
    a.M tree.MSG beautiful.MSG
    'a beautiful tree'

(2) doi copaci frumoşi
    two.M trees.MPL beautiful.MPL
    'two beautiful trees'

(3) o rochie frumoasă
    a.F dress.FSG beautiful.FSG
    'a beautiful dress'

(4) două rochii frumoase
    two.F dresses.FPL beautiful.FPL
    'two beautiful dresses'

There is a third class of nouns shown in (5–6) and glossed as neuter, which show a mixed behaviour:

(5) un scaun   frumos
a.M chair.NSG beautiful.MSG
'a beautiful chair'

(6) două scaune   frumoase
two.F chairs.NPL beautiful.FPL
'two beautiful chairs'

Assignment to a gender class is partly driven by formal factors in Rumanian — nouns ending in [e] are MASC or FEM, those ending in any other vowel are FEM, and nouns ending in a consonant are MASC or NEUT (Farkas and Zec, 1995), but there is also a semantic dimension to syntactic gender assignment: nouns referring to males are MASC in gender while those referring to females are FEM. Nouns referring to inanimate objects may be in any of three classes.

Note that neuter is not an inquorate gender, that is, a gender with a very small number of members (Corbett, 1991, 170), but rather is a class fully on a par with the MASC and the FEM genders. This third class of nouns controls agreement forms identical to the MASC in the singular, and forms identical to the FEM in the plural. The agreement patterns determined by Rumanian nouns are summarised in Table 1.

TABLE 1 Nominal Agreement Patterns

N	Target	N	Target
FSG	FSG	FPL	FPL
MSG	MSG	MPL	MPL
NSG	MSG	NPL	FPL

How should we interpret this third class of nouns? One theoretical possibility is that this (large) class of lexemes simply belongs to two different syntactic genders — they really are MASC in the singular and FEM in the plural (as found with Somali gender polarity), with the existence of this "third" class being essentially a fact internal to the morphology. Such a proposal is found in recent work by Bateman and Polinsky (2005) who propose that Rumanian has just two noun classes in the singular and two in the plural, with membership determined on both formal and semantic grounds. A similar position is adopted in Wechsler and Zlatić (2003, 157): "the so-called neuter is really a class of inquorate nouns that are masculine in the singular but feminine in the plural".

On the other hand, in his wide-ranging study of gender as a morphosyntactic category, Corbett (1991) reasserts the traditional view and argues that the existence of three distinct agreement classes is itself enough to merit recognition of three genders in Rumanian, with a distinction emerging between controller and target genders. There is, furthermore, indication of a three way syntagmatic distinction in the syntax. In particular, there is clear evidence

4 / LOUISA SADLER

from coordination, where the behaviour of neuter singular nouns is evidently distinct from that of masculine singular nouns, that neuter should be differentiated as a third syntactic gender. Note the agreement patterns exemplified in the following data (Farkas and Zec, 1995, 96) for coordinations of singular nouns.

(7) a. Podeaua şi plafonul sînt albe.
    floor.DEF.FSG and ceiling.DEF.MSG are white.FPL
    'The floor and the ceiling are white.'

 b. Scaunul şi dulapul sînt albe.
    chair.DEF.NSG and cupboard.DEF.NSG are white.FPL
    'The chair and the cupboard are white.'

 c. Peretele şi scaunul sînt albe.
    wall.DEF.MSG and chair.DEF.NSG are white.FPL
    'The wall and the chair are white.'

 d. Podeaua şi scaunul sînt albe.
    floor.DEF.FSG and chair.DEF.NSG are white.FPL
    'The floor and the chair are white.'

 e. Podeaua şi uşa sînt albe.
    floor.DEF.FSG and door.DEF.FSG are white.FPL
    'The floor and the ceiling are white.'

 f. Nucul şi prunul sînt uscaţi.
    walnut.DEF.MSG and plum tree.DEF.MSG are dry.MPL
    'The walnut tree and the plum tree are dry.'

This data highlights the difficulty for the view that neuter nouns are simply members of a class MSG/FPL. On this view, a coordination of two MSG nouns should be indistinguishable from a coordination of two NSG nouns, which is clearly not the case (Indeed, Bateman and Polinsky 2005 explicitly leave the resolution behaviour under coordination as a problem in their account.). Table 2 summarizes.

TABLE 2 Nominal Agreement (Inanimates) under Coordination

NP1	NP2	AP	NP1	NP2	AP
NSG	NSG	FPL	FSG	MSG	FPL
FSG	FSG	FPL	FSG	NSG	FPL
MSG	MSG	MPL	MSG	NSG	FPL

The data considered in this section shows that Rumanian is a language which distinguishes three agreement classes (Corbett, 1991, 147) among nouns but has only two target genders — masculine and feminine. This mismatch phenomenon is found in other languages also — Corbett (1991) briefly

discusses Telugu (Dravidian) as having three controller genders and two target genders, and Lak (Caucasian) with four controller genders, three target genders in the singular and two target genders in the plural, as well as a number of other languages.

1.2 Coordination of Animate Nouns

It is well known that in some languages gender resolution in animate coordinate structures is semantically based, rather than taking account of the grammatical gender of the conjuncts. This is evident in particular when natural and grammatical gender diverge, as shown in the following example from French, a language with syntactic resolution for inanimates.

(8) La sentinelle et la personne à la barbe ont été pris
the sentry.FSG and the person.FSG to the beard have been taken.MPL
/*prises en otage.
/taken.FPL hostage
'The sentry and the person with the beard were taken hostage.' (Wechsler, to appear, 10)

In general then, we must allow for syntactic resolution to exist alongside other resolution processes in one and the same language. The following examples illustrate the resolution patterns for coordinations of animate nouns in Rumanian.[1]

(9) Maria şi tata au fost văzuti.
Maria.FSG and father.MSG were seen.MPL
'Maria and father were seen.' (Moosally, 1998, 112)

(10) Maria şi mama au fost văzute.
Maria.FSG and mother.FSG were seen.FPL
'Maria and mother were seen.' (Farkas and Zec, 1995, 94)

(11) Ion şi tata au fost văzuti.
Ion.MSG and father.MSG were seen.MPL
'Ion and father were seen.' (ibid. 95)

(12) un vizitator şi o turistă mult interesaţi
a visitor.MSG and a tourist.FSG very interested.MPL
'a very interested (male) visitor and a very interested (female) tourist'
(Maurice, 2001, 237)

As these examples show, unlike coordination of inanimates, coordinations of animate nouns determine masculine agreement if any of the conjuncts are male-denoting. Confirmation that the determining factor is semantic rather

---

[1]Glosses have been added as appropriate, where they were absent from the original.

than grammatical gender assignment comes both from nominals which are not (semantic) gender specific, but which are feminine in form (*persoană*, 'person'), and those which denote a male individual but are feminine in form (*popă*, 'priest').

Such nouns control agreement of adjectives, determiners, participles and predicative adjectives in terms of their grammatical gender, but participate in semantically based agreement in coordination. The pronominal anaphor referring back to nouns such as *persoană* also reflects the natural gender of the denotata.

(13) Persoană   cu   barbă a fost văzută.   El trebuie arestat
person.DEF.FSG with beard was   seen.FSG he must   arrested.MSG
imediat.
immediately
'The person with a beard was seen. He must be arrested immediately.'
(Farkas and Zec, 1995, 94)

(14) Maria şi   santinelă   au fost căsătoriti   de catre protul
Maria and sentry.DEF.FSG were   married.MPL by   priest.DEF
local.
local
'Maria and the sentry were married by the local priest.' (Wechsler and Zlatić, 2003, 188)

(15) Maria şi   persoană   cu   rochie au   fost văzute.
Maria and person.DEF.FSG with dress   have been seen.FPL
'Maria and the person with a dress have been seen.' (Farkas and Zec, 1995, 94)

In summary, for animates the resolution behaviour under coordination refers to natural rather than grammatical gender, and animate nouns may show a mismatch between grammatical and natural gender: Table 3 compares with Table 2 in the previous section.

TABLE 3 Animate Nominal Agreement under Coordination

NP1	NP2	Target
FEMALE	MALE	MPL
FEMALE	FEMALE	FPL
MALE	MALE	MPL

1.3 Previous Accounts

The facts outlined in the previous sections are described in the general descriptive and typological literature on agreement (Corbett 1991 is a typical example) and have attracted some theoretical attention, including Farkas (1990), Lumsden (1992), Farkas and Zec (1995), Wechsler (to appear), and Wechsler and Zlatić (2003). The most comprehensive discussion is that of Farkas (Farkas, 1990, Farkas and Zec, 1995). In this section I briefly review this approach, which is based on underspecifying the gender value of the nouns.

Farkas (1990) takes agreement to be a directional process of feature copying from the agreement trigger to the agreement target, which initially has unspecified features. Feminine nouns are lexically specified as [+Fem], masculine as [−Fem] and neuter nouns are lexically unspecified for gender. A feature co-occurrence restriction (a feature-filling rule, applying thus only to neuter nouns) gives a gender value for neuter plural nouns:

(16) $\begin{bmatrix} < [+\text{N}][-\text{V}] > \\ +\text{PLURAL} \end{bmatrix} \rightarrow [\,+\text{FEM}\,]$

Neuter singulars are masculine by the Elsewhere Principle, on the assumption that [−Fem] is the default value in the system. This is encoded in the following Feature Specification Default:

(17) $[\quad] \rightarrow [\,-\text{FEM}\,]$

On this view, then, neuter singular Ns are masculine (though not lexically specified as such) and neuter plural Ns are feminine (though again, not by lexical specification). Adjectives and determiners in agreement with neuter nouns will therefore be masculine or feminine depending on the number of the noun and will acquire features in the syntax copied from the controller noun. This approach effectively holds that there are just two syntactic genders in Rumanian: the difference between masculine nouns and neuter nouns in the singular coming down to whether the [−Fem] feature is introduced lexically or by a feature specification default.

Farkas and Zec (1995), which is largely concerned with patterns of agreement under coordination for both animate and inanimate nouns, adopts a slightly revised version of this proposal in the light of this additional data. As before, neuter nouns are lexically unspecified. The following rules are postulated (ordered by the Elsewhere Condition), which provide values for the gender feature:[2]

(18) $[\emptyset\text{F}\,] \rightarrow [-\text{F}]$

---

[2] The approach to agreement between controller and target differs from Farkas (1990) in that it is agnostic on the choice between a directional copying approach and a feature matching approach.

(19) $\begin{bmatrix} \emptyset F \\ \text{Number} \quad [+\text{PL}] \end{bmatrix} \rightarrow [+\text{F}]$

Thus consider an example like (5) repeated here for convenience as (20). The neuter noun is lexically unspecified for gender, but the default in (18) specifies a $-$ value for the feature F: the noun will thus behave syntactically as a MASC noun.

(20) un   scaun   frumos
     a.M chair.NSG beautiful.MSG
     'a beautiful chair'

Consideration of Rumanian agreement patterns leads Farkas and Zec (1995) to abandon the "morphosyntactic resolution rules" approach to coordinate noun phrases. Coordinate structures are taken to be headless: in the absence of a head, the content of morphosyntactic agreement features are determined by the following generalization for animates (Farkas and Zec, 1995, 95):

(21)   **Gender Assignment to groups (animate)**
       a. If the discourse referent includes a male individual, its gender is $[-\text{F}]$.
       b. Otherwise, the referent receives no gender specification.

In case b, the rule in (19) will determine the syntactic gender assignment as feminine.

Because non-coordinate NPs are lexically headed they inherit the agreement features of the head: thus animate 'mismatch' nouns control morphosyntactic agreement (targets agree with the syntactic gender features, so that *persoană cu barbă* occurs with a FSG participle or adjective). On the other hand, pronouns are always governed by discourse factors, so that for animates, male referents determine $[-\text{F}]$ pronouns, and female referents $[+\text{F}]$ pronouns: thus in the case of mismatch nouns, pronouns reflect the natural gender rather than the grammatical gender (see (13)).

For inanimate coordinate phrases, Farkas and Zec (1995, 97) propose the following generalization:

(22)   **Gender Assignment to groups (inanimate)**
       a. If all the components of a composite discourse referent are $[-\text{F}]$, the discourse referent inherits this specification.
       b. Otherwise, the referent receives no gender specification.

Again, the intention is that if the composite discourse referent fails case a, then the rules in (18) and (19) will be relevant and provide a syntactic gender assignment.

The relevant cases concern the following contrasting behaviour between MASC and NEUT nouns under coordination. We provide the lexical specifications according to Farkas and Zec (1995) in parentheses in Table 4.

TABLE 4 Lexical Specifications

NP1	NP2	Target Morphology
MSG ($-$F)	MSG ($-$F)	MPL
MSG ($-$F)	NSG ($\emptyset$F)	FPL
NSG ($\emptyset$F)	NSG ($\emptyset$F)	FPL

The intention is clearly that case b apply whenever there is a neuter conjunct, allowing (19) to determine the syntactic gender assignment to the group as FEM. But for this to happen it is crucial that the default in (18) *fail* to apply at the level of the conjuncts themselves. Otherwise the effect would be to resolve the underspecified $\emptyset$F on all the NSG nouns to $-$F, resulting in the assignment shown in Table 5.

TABLE 5 Specifications After (18)

NP1	NP2	Target Morphology
MSG ($-$F)	MSG ($-$F)	MPL
MSG $-$F	NSG $-$F	MPL
NSG $-$F	NSG $-$F	MPL

The problem is that it is not clear how the rules in (18) and (19) are to be prevented from applying as described to the conjuncts, which lack a lexical specification for the gender feature: note that a modifier such as a numeral, quantifier or attributive adjective shows only a binary distinction between $\pm$F and thus the NP will be determinate for gender, whether the agreement mechanism is feature copying or feature matching.

Working within a constraint-based formalism, Wechsler and Zlatić (2003)[3] develops a related approach, within the wider context of a theory concerning the interaction of syntactic and semantic resolution. For Wechsler and Zlatić (2003), coordinate NPs necessarily lack an inherent gender because they are headless. They postulate the following universal generalizations for such cases:

(23) Gender agreement with an animate NP that lacks inherent gender is always interpreted semantically. (Wechsler and Zlatić, 2003, 150)

---

[3]As an alternative reference, Wechsler (to appear) covers precisely the same ground.

(24) Rule for deriving gender of inanimate aggregate discourse referents:
D.R. [{ [GEND γ_1], ... [GEND γ_n] }] $\Leftrightarrow$
D.R. [GEND γ_1] $\cap$ .... $\cap \gamma_n \cap G_s$]
where γ_1 ... γ_n are null or unary sets and G_s is the set of s-gender features in the grammar (Wechsler and Zlatić, 2003, 152)

The rule in (24) for coordinations of inanimates states that the value of GEND for the coordinate NP is the intersection of the semantically-interpretable genders (typically masculine and feminine) of the conjunct daughters (gender features, on this proposal are the empty set and single-ton sets, e.g. {F}). In the case of Rumanian, as noted above, they assume that neuter nouns are simply members of a mixed class MSG/FPL, and thus it seems that all nouns will have an s-gender feature on this proposal. For inanimate coordinations falling under (24), they take FEM as the resolution class. One problem with this approach is that it predicts that a coordination of MSG with NSG will resolve in precisely the same manner as a coordination of two MSG nouns, because NSG and MSG are indistinguishable.

These accounts, then, are based on an approach which posits only two syntactic genders for Rumanian nouns. The account we develop in the following sections, on the other hand, recognises three nominal genders but only two target genders on adjectives and participles.

1.4 Targets as Underspecified

Rather than take neuter nouns as lexically underspecified for gender, or as members of a mixed class, we will propose instead that the *targets* of agreement underspecify the agreement features of their controllers. Nouns are specified as belonging to one of the three nominal genders (we will modify the expression of this approach to use sets as values for the GEND feature shortly). For example:

(25) *copac*   ($\uparrow$ PRED) = 'TREE'   *rochie*   ($\uparrow$ PRED) = 'DRESS'
            ($\uparrow$ GEND) = MASC   ($\uparrow$ GEND) = FEM
            ($\uparrow$ NUM) = SG   ($\uparrow$ NUM) = SG
    scaun   ($\uparrow$ PRED) = 'CHAIR'
            ($\uparrow$ GEND) = NEUT
            ($\uparrow$ NUM) = SG

Adjectives and determiners place constraints along the lines shown in (26).

(26) *frumoasă*   *frumoşi*
$((\text{ADJ} \in \uparrow) \text{ GEND}) = \text{FEM}$ $((\text{ADJ} \in \uparrow) \text{ GEND}) = \text{MASC}$
$((\text{ADJ} \in \uparrow) \text{ NUM}) = \text{SG}$ $((\text{ADJ} \in \uparrow) \text{ NUM}) = \text{PL}$
$(\uparrow \text{ PRED}) = \text{'BEAUTIFUL'}$ $(\uparrow \text{ PRED}) = \text{'BEAUTIFUL'}$

frumos   *frumoase*
$((\text{ADJ} \in \uparrow) \text{ GEND}) \neg = \text{FEM}$ $((\text{ADJ} \in \uparrow) \text{ GEND}) \neg = \text{MASC}$
$((\text{ADJ} \in \uparrow) \text{ NUM}) = \text{SG}$ $((\text{ADJ} \in \uparrow) \text{ NUM}) = \text{PL}$
$(\uparrow \text{ PRED}) = \text{'BEAUTIFUL'}$ $(\uparrow \text{ PRED}) = \text{'BEAUTIFUL'}$

Entries along the lines of the first two are appropriate for all FSG and MPL modifiers — these share the gender value of their head.[4] The second two entries, for MSG and FPL forms of nominal modifiers, are underspecified: the masculine singular form cannot combine with a feminine noun but combines freely with a masculine or neuter singular, and similarly the feminine plural will combine with the feminine or neuter plural.

Participles and predicate adjectives would be similarly specified. We show below the entry for a MSG predicative adjective:[5]

(27) Un   trandafir alb   e scump.
     a.MSG rose   white.MSG is expensive.MSG
     'A white rose is expensive.' (Farkas, 1990, 539)

(28) *scump* $\neg (\uparrow \text{ SUBJ GEND}) = \text{FEM}$
                $(\uparrow \text{ SUBJ NUM}) = \text{SG}$
                $(\uparrow \text{ PRED}) = \text{'EXPENSIVE'}$

As noted above, coordination of inanimate NPs always results in NPs which control FPL agreement unless all conjuncts are MSG (see Table 2 above, and recall that reference to NP1 and NP2 in this table does not encode facts about the linear order of the conjuncts): FPL morphology on a target can correspond to a neuter or a feminine plural controller. The resolution facts may be summarised as follows:[6]

---

[4]Some determiners are probably inflectional in Rumanian (see the data in (7)). Whether they are inflectional or co-heads, we assume they directly constrain the agreement features of the f-structure they share with the noun.

[5]For simplicity, we treat predicative and attributive adjectives by means of separate lexical entries in this paper. This fails to reflect the fact that they agree in precisely the same way with their controller —what differs is simply the *path* to the controller. The use of paths and local names in lexical entries permits the agreement properties of attributive and predicative adjectives to be given a unitary characterisation: see Otoguro (2006) for an approach to case and agreement along these lines.

[6](29) interprets a coordination of NSG as resolving to NPL —the agreeing FPL form follows from the underspecified requirements placed by FPL targets (see (26) above). An alternative, which we do not pursue here, is to assume that a coordination of NSG conjuncts itself resolves to FPL under the elsewhere clause.

(29)   Rumanian Resolution:
- If all conjuncts have the same gender, the coordinate structure has that gender.
- Otherwise the feminine form is used.

In the following section we replace the atomic gender values with set-valued features to extend our analysis to take account of agreement with coordinate (inanimate) controllers. We then turn to coordinations involving animate conjuncts.

1.5   Agreement and Coordination

Dalrymple and Kaplan (2000) propose an approach to the syntactic resolution of agreement features in coordinate structures which treats GEND as a set-valued rather than an atomic feature. On this approach, syntactic resolution reduces to the simple operation of set union. The value of the GEND feature of the coordinate structure as a whole is defined as the smallest set containing the values of the individual conjuncts, as in (30).

(30) NP $\longrightarrow$ NP   CONJ   NP

$\downarrow \in \uparrow$ $\downarrow \in \uparrow$

$(\downarrow \text{GEND}) \subseteq (\uparrow \text{GEND})$ $(\downarrow \text{GEND}) \subseteq (\uparrow \text{GEND})$

(31) $x \cup y$ is the smallest set z such that $x \subseteq z \ \land \ y \subseteq z$

The approach makes use of a notion of a *set designator* which indicates that the value of a feature is a set and also *exhaustively enumerates* the elements of the set. For example, the equation $(\uparrow \text{CASE}) = \{\text{NOM}, \text{ACC}\}$ (in which $\{\text{NOM}, \text{ACC}\}$ is a set designator) defines the value of CASE (for the f-structure in question) to be the set $\{\text{NOM}, \text{ACC}\}$, and the constraint $(\uparrow \text{SUBJ GEND}) =_c \{\text{M}\}$ requires the value to be the (singleton) set $\{\text{M}\}$ (Dalrymple and Kaplan, 2000).

Following this approach to the GEND feature, we can represent the Rumanian nominal genders as follows:

(32) Rumanian:

MASC	$\{\text{M}\}$
FEM	$\{\text{M}, \text{N}\}$
NEUT	$\{\text{N}\}$

(33) *copac* $(\uparrow \text{GEND}) = \{\text{M}\}$ *rochie* $(\uparrow \text{GEND}) = \{\text{M}, \text{N}\}$

                $(\uparrow \text{NUM}) = \text{SG}$ $(\uparrow \text{NUM}) = \text{SG}$

                $(\uparrow \text{PRED}) = \text{`TREE'}$            $(\uparrow \text{PRED}) = \text{`DRESS'}$

    scaun $(\uparrow \text{GEND}) = \{\text{N}\}$

                $(\uparrow \text{NUM}) = \text{SG}$

                $(\uparrow \text{PRED}) = \text{`CHAIR'}$

The lexical entries for predicative adjectives (as in 34–37) are along the lines shown in (38–41).[7,8]

(34) Un   trandafir alb   e scump.
     a.MSG rose.MSG white.MSG is expensive.MSG
     'A white rose is expensive.' (Farkas, 1990, 539)

(35) O   garoafă   albă   e scumpă.
     a.FSG carnation.FSG white.FSG is expensive.FSG
     'A white carnation is expensive.' (ibid:539)

(36) Un   scaun   confortabil   e folósitor.
     a.MSG chair.NSG comfortable.MSG is useful.MSG
     'A comfortable chair is useful.' (ibid: 540)

(37) Nişte   scaune   confortabile   e folositoare.
     some.FPL chair.NPL comfortable.FPL are useful.FPL
     'Some comfortable chairs are useful.' (ibid: 540)

(38) *scumpă* (SUBJ GEND *must be* FEM)
     ($\uparrow$ SUBJ GEND) $=_c$ {M, N}
     ($\uparrow$ SUBJ NUM) = SG
     ($\uparrow$ PRED) = 'EXPENSIVE'

(39) *scumpi* (SUBJ GEND *must be* MASC)
     ($\uparrow$ SUBJ GEND) $=_c$ {M}
     ($\uparrow$ SUBJ NUM) = PL
     ($\uparrow$ PRED) = 'EXPENSIVE'

(40) *scump* (SUBJ GEND *can't be* FEM)
     ($\uparrow$ SUBJ GEND) $\neg$ = {M, N}
     ($\uparrow$ SUBJ NUM) = SG
     ($\uparrow$ PRED) = 'EXPENSIVE'

(41) *scumpe* (SUBJ GEND *can't be* MASC)
     ($\uparrow$ SUBJ GEND) $\neg$ = {M}
     ($\uparrow$ SUBJ NUM) = PL
     ($\uparrow$ PRED) = 'EXPENSIVE'

---

[7]The XLE does not appear to permit $=_c$ over closed sets as values, as shown in (38) and (43). This is encoded instead as a conjunction of constrained membership statements in the XLE:

   (i) {N} $\in_c$ ($\uparrow$ GEND)
      {M} $\in_c$ ($\uparrow$ GEND)

[8]XLE does not implement negation of closed sets as shown in (40) and (44). The negation shown on the entry for *frumos* (MSG) can be re-expressed as a negation over a conjunction of membership statements:

   (i) $\neg$ [{N} $\in$ ($\uparrow$ GEND) $\wedge$ {M} $\in$ ($\uparrow$ GEND)]

The negation shown for *frumoase* FPL can be re-expressed as a positive requirement that {N} is in the set.

14 / Louisa Sadler

For example in (37) *scaune* is lexically specified $(f1\ \text{GEND}) = \{\text{N}\}$ and the FPL adjective *folositoare* (like *scumpe* in (41)) specifies $(f1\ \text{GEND})\ \neg = \{\text{M}\}$, that is, requires the GEND value not to be the closed set containing the single element $\{\text{M}\}$, hence allowing the GEND value to be either $\{\text{N}\}$ or $\{\text{M}, \text{N}\}$. Given that there is a limited set of possibilities here, we can alternatively express this negative constraint as the equivalent:

(42) $\{\text{N}\} \in_c (\uparrow \text{SUBJ GEND})$

Attributive adjectives place constraints along the lines shown in (43) and (44), and other NP internal modifiers such as numerals, demonstratives and quantifiers will be similar.

(43) *frumoasă* (FSG) *frumoşi* (MPL)
  $((\text{ADJ} \in \uparrow)\ \text{GEND}) =_c \{\text{M, N}\}$ $((\text{ADJ} \in \uparrow)\ \text{GEND}) =_c \{\text{M}\}$
  $((\text{ADJ} \in \uparrow)\ \text{NUM}) = \text{SG}$ $((\text{ADJ} \in \uparrow)\ \text{NUM}) = \text{PL}$

(44) *frumos* (MSG) *frumoase* (FPL)
  $((\text{ADJ} \in \uparrow)\ \text{GEND})\ \neg = \{\text{M, N}\}$ $((\text{ADJ} \in \uparrow)\ \text{GEND})\ \neg = \{\text{M}\}$
  $((\text{ADJ} \in \uparrow)\ \text{NUM}) = \text{SG}$ $((\text{ADJ} \in \uparrow)\ \text{NUM}) = \text{PL}$

We now turn to the coordination examples in (7), restricting attention for the moment to the behaviour of inanimate conjuncts. According to the analysis of syntactic resolution in Dalrymple and Kaplan (2000), the GEND feature of the coordinate NP as a whole is the smallest set which has the GEND values of the conjunct daughters as subsets (see (30) and (31)). Table 6 summarises the results of this analysis.

TABLE 6 Nominal Coordination with Set Values

NP1		NP2		NPCoord	Target Morph
$\{\text{M N}\}$	(FSG)	$\{\text{M}\}$	(MSG)	$\{\text{M N}\}$	FPL
$\{\text{M}\}$	(MSG)	$\{\text{N}\}$	(NSG)	$\{\text{M N}\}$	FPL
$\{\text{M N}\}$	(FSG)	$\{\text{N}\}$	(NSG)	$\{\text{M N}\}$	FPL
$\{\text{N}\}$	(NSG)	$\{\text{N}\}$	(NSG)	$\{\text{N}\}$	FPL
$\{\text{M N}\}$	(FSG)	$\{\text{M N}\}$	(FSG)	$\{\text{M N}\}$	FPL
$\{\text{M}\}$	(MSG)	$\{\text{M}\}$	(MSG)	$\{\text{M}\}$	MPL

On this first pass, the phrase structure rule for Rumanian is constrained to apply only to inanimate NPs because, as we have seen, animate NPs undergo semantic resolution under coordination. We assume that nouns are lexically specified as ANIM + or −. The following rule is restricted so that only coordinate structures in which all conjuncts are inanimate undergo resolution by

set union.[9]

(45) NP $\longrightarrow$

NP	CONJ	NP
$\downarrow \in \uparrow$		$\downarrow \in \uparrow$
$(\downarrow \text{GEND}) \subseteq (\uparrow \text{GEND})$		$(\downarrow \text{GEND}) \subseteq (\uparrow \text{GEND})$
$(\downarrow\text{ANIM}) = -$		$(\downarrow\text{ANIM}) = -$

To conclude this section, we observe that a simple account of the different numbers of controllers and targets can be given by the simple method of using negative conditions. Moreover, the otherwise slightly puzzling (inanimate) agreement pattern of two neuters under coordination is straightforwardly accommodated under an approach using closed sets for agreement features and set union for syntactic resolution.

1.6 Semantic Resolution

In very many languages, the sort of syntactic resolution under coordination of gender features modelled in the proposal of Dalrymple and Kaplan (2000) by set descriptors and set union is one aspect of the phenomenon and exists alongside other processes and in particular semantically-based resolution in the case of conjoined animates (see Corbett 1991, Wechsler and Zlatić 2003 for some discussion). Coordinations of animate NPs in Rumanian do not resolve syntactically, but according to the following generalization:

(46) a. If one conjunct denotes a male animate then M is used.
   b. If all conjuncts are M, then M is used.
   c. Otherwise, F is used.   (Corbett, 1991, 289)

We now consider how the approach to syntactic resolution in the previous section, directly modelled on Dalrymple and Kaplan (2000), can be combined with a formulation of semantic resolution.[10] Our approach starts from the observation that if any of the conjuncts refers to a MALE individual, then the f-structure corresponding to the coordinate structure as a whole is marked as having masculine gender. To encode the notion of reference to a male individual (or set of individuals) I posit an additional f-structure feature SEMGEND

---

[9]An alternative is to declare the feature ANIM as distributive, which would additionally rule out mixed animacy coordination. We will return to this issue shortly.

[10]Wechsler and Zlatíc (2003) discusses languages which exhibit both syntactic and semantic resolution under coordination. Although their approach is not formalized in detail, it takes LFG as its framework of reference. The essence of their proposal is that the GEND feature of an animate coordinate structure will have a semantic value while the GEND feature of a inanimate coordinate structure will have a set-valued feature. Semantically assigned values are taken to be semantic forms such as 'female', 'non-female', with the assumption that 'the negatively defined semantic feature 'non-female' is not distributive (since negation itself is not distributive): a 'non-female' group is a group that fails to meet the description of a 'female' group (namely a group of females). Thus any group containing at least one male is a 'non-female' group" (Wechsler and Zlatíc, 2003, 151). There are clearly a number of issues concerning how such an account might be formalized, but discussion of these matters would take us too far afield.

with values MALE and FEMALE. A similar feature is used to encode semantic gender in Network Morphology lexical networks, although such analyses do not deal with semantic resolution in the syntax for coordinate structures (Corbett and Fraser, 2000). I assume that lexical entries which denote male individuals are lexically specified as ($\uparrow$ SEMGEND) = MALE (including mismatch nouns which are syntactically FEM), and those which denote female individuals are likewise marked as ($\uparrow$ SEMGEND) = FEMALE. Nouns which lack an inherent semantic gender are ambiguous out of context and thus in principle may undergo either coordination schema.[11]

For clarity, I proceed by considering first what sorts of annotations would be necessary to encode the generalization in (46). In (47), the functional uncertainty on the second conjunct daughter will be interpreted existentially: it succeeds if there is a member of the set with SEMGEND = MALE.[12]

(47) NP $\longrightarrow$

	NP	CONJ	NP
	$\downarrow \in \uparrow$		$\downarrow \in \uparrow$
	$(\downarrow \text{ANIM}) = +$		$(\downarrow \text{ANIM}) = +$
			$(\uparrow \in \text{SEMGEND}) = \text{MALE}$
			$(\uparrow \text{ GEND}) = \{\text{M}\}$

This rule will only succeed if one member (at least) is MALE. Otherwise, all the daughters have SEMGEND = FEMALE and the syntactic gender is set to feminine ($\{\text{M N}\}$) for the set as a whole.

(48) NP $\longrightarrow$

	NP	CONJ	NP
	$\downarrow \in \uparrow$		$\downarrow \in \uparrow$
	$(\downarrow \text{ANIM}) = +$		$(\downarrow \text{ANIM}) = +$
	$(\downarrow \text{ SEMGEND}) = \text{FEMALE}$		$(\downarrow \text{ SEMGEND}) = \text{FEMALE}$
			$(\uparrow \text{ GEND}) = \{\text{M, N}\}$

These rules can be combined into one. (48) requires all the members of the coordinate set to have SEMGEND = FEMALE. Since negation in a functional uncertainty is given a wide scope interpretation, that is, is interpreted as a universal (not an existential), we can express this condition as:

(49) ($\uparrow \in$ SEMGEND) $\neg=$ (MALE)

      (there is no member of the set for which SEMGEND = MALE is true)

---

[11] The rules in (47) and (48) as formulated predict that if a mismatch noun which is MASC in syntactic gender but refers to a FEMALE individual is coordinated with another noun which refers to a FEMALE individual, the set as a whole will control FEM agreement. I do not currently have any grammatical/natural gender mismatch data to confirm or contradict this.

[12] The constraint ($\uparrow \in$ SEMGEND) = MALE is arbitrarily placed on the second conjunct and could as well be associated with the CONJ daughter. Clearly the rules can also be extended to cover additional conjuncts by adding a Kleene-plus to the first conjunct.

(50) NP $\longrightarrow$ NP   CONJ   NP
                            $\downarrow \in \uparrow$ $\downarrow \in \uparrow$

$(\downarrow \text{ANIM}) = +$ $(\downarrow \text{ANIM}) = +$

$[(\uparrow \in \text{SEMGEND}) \neg= \text{MALE}$

$(\uparrow \text{GEND}) = \{\text{M, N}\}$

$| \quad (\uparrow \text{GEND}) = \{\text{M}\}]$

1.7 Combining Syntactic and Semantic Resolution

We have proposed two rules for coordinate structures, one for animate conjuncts and one for inanimate conjuncts, capturing the generalizations concerning agreement and especially agreement under coordination discussed in Farkas and Zec (1995). These rules, repeated as (51) and (52) below, follow the assumption in Farkas and Zec (1995) that coordinate NPs combine either animate or inanimate conjuncts but do not mix the two: the two separate rules in (52) and (51) ensured that the animacy features of the conjuncts matched.

(51) NP $\longrightarrow$ NP   CONJ   NP
                            $\downarrow \in \uparrow$ $\downarrow \in \uparrow$

$(\downarrow \text{ANIM}) = +$ $(\downarrow \text{ANIM}) = +$

$[(\uparrow \in \text{SEMGEND}) \neg= \text{MALE}$

$(\uparrow \text{GEND}) = \{\text{M, N}\}$

$| \quad (\uparrow \text{GEND}) = \{\text{M}\}]$

(52) NP $\longrightarrow$ NP   CONJ   NP
                            $\downarrow \in \uparrow$ $\downarrow \in \uparrow$

$(\downarrow \text{ANIM}) = -$ $(\downarrow \text{ANIM}) = -$

$(\downarrow \text{GEND}) \subseteq (\uparrow \text{GEND})$ $(\downarrow \text{GEND}) \subseteq (\uparrow \text{GEND})$

However though this is the assumption made in Farkas and Zec (1995), Moosally (1998) observes that this does not seem to be correct and gives the following examples of mixed animate/inanimate coordinations:

(53) Mašina şi sofer-ul au fost văzuti ieri.
car.FSG and driver.MSG have.PL been seen.MPL yesterday
'The car and the driver (male referent) were seen yesterday.' (Moosally, 1998, 113)

(54) Mašina şi sofer-ita au fost văzute ieri.
car.FSG and driver.FSG have.PL been seen.FPL yesterday
'The car and the driver (female referent) were seen yesterday.' (Moosally, 1998, 113)

The following example also shows that the "animate" strategy is not limited to human animates but extends to (at least some) animals:

(55) Pisica şi ciinele sing inometati.
cat.FSG and dog.MSG are hungry.MPL
'The cat and the dog are hungry.' (Moosally, 1998, 114)

18 / LOUISA SADLER

Clearly more research is needed to determine precisely what the facts are here, but on the assumption that Moosally's data are correct, it is possible to give a rather succinct statement of the resolution facts for Rumanian, using templates to encode generalizations. We define first the following templates:[13]

(56) a. CONJUNCT $=$ $\downarrow \in \uparrow$
b. RES-GEND $=$ $(\uparrow \in \text{SEMGEND}) \neg = \text{MALE}$
$(\downarrow \text{GEND}) \subseteq (\uparrow \text{GEND})$
c. SEM-GEND $=$ $(\uparrow \in \text{SEMGEND}) =_c \text{MALE}$
$(\uparrow \text{GEND}) = \{\text{M}\}$

(57) NP-CONJUNCT   @ CONJUNCT
@ RES-GEND │ SEM-GEND

(56b) and (56c) together state the disjunction: either some conjunct has the feature SEMGEND = MALE defined, in which case the GEND of the coordinate structure is $\{\text{M}\}$ (semantic resolution), or no conjunct has the feature SEM-GEND = MALE defined, in which case the GEND of the coordinate structure is given by syntactic resolution (set union).

(58) NP $\longrightarrow$ NP   CONJ   NP
@NP-CONJUNCT   @NP-CONJUNCT

Clearly, it is also possible to encode relatively succinctly (by using templates) the situation holding in a language in which it is impossible to mix animate and inanimate conjuncts (i.e. a language in which (51) and (52) are the operative rules).

1.8 Conclusion and Further Data

Finally, it is worth noting that the discussion in the previous literature (Farkas, 1990, Farkas and Zec, 1995, Lumsden, 1992, Wechsler, to appear, Wechsler and Zlatić, 2003) is concerned with coordinations of non-coreferring singular NPs, but additional data suggests the existence of further agreement patterns. For example, Maurice (2001) notes that with inanimates in a coordination of SG and PL it is the PL which determines agreeing forms, as shown by the following contrast:[14]

(59) Satelitul   şi avionele   au fost doborîte.
satellite.DET.MSG and airplane.DET.NPL have been shot.down.FPL
'The satellite and the airplanes have been shot down.' (Maurice, 2001, 238)

---

[13]We ignore PERS and NUM here.

[14]In (59) we gloss the agreement form as FPL in line with our practice elsewhere, but Maurice glosses it as NPL. This has no bearing on her point.

(60) Satelitii   şi avionul   au fost doborîţi.
satellite.DEF.MPL and airplane.DEF.NSG have been shot.down.MPL
'The satellites and the airplane have been shot down.' (ibid:238)

If two plurals are combined the predicate agrees with the closest conjunct:

(61) Sateliţii   şi avioanele   au fost doborîte.
satellite.DEF.MPL and airplane.DEF.NPL have been shot.down.FPL
'The satellites and the airplanes have been shot down.' (ibid:238)

(62) Avioanele   şi sateliţii   au fost doborîţi.
airplane.DEF.NPL and satellite.DEF.MPL have been shot.down.MPL
'The airplanes and the satellites have been shot down.' (ibid:238)

Aurora Petan (p.c.) gives the following, with two nouns denoting the same entity, with closest conjunct agreement and a singular verb:

(63) Speranta şi viitorul   meu   este acest copil.
hope.FSG and future.NSG my.MSG is   this child.
'My hope and future is this child.'

(64) Viitorul şi speranta mea   este acest copil.
future.NSG and hope.FSG my.FSG is   this child.
'My future and hope is this child.'

I leave these patterns to one side, but clearly a more comprehensive account of Rumanian agreement under coordination would have to take account of these patterns and their distribution.

This paper has shown that LFG permits a relatively simple and straightforward account of the intricacies of gender agreement in Rumanian, a language which both displays a separation between the number of target and controller gender and in which animate and inanimate noun phrases undergo different gender resolution patterns under coordination. The account posits three controller genders and two target genders and uses underspecification on targets to capture the agreement facts. The treatment of resolution under coordination builds on the set-based approach to resolution of Dalrymple and Kaplan (2000) and reduces these resolution patterns to a simple disjunction: if any conjunct is animate male, then the coordinate structure is marked as MASC, and otherwise, gender is resolved by set union.

Acknowledgments

20 / Louisa Sadler

to communicating formal notions. Over the years I have benefitted greatly from such interactions. Thanks to Nicoleta Bateman, Kakia Chatsiou, Mary Dalrymple, Despina Kazana, Tracy Holloway King, John Maxwell, Aurora Petan, Andrew Spencer and two anonymous reviewers for data, comments and feedback which have improved this paper — all errors are mine.

References

Bateman, Nicoleta and Maria Polinsky. 2005. Rumanian as a Two-Gender Language. unpublished LSA handout.

Corbett, Greville and Norman M. Fraser. 2000. Gender assignment: a typology and a model. In G. Senft, ed., *Systems of Nominal Classification*, pages 293–325. Cambridge: Cambridge University Press.

Corbett, Greville G. 1991. *Gender*. Cambridge, UK: Cambridge University Press.

Crouch, Dick, Mary Dalrymple, Ron Kaplan, Tracy King, John Maxwell, and Paula Newman. 2006. XLE documentation. Tech. rep., Palo Alto Research Center, Palo Alto, CA.

Dalrymple, Mary and Ron Kaplan. 2000. Feature indeterminacy and feature resolution in description-based syntax. *Language* 76(4):759–798.

Farkas, Donka. 1990. Two Cases of Underspecification in Morphology. *Linguistic Inquiry* 21(4):539–550.

Farkas, Donka and Draga Zec. 1995. Agreement and Pronominal Reference. In Guglielmo Cinque and Giuliana Giusti, ed., *Advances in Roumanian Linguistics*, pages 83–101. Philadelphia: John Benjamins.

Lumsden, John. 1992. Underspecification in Grammatical and Natural Gender. *Linguistic Inquiry* 23(3):469–86.

Maurice, Florence. 2001. Deconstructing Gender - The case of Rumanian. In Marlis Hellinger and Hadumod Bussmann, ed., *Gender across Languages: The linguistic representation of men and women*, vol. Vol 1, pages 229–52. Amsterdam: John Benjamins.

Moosally, Michelle. 1998. *Noun Phrase Coordination: Ndebele Agreement Patterns and Cross-Linguistic Variation*. Ph.D. thesis, University of Texas at Austin, Austin, Tx.

Otoguro, Ryo. 2006. *Morphosyntax of Case: A theoretical investigation of the concept*. Ph.D. thesis, University of Essex.

Wechsler, Stephen. to appear. 'Elsewhere' in Gender Resolution. In K. Hanson and S. Inkelas, eds., *The Nature of the Word — Essays in Honor of Paul Kiparsky*, pages 1–39. Boston,MA: MIT Press.

Wechsler, Stephen and Larisa Zlati´c. 2003. *The Many Faces of Agreement*. Stanford, CA: CSLI Publications.