Feature resolution by lists: the case of French coordination

Gabriel Aguila-Multner & Berthold Crysmann

Laboratoire de linguistique formelle, CNRS & U Paris-Diderot

Abstract. In this paper[1], we shall address resolution of gender and person in French coordination and suggest that a list-based encoding of feature values provides for a very simple and intuitive resolution mechanism in coordinate structures by means of simple list concatenation, while it leaves the treatment of agreement in head-compositional structures entirely unaffected. We shall discuss the implementation of this approach in the context of an emerging computational HPSG of French based on the LinGO Grammar Matrix (Bender et al., 2002), and argue that the problem at hand calls for concatenation by recursive copying (Emerson, 2017), as opposed to difference lists (Clocksin & Mellish, 1981). Finally, we conclude that the list-based encoding of person and gender values can act as a drop-in replacement for the standard sort-based encoding, since it is not only more flexible in the treatment of feature resolution, but also bears the further potential of representing more elaborate person systems, like the inclusive/exclusive distinction.

1 Feature resolution in coordination

Probably the most basic function of coordination is to combine individuals or events into aggregates. With individuals this typically creates aggregates that are treated as plurals, e.g. for the purposes of agreement. E.g. consider the examples in (1) and (2).

(1) Le chien et le chat dorment.
the dog.SG and the cat.SG sleep.PRS.3PL
'The dog and the cat sleep.'

(2) Le chien et le chat endormis se réveillent.
the dog.SG and the cat.SG asleep.PL.M REFL awaken.PRS.3PL
'The dog and the cat that were asleep are waking up.'

While the morphosyntactic number value of aggregates standardly reflects their semantic plurality (see, however, An & Abeillé 2017 for closest conjunct agreement in French NPs), there is no a priori expectation as to gender or

[1] The research reported on in this paper has been partially carried out within the excellency cluster (LabEx) "Empirical Foundations in Linguistics", supported by a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program (reference: ANR-10-LABX-0083).

person values of the coordinate structure unless, of course, they are composed of individuals of the same kind.

Languages like English do not show gender agreement in the plural, whereas French does, as illustrated by the agreement between subject and verb (3) or between subject and a predicative adjective (4).

(3) a. Les chevaux sont partis.
the horse(M).PL be.PRS.3PL leave.PTCP.PL.M
'The horses left.'

b. Les tortues sont parties.
the turtle(F).PL be.PRS.3PL leave.PTCP.PL.F
'The turtles left.'

(4) a. Les frelons sont dangereux.
the hornet(M).PL be.PRS.3PL dangerous.PL.M
'Hornets are dangerous.'

b. Les guêpes sont dangereuses.
the wasp(F).PL be.PRS.3PL dangerous.PL.F
'Wasps are dangerous.'

Gender agreement carries over to coordinate structures, as shown in (5).

(5) a. Le cheval et l'âne sont partis.
the horse(M) and the donkey(M) be.PRS.3PL leave.PTCP.PL.M
'The horse and the donkey left.'

b. La tortue et la salamandre sont parties.
the turtle(F) and the salamander(F) be.PRS.3PL leave.PTCP.PL.F
'The turtle and the salamander left.'

For coordination to be functional, in a linguistic sense, there need to be resolution strategies to determine agreement not only in the case of matching gender (or person) specifications, as in (5), but also in case of mismatch.

1.1 Gender resolution

The resolution of gender in French follows a pattern illustrated below. For a typological survey of gender systems and resolution strategies see Corbett (1991).

(6) a. Les juments et les ânesses sont parties.
the mare(F).PL and the donkey(F).PL be.PRS.3PL leave.PTCP.PL.F
'The mares and the female donkeys left.'

b. * Les juments et les ânesses sont partis.
the mare(F).PL and the donkey(F).PL be.PRS.3PL leave.PTCP.PL.M

(7) a. Les chevaux et les ânesses sont partis.
the horse(M).PL and the donkey(F).PL be.PRS.3PL leave.PTCP.PL.M
'The horses and the female donkeys left.'

b. * Les chevaux et les ânesses sont parties.
the horse(M).PL and the donkey(F).PL be.PRS.3PL leave.PTCP.PL.F

(8) a. Les juments et les ânes sont partis.
the mare(F).PL and the donkey(M).PL be.PRS.3PL leave.PTCP.PL.M
'The mares and the donkeys left.'

b. *Les juments et les ânes sont parties.
the mare(F).PL and the donkey(M).PL be.PRS.3PL leave.PTCP.PL.F

(9) a. Les chevaux et les ânes sont partis.
the horse(M).PL and the donkey(M).PL be.PRS.3PL leave.PTCP.PL.M
'The horses and the donkeys left.'

b. *Les chevaux et les ânes sont parties.
the horse(M).PL and the donkey(M).PL be.PRS.3PL leave.PTCP.PL.M

As can be seen, any occurrence of a masculine inside the coordinate structure resolves to masculine for the entire coordination, and only coordinations of exclusively feminine NPs (6) show feminine agreement. This is true at any level of embedding inside the coordinate structure, as example (10) testifies, and the constraints on agreement hold locally, as well as across non-local dependencies, as illustrated by the relative clause in (11).

(10) a. Les juments, les ânesses et les poneys sont
the mare(F).PL the donkey(F).PL and the pony(M).PL be.PRS.3PL
partis.
leave.PTCP.PL.M
'The mares, the female donkeys and the ponies left.'

b. *Les juments, les ânesses et les poneys sont
the mare(F).PL the donkey(F).PL and the pony(M).PL be.PRS.3PL
parties.
leave.PTCP.PL.F

(11) a. Le chien et la tortue, qui étaient endormis, se sont
the dog(M) and the turtle(F) who be.IPFV.3PL asleep.M.PL be.PRS.3PL
réveillés.
awaken.PTCP.PL.M
'The dog and the turtle, who were asleep, woke up.'

b. *Le chien et la tortue, qui étaient endormies, se sont
the dog(M) and the turtle(F) who be.IPFV.3PL asleep.F.PL be.PRS.3PL
réveillées.
awaken.PTCP.PL.F

1.2 Person resolution

Person resolution strategies are somewhat more complex than gender resolution strategies owing to the ternary distinction of person values in French. Person agreement is illustrated for simple non-coordinated subjects in examples (12), while the resolution pattern in coordinate structures can be observed in (13-15):

(12) a. Nous nous entendons bien.
1PL get.along.PRS.1PL well
'We get along well.'

b. Vous vous entendez bien.
2PL get.along.PRS.2PL well
'You get along well.'

c. Elles s'entendent bien.
3PL.F get.along.PRS.3PL well
'They get along well.'

(13) a. Toi et moi allons bien nous entendre.
you and I will.PRS.1PL well get.along.INF.1PL
'You and I will get along well.'

b. * Toi et moi allez bien vous entendre.
you and I will.PRS.2PL well get.along.INF.2PL

c. * Toi et moi vont bien s'entendre.
you and I will.PRS.3PL well get.along.INF.3PL

(14) a. Les enfants et moi allons bien nous entendre.
the child.PL and I will.PRS.1PL well get.along.INF.1PL
'The children and I will get along well.'

b. * Les enfants et moi vont bien s'entendre.
the child.PL and I will.PRS.3PL well get.along.INF.3PL

(15) a. Toi et les enfants allez bien vous entendre.
you and the child.PL will.PRS.2PL well get.along.INF.2PL
'You and the children will get along well.'

b. * Toi et les enfants vont bien s'entendre.
you and the child.PL will.PRS.3PL well get.along.INF.3PL

The generalisation can be formulated in terms of the person hierarchy (1>2>3):

- any first person conjunct triggers first person agreement;
- in the absence of any first person conjunct, any second person conjunct triggers second person agreement;
- otherwise (i.e. if all conjuncts are third person), third person agreement is used.

Once again, neither the depth of embedding in the coordinate structure (16) nor the locality of the agreement relation (17) seem to affect this pattern.

(16) a. Les enfants, les parents et moi nous entendons bien.
the child.PL the parent.PL and I get.along.PRS.1PL well
'The children, the parents, and I get along well.'

b. * Les enfants, les parents et moi s'entendent bien.
the child.PL the parent.PL and I get.along.PRS.3PL well

(17) a. Les enfants et moi, qui nous sommes rencontrés hier,
the child.PL and I who be.PRS.1PL meet.PTCP.PL.M yesterday
nous entendons bien.
get along.PRS.1PL well
'The children and I who have met yesterday get along well.'

b. * Les enfants et moi, qui se sont rencontrés hier, s'entendent bien.
the child.PL and I who be.PRS.3PL meet.PTCP.PL.M yesterday get along.PRS.3PL well

The pattern for person resolution we observe for French is actually more widely attested across languages and commonly referred to in the context of the person hierarchy: e.g. English antecedent-anaphora agreement follows this pattern (Zwicky, 1977), and so does subject-verb agreement in languages such as German or Russian (King & Dalrymple, 2004).

1.3 Discussion

Lexical-Functional Grammar uses rather sets to represent coordinations in *f*-structure. Properties imposed on the set can be distributed over the members of the set, e.g. case specifications, or not, as e.g. person, number or gender specifications (Dalrymple & Kaplan, 2000; King & Dalrymple, 2004). Since these sets tend to be flat, membership constraints may suffice to percolate non-distributive features onto set members.[2]

HPSG does not recognise an intermediate level of representation such as *f*-structure but rather builds up semantics in parallel with syntactic structure. E.g. in MRS (Copestake et al., 2005), coordinations of individuals are represented as a group individual (together with its quantifier) that embeds the semantic contribution of its left and right daughters via the L-INDEX and R-INDEX features respectively (cf. (18)). The HOOK features INDEX and LTOP, which define the syntax-semantics interface (Copestake et al., 2001), however, solely expose the index and label of the coordinate structure as a whole, as illustrated by the sample MRS in (18). As a masculine noun triggering masculine agreement of the entire coordination can be embedded arbitrarily deep in a coordination of feminines (see example (10)), access to any person or gender features of conjuncts would necessitate traversing the MRS graph, e.g. by means of functional uncertainty, a solution that runs counter the idea of a lean interface between syntax and semantics, as advanced by Copestake et al. (2001).

[2] Nevertheless, distribution of features in LFG will need to differentiate according to feature values, making the statement of resolution quite clumsy. E.g. feminine gender values will be distributive, whereas masculine values on the coordinate structure will only require membership on one of the *f*-structure sets contributed by the conjunct-daughters. One can imagine that such a regime will become even more unwieldy, once we move to a tri-fold resolution scheme, as observed with person.

(18) MRS for *la girafe et l'éléphant* (quantifiers omitted)

$$
\begin{bmatrix}
\textit{mrs} & \\
\text{HOOK} & \begin{bmatrix} \text{INDEX} & \boxed{c}\begin{bmatrix}\textit{ref-ind} & \\ \text{NUM} & \textit{pl}\end{bmatrix} \\ \text{LTOP} & \boxed{l} \end{bmatrix} \\
\text{RELS} & \left\langle
\begin{bmatrix} \text{PRED} & \textit{and\_coord\_rel} \\ \text{ARG0} & \boxed{c} \\ \text{LBL} & \boxed{l} \\ \text{L-INDEX} & \boxed{g} \\ \text{R-INDEX} & \boxed{e} \end{bmatrix},
\begin{bmatrix} \text{PRED} & \textit{girafe\_n\_rel} \\ \text{LBL} & \textit{handle} \\ \text{ARG0} & \boxed{g}\begin{bmatrix}\textit{ref-ind} & \\ \text{NUM} & \textit{sg} \\ \text{GEND} & \textit{f}\end{bmatrix} \end{bmatrix},
\begin{bmatrix} \text{PRED} & \textit{éléphant\_n\_rel} \\ \text{LBL} & \textit{handle} \\ \text{ARG0} & \boxed{e}\begin{bmatrix}\textit{ref-ind} & \\ \text{NUM} & \textit{sg} \\ \text{GEND} & \textit{m}\end{bmatrix} \end{bmatrix}, \ldots
\right\rangle
\end{bmatrix}
$$

A straightforward alternative solution would shift the burden to the syntax: for gender alone, a rather brute-force approach would expand what is now a single coordination rule into three, projecting *fem* to the coordination in the case of all feminine daughters, projecting *mas* from a masculine left daughter, and finally, projecting *mas* from the right daughter with a feminine left daughter[3]. The same needs of course to be done for person, yet with combinations of three values to be taken care of, instead of just two. Worse, since both gender and person may need to be resolved at the same time, as illustrated by the example in (19), the rules may actually multiply out, in the worst case.[4]

(19)

Elle,	son frère	et	moi	nous sommes	bien	entendus.
3SG.F	her brother(M.SG)	and	1.SG	be.PRS.1PL	well	get.along.PTCP.PL.M

[3] This last restriction avoids spurious ambiguities with all masculine coordinations, using three rules for the four logical combinations.

[4] These observations are only true, in a strict sense, for pure unification formalisms. In systems like Trale (Penn, 2004) the disjunctions between rules could be relegated to attached relational constraints or even better, implicational constraints, as pointed out to us by an anonymous reviewer. However, once a general solution has been found for formalisms without these more elaborate constraints, it certainly helps towards closing the gap in expressiveness between the two competing approaches to HPSG implementation.

‘She, her brother, and I got on well.’

This is not just uneconomical, but the need to multiply out rules for combinations of feature values is something that unification grammar set out to eliminate in order to improve over CFGs. Furthermore, enumeration of combinations will end up obfuscating the linguistically rather clear person and gender hierarchies that govern resolution.

A rather radical approach to feature resolution has been proposed by Sag (2003): instead of unification, he proposed subsumption checks, an approach that bears some similarity to Ingria (1990). This change in the underlying logic of typed feature formalism, however, has not been widely adopted. Furthermore, there is no implementation to date that supports this. Fortunately, as we shall show below, a simple extension to the representation of gender and person features is sufficient to address the issue of feature resolution using a standard unification formalism.

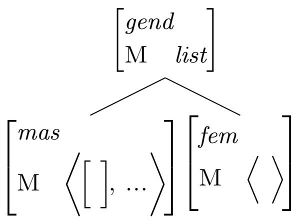
In the following section, we shall therefore develop a theory of feature resolution that crucially distinguishes between a feature itself and its resolution potential. More concretely, we shall enrich the representation of *per* and *gend* values in order to distinguish between e.g. *being* first person or *being* masculine and containing a first person or masculine. We shall show that once the signature of these features is slightly enriched, feature resolution in coordinate structures can be done deterministically. This move leaves untouched standard phrase structure rules targeting entire INDEX values, including all of person, number, and gender features, whereas coordinate structures will have the required flexibility to determine the resolution potential for each feature either holistically, as in the case of semantically motivated number (*pl*), or else in terms of a syntactic resolution strategy.

2 Analysis

2.1 The basic approach: using lists to express existential constraints

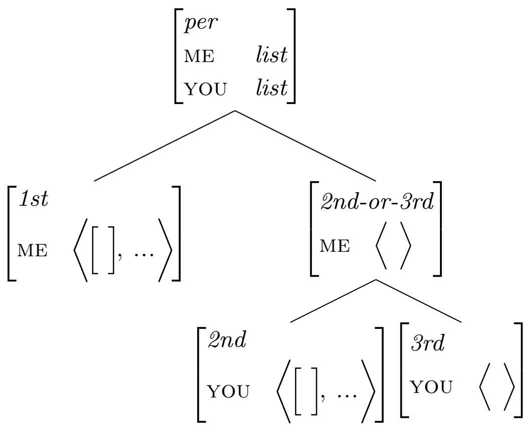
To apply this idea to gender resolution, we first enrich the type *gend* with a feature M taking a list as its value, cf. (20). The type *mas* is then constrained to have a non-empty M list, while the type *fem* is required to have an empty M list

(20) Type constraints for gender



Similarly, we enrich the type *per* with the two list-valued features ME and YOU, cf. (21); the type *1st* requires a non-empty ME list, but does not constrain the YOU list; the type *2nd* requires a non-empty YOU list and an empty ME list; finally the type *3rd* requires both lists to be empty.

(21) Type constraints for person



While this elaborate structure will change nothing with respect to standard projection of INDEX values in head-compositional structures, we gain added flexibility when dealing with coordination: recall that coordinations of individuals (or events) introduce their own INDEX variable, which represents the aggregate. Thus, what needs to be done to capture feature resolution is to determine the PER and GEND values of the group variable on the basis of the respective features of the group members, whereas the NUM value transparently represents plural semantics of the aggregate.

Given our list representation, we expand the coordination rule types of the Grammar Matrix[5] (Drellishak & Bender, 2005) by the constraint in (22) above, enabling us to directly compute the values of agreement features by list concatenation.

[5] The LinGO Grammar Matrix (Bender et al., 2002) is a starter kit for the development of implemented HPSG-style grammars, which has been distilled, originally, from the implemented grammars of English (Copestake & Flickinger, 2000) and Japanese (Siegel & Bender, 2002). On the syntactic side, the Matrix provides type definitions for grammars developed in the spirit of Ginzburg & Sag (2000). With respect to semantics, the Matrix provides compositional principles for semantics construction in Minimal Recursion Semantics (Copestake et al., 2005), ensuring both reversibility and cross-linguistic interoperability.

(22)
$$\begin{bmatrix} \textit{coord} & \\ \text{SS.LOC} & \begin{bmatrix} \text{CAT} & \boxed{c} \\ \text{CONT} & \begin{bmatrix} \text{HOOK.INDEX} & \begin{bmatrix} \textit{ref-ind} & \\ \text{NUM} & \textit{pl} \\ \text{GEND} & \begin{bmatrix} \text{M} & \boxed{1} \oplus \boxed{2} \end{bmatrix} \\ \text{PER} & \begin{bmatrix} \text{ME} & \boxed{3} \oplus \boxed{4} \\ \text{YOU} & \boxed{5} \oplus \boxed{6} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{DTRS} & \left\langle \begin{bmatrix} \text{SS.LOC} & \begin{bmatrix} \text{CAT} & \boxed{c} \\ \text{CONT} & \begin{bmatrix} \text{HOOK.INDEX} & \begin{bmatrix} \textit{ref-ind} & \\ \text{GEND} & \begin{bmatrix} \text{M} & \boxed{1} \end{bmatrix} \\ \text{PER} & \begin{bmatrix} \text{ME} & \boxed{3} \\ \text{YOU} & \boxed{5} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}, \begin{bmatrix} \text{SS.LOC} & \begin{bmatrix} \text{CAT} & \boxed{c} \\ \text{CONT} & \begin{bmatrix} \text{HOOK.INDEX} & \begin{bmatrix} \textit{ref-ind} & \\ \text{GEND} & \begin{bmatrix} \text{M} & \boxed{2} \end{bmatrix} \\ \text{PER} & \begin{bmatrix} \text{ME} & \boxed{4} \\ \text{YOU} & \boxed{6} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

Furthermore, to ensure lists are lexically well-terminated, we constrain lexical types accordingly: (23) defines *basic-noun-lex*, a type from which lexical types fro nouns and pronouns inherit.[6]

(23) *basic-noun-lex* →
$$\begin{bmatrix} \text{SS.LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HD} & \textit{noun} \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{HOOK} & \begin{bmatrix} \text{INDEX} & \begin{bmatrix} \text{GEND} & \begin{bmatrix} \text{M} & \textit{0-1-list} \end{bmatrix} \\ \text{PER} & \begin{bmatrix} \text{ME} & \textit{0-1-list} \\ \text{YOU} & \textit{0-1-list} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

[6] The list type *0-1-list* denotes a list with length of at most 1. It is straightforwardly defined in TDL as follows:

i. *0-1-list* := *list*.

ii. *1-nelist* := *0-1-list* ∧ *nelist* ∧ [REST *elist*].

iii. *elist* := *0-1-list*.

With these basic constraints in place, we will obtain the following results for gender in a coordination of two NPs:

(24) a. $$\begin{bmatrix}\textsc{gend} & m\begin{bmatrix}\text{M}\left\langle\begin{bmatrix}\ \end{bmatrix}\right\rangle\end{bmatrix}\end{bmatrix} + \begin{bmatrix}\textsc{gend} & m\begin{bmatrix}\text{M}\left\langle\begin{bmatrix}\ \end{bmatrix}\right\rangle\end{bmatrix}\end{bmatrix}$$
$$= \begin{bmatrix}\textsc{gend} & \begin{bmatrix}\text{M}\left\langle\begin{bmatrix}\ \end{bmatrix}, \begin{bmatrix}\ \end{bmatrix}\right\rangle\end{bmatrix}\end{bmatrix}$$

b. $$\begin{bmatrix}\textsc{gend} & m\begin{bmatrix}\text{M}\left\langle\begin{bmatrix}\ \end{bmatrix}\right\rangle\end{bmatrix}\end{bmatrix} + \begin{bmatrix}\textsc{gend} & f\begin{bmatrix}\text{M}\left\langle\ \right\rangle\end{bmatrix}\end{bmatrix}$$
$$= \begin{bmatrix}\textsc{gend} & \begin{bmatrix}\text{M}\left\langle\begin{bmatrix}\ \end{bmatrix}\right\rangle\end{bmatrix}\end{bmatrix}$$

c. $$\begin{bmatrix}\textsc{gend} & f\begin{bmatrix}\text{M}\left\langle\ \right\rangle\end{bmatrix}\end{bmatrix} + \begin{bmatrix}\textsc{gend} & m\begin{bmatrix}\text{M}\left\langle\begin{bmatrix}\ \end{bmatrix}\right\rangle\end{bmatrix}\end{bmatrix}$$
$$= \begin{bmatrix}\textsc{gend} & \begin{bmatrix}\text{M}\left\langle\begin{bmatrix}\ \end{bmatrix}\right\rangle\end{bmatrix}\end{bmatrix}$$

d. $$\begin{bmatrix}\textsc{gend} & f\begin{bmatrix}\text{M}\left\langle\ \right\rangle\end{bmatrix}\end{bmatrix} + \begin{bmatrix}\textsc{gend} & f\begin{bmatrix}\text{M}\left\langle\ \right\rangle\end{bmatrix}\end{bmatrix}$$
$$= \begin{bmatrix}\textsc{gend} & \begin{bmatrix}\text{M}\left\langle\ \right\rangle\end{bmatrix}\end{bmatrix}$$

Note that the result of list concatenation is underspecified as to the *gend* type[7], but the resulting non-empty M lists in (24a-c) are only compatible with the type constraints of [GEND *mas*], thus triggering masculine agreement, whereas the empty list in (24d) is only compatible with the type constraint for [GEND *fem*].

2.2 A closer look at list concatenation in TDL[8]

Pure unification formalisms, such as the LinGO LKB (=Linguistic Knowledge Builder; Copestake, 2002) do not recognise lists as primitive data structures,

[7] The LKB, unlike Trale, does not allow inference from features to the types that introduce or constrain them.

[8] TDL (=Type Description Language; Krieger, 1996) was the original description language of the PAGE system (Uszkoreit et al., 1994) and currently is the standard description language for typed feature structure grammar development and runtime platforms within the DELPH-IN collaboration, such as the LKB (Copestake, 2002), Pet (Callmeier, 2000), and Ace (Crysmann & Packard, 2012). Grammars specified in TDL include the English Resource Grammar (Copestake & Flickinger, 2000) among several others, as well as the LinGO Grammar Matrix (Bender et al., 2002).

nor do they provide any specific list operations like, e.g. member/2 or append/3. Rather, lists are encoded as feature structures in FIRST/REST (or HD/TL) notation, which provides easy access for push and pop operations. List concatenation, however, is typically done via difference lists (Clocksin & Mellish, 1981), which maintain an additional pointer to the end of an open list onto which additional lists can be unified.

In a first attempt, we have therefore used difference lists to concatenate the respective person and gender lists, such as GEND.M or PER.ME in coordinate structures. The problem that soon transpired was that difference lists are open lists by necessity, such that any attempt to constrain to a non-empty difference list of indeterminate length (25a) was even successful with empty difference lists (25b), unifying a list element onto the pointer to the end of the list, as shown in (25c).[9]

(25) a. Non-empty difference list underspecified for length

$$\begin{bmatrix} \text{LIST} & \begin{bmatrix} \text{FIRST} & [\;] \\ \text{REST} & \textit{list} \end{bmatrix} \\ \text{LAST} & \textit{list} \end{bmatrix}$$

b. Empty difference list

$$\begin{bmatrix} \text{LIST} & \boxed{l} \\ \text{LAST} & \boxed{l} \end{bmatrix}$$

c. Unifying a non-empty list onto an empty difference list

$$\begin{bmatrix} \text{LIST} & \boxed{l} \begin{bmatrix} \text{FIRST} & [\;] \\ \text{REST} & \textit{list} \end{bmatrix} \\ \text{LAST} & \boxed{l} \end{bmatrix}$$

The only possible solution would have been to terminate this pointer at some point, which proved hard to do in a general and principled fashion. Furthermore, it was difficult to express length constraints on difference lists, as used, e.g. in (23) above.

Fortunately, Emerson (2017) has recently proposed a method to perform concatenation directly on lists, using recursive copying of list members. Following his proposal, we implemented the constraint in (22) using the list definitions in (26), yielding an implementation of the coordinate structure constraint as in (27).

(26) List concatenation in TDL (Emerson, 2017)

a. *list-copy* := *list* ∧

$$\begin{bmatrix} \text{COPY} & \textit{list} \\ \text{NEXT} & \textit{list} \end{bmatrix}$$

[9] Note that checking for cyclic feature structures – a check which the LKB indeed performs — will not provide a solution: once we need to underspecify the length of the list, reentrancy between the REST and LAST cannot be stated.

b. *nelist-copy* := *list-copy* ∧ *nelist* ∧

$$\begin{bmatrix} \text{FIRST} & \boxed{f} \\ \text{REST} & \begin{bmatrix} \textit{list-copy} & \\ \text{COPY} & \boxed{r} \\ \text{NEXT} & \boxed{n} \end{bmatrix} \\ \text{COPY} & \begin{bmatrix} \text{FIRST} & \boxed{f} \\ \text{REST} & \boxed{r} \end{bmatrix} \\ \text{NEXT} & \boxed{n} \end{bmatrix}$$

c. *elist-copy* := *list-copy* ∧ *elist* ∧

$$\begin{bmatrix} \text{COPY} & \boxed{n} \\ \text{NEXT} & \boxed{n} \end{bmatrix}$$

The declarations in (26) faithfully replicate the proposal by Emerson (2017) for purely list-based append: the core idea is to augment a typed feature structure list representation with features for a successor list (NEXT) and a result list COPY. In essence, the FIRST/REST part of the enriched structure represents the first list, the NEXT feature the second list and the COPY holds the resulting concatenation. The type *list-copy* merely introduces the appropriate features (26a). The second clause (26b) recurses over the first list, token-identifying element by element the members of the first list with the members of COPY, the result list. Once the end of the first list has been reached, a subtype of *elist*, (26c) identifies the second list (NEXT) with the result list (COPY). The COPY feature of the entire list will thus consist of the second list, plus the elements of FIRST prepended to it member by member.

In the implementation of feature resolution in French, we consequently use Emerson-style list concatenation, as illustrated in (27). Using gender as an example, the NEXT feature of the GEND.M list of the left conjunct is equated, in coordinate structures, with the GEND.M list of the right conjunct, and the resulting list concatenation in the left daughter's GEND.M.COPY will be token-identical to the GEND.M list on the mother of the coordinating construction.

(27)
$$\begin{bmatrix} \textit{coord} & \\ \textsc{ss.loc} & \begin{bmatrix} \textsc{cat} & \boxed{c} \\ \textsc{cont} & \begin{bmatrix} \textsc{hook.index} & \begin{bmatrix} \textit{ref-ind} & \\ \textsc{num} & \textit{pl} \\ \textsc{gend} & \begin{bmatrix} \textsc{m} & \boxed{1} \end{bmatrix} \\ \textsc{per} & \begin{bmatrix} \textsc{me} & \boxed{3} \\ \textsc{you} & \boxed{5} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \textsc{dtrs} & \left\langle \begin{bmatrix} \textsc{ss.loc} & \begin{bmatrix} \textsc{cat} & \boxed{c} \\ \textsc{cont} & \begin{bmatrix} \textsc{hook.index} & \begin{bmatrix} \textit{ref-ind} & \\ \textsc{gend} & \begin{bmatrix} \textsc{m} & \begin{bmatrix} \textit{list-copy} & \\ \textsc{copy} & \boxed{1} \\ \textsc{new} & \boxed{2} \end{bmatrix} \end{bmatrix} \\ \textsc{per} & \begin{bmatrix} \textsc{me} & \begin{bmatrix} \textit{list-copy} & \\ \textsc{copy} & \boxed{3} \\ \textsc{new} & \boxed{4} \end{bmatrix} \\ \textsc{you} & \begin{bmatrix} \textit{list-copy} & \\ \textsc{copy} & \boxed{5} \\ \textsc{new} & \boxed{6} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}, \begin{bmatrix} \textsc{ss.loc} & \begin{bmatrix} \textsc{cat} & \boxed{c} \\ \textsc{cont} & \begin{bmatrix} \textsc{hook.index} & \begin{bmatrix} \textit{ref-ind} & \\ \textsc{gend} & \begin{bmatrix} \textsc{m} & \boxed{2} \end{bmatrix} \\ \textsc{per} & \begin{bmatrix} \textsc{me} & \boxed{4} \\ \textsc{you} & \boxed{6} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

Fully in parallel to gender, the constraint in (27) equally describes concatenation of person values, broken down to ME and YOU features.

3 Conclusion

In this paper we have discussed resolution of gender and person features in French coordination and suggested to augment the representation of their values with a list-based encoding, and we have shown how this simple extension in the type signature enables us to address the issue of resolution in terms of simple list concatenation. Both the simplicity of the approach and the cross-linguistic recurrence of the phenomenon will make this solution easily applicable to a wider range of languages, both theoretically and within the context of multilingual grammar engineering.

The redundancy between type-based encoding of person and gender features and the list-based one raises the obvious question whether the latter can fully substitute for the former. As for French, we can answer this question in the affirmative, since the lists we propose clearly cover the full inventory of distinctions, yet provide the additional option of distinguishing between exclusively having some property (closed list) and containing some property (open list). Furthermore, feature resolution by concatenation constitutes a simple and uniform mode of composition. The decomposition of person into ME and YOU features bears the further potential to provide an encoding of inclusive and exclusive person distinctions in the plural, as suggested, e.g. by Anderson (1992). Finally, the present approach clearly shows that feature resolution can be done with unification alone, obviating the need for subsumption checks.

Bibliography

An, Aixiu & Anne Abeillé. 2017. Agreement and interpretation of binominals in french. In Stefan Müller (ed.), *Proceedings of the 24th International Conference on Head-Driven Phrase Structure Grammar, University of Kentucky, Lexington*, 26–43. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2017/hpsg2017-an-abeille.pdf.

Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Bender, Emily M., Dan Flickinger & Stephan Oepen. 2002. The grammar matrix: An open-source starter-kit for the rapid development of cross-linguistically consistent broad-coverage precision grammar. In John Carroll, Nelleke Oostdijk & Richard Sutcliffe (eds.), *Proceedings of the Workshop on Grammar Engineering and Evaluation at the 19th International Conference on Computational Linguistics*, 8–14.

Callmeier, Ulrich. 2000. PET — a platform for experimentation with efficient HPSG processing techniques. *Journal of Natural Language Engineering* 6(1). 99–108.

Clocksin, William F. & Christopher S. Mellish. 1981. *Programming in Prolog*. Heidelberg: Springer.

Copestake, Ann. 2002. *Implementing typed feature structure grammars*. Stanford: CSLI Publications.

Copestake, Ann & Dan Flickinger. 2000. An open-source grammar development environment and broad-coverage English grammar using HPSG. In *Proceedings of the Second conference on Language Resources and Evaluation (LREC-2000)*, Athens.

Copestake, Ann, Dan Flickinger, Carl Pollard & Ivan Sag. 2005. Minimal recursion semantics: an introduction. *Research on Language and Computation* 3(4). 281–332.

Copestake, Ann, Alex Lascarides & Dan Flickinger. 2001. An algebra for semantic construction in constraint-based grammars. In *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics, Toulouse, France*, .

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

Crysmann, Berthold & Woodley Packard. 2012. Towards efficient HPSG generation for German, a non-configurational language. In *Proceedings of the 24th International Conference on Computational Linguistics (COLING 2012)*, 695–710. Mumbai, India.

Dalrymple, Mary & Ron Kaplan. 2000. Feature indeterminacy and feature resolution. *Language* 76(4). 759–798.

Drellishak, Scott & Emily M. Bender. 2005. A coordination module for a crosslinguistic grammar resource. In Stefan Müller (ed.), *Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar, Department of Informatics, University of Lisbon*, 108–128. http://cslipublications.stanford.edu/HPSG/2005/drellishak-bender.pdf.

Emerson, Guy. 2017. (Diff-)List Appends in TDL. 13th DELPH-IN summit, Oslo.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations. the form, meaning, and use of english interrogatives*. Stanford: CSLI Publications.

Ingria, Robert J. P. 1990. The limits of unification. In *Proceedings of the twenty-eigth annual meeting of the Association for Computational Linguistics*, 194–204.

King, Tracy Holloway & Mary Dalrymple. 2004. Determiner agreement and noun conjunction. *Journal of Linguistics* 40. 69–104.

Krieger, Hans-Ulrich. 1996. *TDL — a type description language for constraint-based grammars*, vol. 2 Saarbrücken Dissertations in Computational Linguistics and Language Technology. Saarbrücken: DFKI GmbH.

Penn, Gerald. 2004. Balancing clarity and efficiency in typed feature logic through delaying. In *Proceedings of the 42nd Meeting of the Association for Computational Linguistics (ACL'04)*, 239–246. Barcelona, Spain. doi: 10.3115/1218955.1218986. `http://www.aclweb.org/anthology/P04-1031`.

Sag, Ivan A. 2003. Coordination and underspecification. In Jong-Bok Kim & Stephen Wechsler (eds.), *Proceedings of the 9th International Conference on Head-driven Phrase Structure Grammar, Kyung Hee University, Seoul, 5–7 August, 2002*, 267–291. Stanford: CSLI Publications.

Siegel, Melanie & Emily M. Bender. 2002. Efficient deep processing of Japanese. In *Proceedings of the 3rd workshop on asian language resources and international standardization. coling 2002 post-conference workshop*, Taipei.

Uszkoreit, Hans, Rolf Backofen, Stephan Busemann, Abdel Kader Diagne, Elizabeth Hinkelman, Walter Kasper, Bernd Kiefer, Hans-Ulrich Krieger, Klaus Netter, Günter Neumann, Stephan Oepen & Stephen P. Spackman. 1994. Disco - an HPSG-based NLP system and its application for appointment scheduling. In *Proceedings of the 15th International Conference on Computational Linguistics (COLING'94), august 5-9*, vol. 1, 436–440. Kyoto, Japan.

Zwicky, Arnold. 1977. Hierarchies of person. In *Chicago linguistic society*, vol. 13, 714–733.