Social gender and agreement with gender-fair forms: an LFG account

Yanis da Cunha
Laboratoire de Linguistique Formelle, Université Paris Cité

Languages, with or without grammatical gender, are able to refer to social gender, an extra-linguistic categorization of humans (Ackerman 2019). In particular, grammar allows speakers to express a referent’s social gender through nouns, pronouns or agreement markers. Despite its pervasive presence across languages, reference to social gender has not yet been fully accounted for in formal approaches. Moreover, the rise of interest for gender-fair forms, such as non-binary pronouns (English singular they, Bodine 1975, or French iel, Greco 2015) or inclusive forms (les étudiant-es ‘the students’), poses several challenges for a formal account of social gender reference (Bjorkman 2017).

The present work addresses these questions and sketches a formal analysis of social gender using LFG. We first review different types of relations between grammatical and social gender for nouns. Then, building on the description-by-analysis approach to the syntax-semantics interface (Andrews 2008), we suggest a new formalization of the social gender interpretation of grammatical gender. Finally, we turn to the analysis of agreement and interpretation of gender-fair forms, using new French data about inclusive forms and the neopronoun iel.

1 Expressing social gender through grammatical gender

We start our analysis by reviewing a range of ways grammatical gender languages can use grammatical gender (g-gender) to convey a social gender (s-gender) classification of human referents.

First, the most common case is represented by gender-iconic nouns (Bonami & Boyé 2019). These are human-denoting nouns (mainly occupational, kinship and title nouns) for which g-gender is straightforwardly interpreted as s-gender (Corbett 1991). A feminine iconic noun is female-referring while a masculine iconic noun is male-referring. The specification of g-gender can be due either to lexical meaning (“definitional gender” Kreiner, Sturt & Garrod 2008) or to morphology (gender-iconic pairs Bonami & Boyé 2019). French examples in (1-2) illustrate these two possibilities, respectively.

(1)  reine ‘queen’, mère ‘mother’, frère ‘brother’
(2)  boulanger/boulangère ‘baker’, chanteur/chanteuse ‘singer’

The second case corresponds to common gender nouns (3-4). For these nouns, neither semantics nor morphology are cues for g-gender. Hence, they are unspecified: they can refer to any s-gender. There is often an a priori bias as to which s-gender is more likely to be referred to, known as stereotypical gender (Kreiner, Sturt & Garrod 2008, Stahlberg et al. 2007), but we leave this aspect aside. Importantly, although lexically unspecified, these nouns often acquire a s-gender interpretation in syntax, where agreement markers are made available. Concretely, feminine agreement markers trigger a female interpretation and masculine agreement markers trigger a male interpretation. The French examples in (3) show a s-gender specification by noun-determiner and noun-adjective agreement, while the Russian examples in (4) show it by subject-verb agreement (Asarina 2009).

(3)  a. La ministre est intelligente
     the.f minister is smart.f
     ‘The (female) minister is smart’
   b. Le ministre est intelligent
     the.M minister is smart.M
     ‘The (male) minister is smart’
These examples are an illustration of the Agreement Marking Principle (Wechsler 2011), which states that when a controller is unspecified for some grammatical feature (e.g. g-gender in common gender nouns), target agreement features give rise to a semantic interpretation. Although this principle captures the empirical data in (3-4), we will show there is no need to integrate it in a formal analysis if a general meaning is given to g-gender features.

The last case to account for is represented by non-iconic nouns. These nouns challenge any analysis of g-gender meaning, as their g-gender does not correspond to s-gender. Two sub-types can be distinguished. Generic nouns have a fixed g-gender but can refer to any s-gender (5a). Although feminine, the French noun recrue ‘recruit’ can refer to both women and men (5b). Mismatched nouns, by contrast, have a s-gender different from what would be predicted by their g-gender. For example, German Mädchen ‘girl. N’ is a female-referring neuter noun while Irish cailín ‘girl. M’ is a female-referring masculine noun.

2 A description-by-analysis approach to grammatical gender meaning

We use the Glue framework for the syntax-semantics interface in LFG (Dalrymple 1999). Its standard approach is co-description: lexical entries contain both functional equations (the first two lines in 6), representing syntactic information, and meaning constructors (the last two lines in 6), which define meanings and ensure their semantic composition. We use first-order Glue to simplify the meaning constructors (Kokkonidis 2008). In (6), the meaning of reine ‘reine’ is represented by a gender-general predicate (monarch) of type <et>. The meaning contribution of g-gender is represented as a restricting modifier of type <<et>,<et>>, which we take as the instruction for s-gender categorization (Cooper 2013: p. 182). We use the partial operator $\partial$ to account for its presupposed status. Composition then gives the following meaning: $\lambda x.\text{monarch}(x) \land \partial \text{female}(x) : e_1 \rightarrow o t_1$.

This standard approach seems however unsatisfactory since in (6), the female meaning and the feminine grammatical feature appear unrelated: their co-occurrence seems to be a mere accident. But, we have shown there is a generalization to capture, as g-gender is a cue for s-gender. As such, co-description is not well-suited to express a generalization about relations between lexical descriptions. One possibility is to use templates (Dalrymple, Kaplan & King 2004), but this approach leaves complications for the treatment of common gender nouns.

We thus take the description-by-analysis (DBA) approach developed by Andrews (2007, 2008). In this approach, instead of specifying each lexical entry redundantly by functional descriptions and meaning constructors, we explicitly formalize the meaning contribution of grammatical features. For example, the co-occurrence of (f GENDER) = F and (f HUMAN) = + in an f-structure $f$ gives rise to a female interpretation. To express this generalization, we use Semantic Lexical Entries (SLE) which define a correspondence between f-structure features and meaning constructors. The SLE in (7) delivers a female meaning constructor for feminine human nouns.
With this SLE and a SLE interpreting the \texttt{PRED} feature (8a), the Morphological Lexical Entry in (6) reduces to (8b). The main consequence is that this lexical entry only contains functional descriptions, and semantic interpretation of g-gender is left to the more general rule (7).

\begin{align*}
\text{(7)} \quad (f \textsc{gender}) & = F \\
\quad (f \textsc{human}) & = +
\end{align*}

\[ \iff \lambda P \lambda x. P(x) \land \partial \text{female}(x) : (e_f \rightarrow t_f) \rightarrow (e_f \rightarrow t_f) \]

S-gender specification of common gender nouns is now made easy to account for: although these nouns are lexically unspecified, a g-gender feature can be added by agreement markers. This feature is then interpreted by the g-gender to s-gender SLE (7). This mechanism basically implements the Agreement Marker Principle (Wechsler 2011) without stipulating it separately. Example (9) gives a simplified LFG analysis of s-gender inference by g-gender agreement. So called “semantic agreement” just reduces here to the usual interpretation of a feature added by an agreement target.

\begin{align*}
\text{(8) a.} \quad (f \textsc{pred}) & = \text{'queen'} \iff \lambda x. \text{monarch}(x) : e_f \rightarrow t_f \\
\quad \text{(f \textsc{gender})} & = F \\
\text{(8) b.} \quad \text{reine N (} \uparrow \text{pred}) & = \text{'queen'} \\
\quad \text{(f \textsc{gender})} & = F
\end{align*}

Finally, non-iconic nouns show a s-/g-gender mismatch: they bear a g-gender feature active in syntax, but inactive in semantics. This case challenges our analysis, since we posited a general SLE that interprets g-gender on human nouns as s-gender (7). Crucial for the analysis is that SLEs are resource-sensitive: each feature of an f-structure can only be used once to provide meaning constructors (otherwise, we would face a surplus of resources). So, non-iconic nouns require specific SLEs that conjointly interpret their \texttt{pred} with their \texttt{gender} feature (10).

\(1\) Consequently, the \texttt{gender} feature is made inaccessible for the more general SLE (7). Idiosyncratic meanings for non-iconic nouns can then be specified. The feminine French noun \textit{recrue} ‘recruit’ is interpreted without s-gender (10a) while the neuter German noun \textit{Mädchen} ‘girl’ is interpreted with a female s-gender (10b).

\begin{align*}
\text{(10) a.} \quad (f \textsc{pred}) & = \text{'recrue'} \\
\quad (f \textsc{gender}) & = F \\
\iff \lambda x. \text{recruit}(x) : e_f \rightarrow t_f \\
\text{b.} \quad (f \textsc{pred}) & = \text{'mädchen'} \\
\quad (f \textsc{gender}) & = I \\
\iff \lambda x. \text{child}(x) \land \partial \text{female}(x) : e_f \rightarrow t_f
\end{align*}

This proposal makes it possible to account for “semantic derogation” (Schulz 1975), where (feminine) g-gender gets idiomatic meanings when used with certain nouns (\textit{professionnelle} ‘female professional’ or ‘prostitute’). For non-iconic nouns, it is also useful to split agreement features (Wechsler 2011), with \texttt{concord} for attributive agreement, grammatically required, and \texttt{index} for predicative/anaphoric agreement, which can flexibly trigger a s-gender interpretation (11).

\begin{align*}
\text{(11) a.} \quad \text{La recrue arrive. Il est très compétent. ‘The.F recruit. F arrives. He.M is very competent.M’} \\
\quad \text{b.} \quad \text{Sa Majesté est généreux. ‘His.F Majesty. F is generous.M (eg. to a king)’}
\end{align*}

\(1\) See Andrews (2008) for a similar analysis of \textit{pluralia tantum} nouns, which exhibit a mismatch for number features.
3 Inclusive forms in French

Inclusive forms represent a wide range of strategies that allow speakers to adopt gender-fair attitudes through language (Abbou et al. 2018). These strategies mainly aim at escaping the s-gender inferences made by g-gender, to balance gender reference, make women more visible (Xiao, Strickland & Peperkamp 2023) or include non-binary identities. For example, male interpretation of occupational nouns can be avoided by using feminine forms (new or not) such as autrice instead of auteur ‘author’. Important for our purpose is the use of graphic signs to abbreviate double mentions of feminine and masculine forms as shown in (12).

(12) a. Les étudiantes et étudiants ‘the (female) students and (male) students’
   b. Les étudiant-es ‘the students (any gender)’

These forms are interesting because they exhibit a mixed behavior regarding the previous categories introduced. On one hand, they share with common gender nouns the compatibility with any agreement marker. They can co-occur with masculine, feminine or other inclusive forms (13). Their agreement is however not unconstrained, as they are not used with incompatible g-genders (14). On the other hand, they resemble generic nouns in being unspecified for s-gender. In example (15), enseignant-es ‘teachers’ implies a mixed group containing men and women. Stereotypical gender is here likely to play a role in the interpretation or the choice of agreement (Richy & Burnett 2021).

(13) a. je laisse de bons souvenirs à certains étudiant-es (Twitter, 10/20/2019)
   ‘I leave good memories to some M students’
   b. la […] majorité des étudiant-es sont restés chez eux.M (Twitter, 01/13/2021)
   ‘the majority of students stayed M home’
   c. […] ces futures enseignant-es restent. (Twitter, 04/12/2019)
   ‘These future F teachers stay here’
   d. 30% des étudiant-es ne sont pas aidé-es par leurs parents (Twitter, 05/03/2023)
   ‘30% of the student are not helped by their parents’

(14) a. * Certains nouvelles étudiant-es
   ‘Some M new F students’
   ‘The new .M teachers are competent F’

(15) 82% des enseignant-es sont des femmes en primaire. (Twitter, 05/03/2017)
   ‘82% of teachers are women in primary school’

An account of these data will thus build on the analysis of both common gender nouns and generic nouns. As common gender nouns, inclusive forms are left unspecified for g-gender. This makes them compatible with masculine or feminine agreement markers, but not both at the same time (as in regular agreement). However, as generic nouns, their eventual g-gender is not interpreted. We use a FORM feature to trigger the SLE in (16a) consuming the GENDER feature (as was done by PRED in 10a), whatever its value. The added meaning constructor does not specify any s-gender: it is an identity function \( \lambda P \lambda x. P(x) \). If the inclusive noun has not been specified for g-gender, the standard interpretation of its PRED (16b) is enough to give it a gender-general meaning.

(16) a. \( (f_{\text{FORM}}) = \text{inclusive}, (f_{\text{HUMAN}}) = + \) \( (f_{\text{GENDER}}) \)

   \[ \iff \lambda P \lambda x. P(x) : (e_f \circ o t_f) \circ (e_f \circ o t_f) \]

   b. \( (f_{\text{PRED}}) = \text{‘student’} \iff \lambda x. \text{student}(x) : e_f \circ o t_f \)

---

\(^2\text{We show here hyphen inclusive forms (-) but more varied signs are found in use (Burnett & Pozniak 2021).}\)
Importantly, we have to give priority to SLE (16a) over the more general SLE (7) when providing meaning constructors for an f-structure. As suggested by a reviewer, this can be achieved by stating that SLEs including a FORM feature have priority over other SLEs. This way, we can ensure that étudiant-es does not receive a s-gender specification even if agreement markers are present, like certains.m in (16b).

The analysis we suggested above can readily extend to the neopronoun iel(s) ‘they’. It shares some properties with inclusive forms, such as compatibility with any agreement marker (17). However, beside its gender-general use, it can also refer to a person whose s-gender is non-binary (18). This is mostly the case for specific singular reference (18a), but it is also found in plural (18b).

As we took a meaning-based approach to s-gender categorization, adding a nonbinary predicate does not imply any change for the grammar of g-gender. A possible SLE for interpreting iel(s) is presented in (19). It takes the entity denoted by the pronoun (e_f) and either leaves it unchanged (gender-general use) or restricts it to be non-binary (nonbinary(x)).

Conclusion We sketched a new analysis of g-gender and s-gender in French framed in a DBA approach to the syntax-semantics interface in LFG (Andrews 2007, 2008). We showed how s-gender can be inferred from g-gender either lexically (iconic nouns) or syntactically (common gender nouns), straightforwardly accounting for the Agreement Marking Principle (Wechsler 2011). Exceptions to the general g-gender meaning are treated by a specific interpretation of PRED and GENDER features combination. This analysis has been applied to inclusive forms and neopronouns in French, with successful results. However several aspects remain to be taken into account, such as stereotypical gender, masculine generics, gender shift by anaphora or coordinations. These phenomena will probably require a more complex representation of the structure and interpretation of g-gender features. Finally the provided data, especially for inclusive forms in French, show a great degree of variation across speakers, depending on attitudes about sexism in language, identities or communicative goals. But, we hope to have provided a promising framework to deal with these data.
References


