

Constraining AGReement in the SRG

for the 20th DELPH-IN summit in Olomouc

Olga Zamaraeva, Lorena S. Allegue
Department of Informatics/CITIC, Department of Philology
Universidade da Coruña

July 3 2024

AGReement in the SRG

Constraining
AGReement in the
SRG

for the 20th
DELPH-IN summit
in Olomouc

Intro

- ▶ Context: inheriting a large grammar where many decisions have been made
- ▶ Agreement: a phenomenon that is related to CONT|RELS and INDEX
 - ▶ ...and as such is an analysis difficult to change
 - ▶ delicate mechanics in RELS
 - ▶ *event* vs. *referent* semantics
 - ▶ can be hard to copy up large portions of structure
- ▶ ...if we are done quickly talking about AGR, let's talk about working with large grammars more generally

Spanish agreement in the NP/AP

Constraining
AGReement in the
SRG

for the 20th
DELPH-IN summit
in Olomouc

Intro

Spanish nouns have gender, and modifiers and specifiers usually exhibit morphological agreement with the inherent gender of the noun

- (1) Este público es maravilloso.
this.MASC.SG public:MASC.sg COP wonderful.MASC.SG
'This audience is wonderful.' (spa; TIBIDABO 5462)

Adjectives cannot insist on referential semantics for the modificand:

- (2) Hacer ejercicio es bueno?
do.INFINITIVE exercise COP good.MASC.SG
'Is it good to exercise?' (spa; TIBIDABO corpus item
716)

Modifying a noun:

- (3) Era una fila
COP INDEF.FEM.SG vein:FEM.SG
muerta, sin uranio.
die.PART.PST.FEM.SG without uranuim:MASC.SG
'It was a dead vein, with no uranium.' [spa; TIBIDABO
2010]

Complement of the copula, agreeing with the subject (ARG2 of the verb):

- (4) Su hijo también fue
 POSS son:MASC.SG also COP
 asesinado
 assassinate.PART.PST.SG.MASC
 (*a/*os/*as)
 (*SG.FEM/*PL.MASC/*PL.FEM)
 'Their son was also assassinated.' (spa; TIBIDABO
 13784)

Spanish past participles: SUBJ, COMPS, ARG1 and ARG2

Constraining
AGReement in the
SRG
for the 20th
DELPH-IN summit
in Olomouc

Intro

- (5) La operación fue
DEF.SG.FEM operation:FEM.SG COP
realizada por un
realize.PART.PST.SG.FEM by INDEF.MASC.SG
experto.
expert:MASC.SG
'The operation was performed by an expert.' [spa;
TIBIDABO item 11192]

Spanish perfective construction with HABER

Constraining
AGReement in the
SRG

for the 20th
DELPH-IN summit
in Olomouc

Intro

Perfective constructions, with auxiliary HABER, not agreeing with the subject:

- (6) Todos habían ganado
all.MASC.PL AUX.IPFV.3PL win.PART.PST.MASC.SG
(*os).
(*PPART.MASC.PL)
'Everyone had won.' (spa; TIBIDABO item 12448)

- ▶ Semantic and morphosyntactic (Pollard and Sag, 1994)
 - ▶ LOCAL|CONT|HOOK|INDEX|PNG|GEN
 - ▶ LOCAL|(CAT|HEAD)|AGR|PNG|GEN

AGReement in the SRG

Constraining
AGReement in the
SRG

for the 20th
DELPH-IN summit
in Olomouc

- ▶ Underspecified in the lexicon

Intro

AGReement in the SRG

Constraining
AGReement in the
SRG

for the 20th
DELPH-IN summit
in Olomouc

Intro

- ▶ Underspecified in the lexicon
- ▶ External analyzer (Freeling 4.1) assigns tags
 - ▶ e.g. AQ0MS00 (qualitative adjective masculine singular)

- ▶ Underspecified in the lexicon
- ▶ External analyzer (Freeling 4.1) assigns tags
 - ▶ e.g. AQ0MS00 (qualitative adjective masculine singular)
- ▶ Lexical rules associated with the tag constrain the gender values

- ▶ Underspecified in the lexicon
- ▶ External analyzer (Freeling 4.1) assigns tags
 - ▶ e.g. AQ0MS00 (qualitative adjective masculine singular)
- ▶ Lexical rules associated with the tag constrain the gender values
- ▶ The lexical rules use a compact hierarchy of rules like *sg-illr* and constrain the values at LOCAL|AGR

- ▶ Underspecified in the lexicon
- ▶ External analyzer (Freeling 4.1) assigns tags
 - ▶ e.g. AQ0MS00 (qualitative adjective masculine singular)
- ▶ Lexical rules associated with the tag constrain the gender values
- ▶ The lexical rules use a compact hierarchy of rules like *sg-illr* and constrain the values at LOCAL|AGR
 - ▶ including for adjectives and verbs

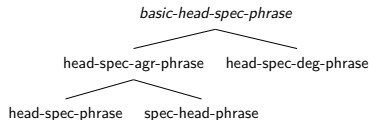
- ▶ Underspecified in the lexicon
- ▶ External analyzer (Freeling 4.1) assigns tags
 - ▶ e.g. AQ0MS00 (qualitative adjective masculine singular)
- ▶ Lexical rules associated with the tag constrain the gender values
- ▶ The lexical rules use a compact hierarchy of rules like *sg-illr* and constrain the values at LOCAL|AGR
 - ▶ including for adjectives and verbs
 - ▶ lexical entries may associate the AGR values with e.g. MOD or SUBJ

SRG: AGR in phrase structure rules

Constraining
AGReement in the
SRG

for the 20th
DELPH-IN summit
in Olomouc

- ▶ AGR was underspecified:
 - ▶ optcomp
 - ▶ copulas
 - ▶ participle inflectional and derivational rules
 - ▶ “adjpart” phrase (participle to modify noun)
 - ▶ head-specifier rules
 - ▶ “spec-degree” has *event* semantics, others *ref*
- ▶ But why is it needed there?
 - ▶ Look in the LKB: *La gata es negro.* ('The cat is black.')



head-spec-agr-phrase
SYNSEM | LOCAL | AGR [1]
HEAD-DTR | SYNSEM | LOCAL | AGR [1]

Modifying a noun:

(7) Era una fila
COP INDEF.FEM.SG vein:FEM.SG
muerta, sin uranio.
die.PART.PST.FEM.SG without uranuim:MASC.SG
'It was a dead vein, with no uranium.' [spa; TIBIDABO
2010]

Complement of the copula, agreeing with the subject (ARG2 of the verb):

- (8) Su hijo también fue
 POSS son:MASC.SG also COP
 asesinado
 assassinate.PART.PST.SG.MASC
 (*a/*os/*as)
 (*SG.FEM/*PL.MASC/*PL.FEM)
 'Their son was also assassinated.' (spa; TIBIDABO
 13784)

Perfective constructions, with auxiliary HABER, not agreeing with the subject:

- (9) Todos habían ganado
all.MASC.PL AUX.IPFV.3PL win.PART.PST.MASC.SG
(*os).
(*PPART.MASC.PL)
'Everyone had won.' (spa; TIBIDABO item 12448)

- ▶ Verbs link LOCAL|AGR to SUBJ's INDEX
- ▶ but a past participle cannot then be formed from such a verb
 - ▶ (unless AGR was underspecified and grammar overgenerates)

(10) La operación fue
DEF.SG.FEM operation:FEM.SG COP
realizada por un
realize.PART.PST.SG.FEM by INDEF.MASC.SG
experto.
expert:MASC.SG
'The operation was performed by an expert.' [spa;
TIBIDABO item 11192]

Intro

- ▶ AGR info comes in with the Freeling tag
 - ▶ preterminal node
- ▶ that infl-lex-rule then cannot really be a *verb-itr*
- ▶ a complex hierarchy of lex-rules copying up structure bit by bit
 - ▶ “participle”
 - ▶ not directly used(?)
 - ▶ passive participle (with ARG1 and ARG2)
 - ▶ passive participle-looking things
 - ▶ unaccusative verbs (empty COMPS)
 - ▶ perfective (flag feature)

AGReement in the SRG: items for discussion

Constraining
AGReement in the
SRG

for the 20th
DELPH-IN summit
in Olomouc

Intro

- ▶ Where does AGR go on nouns, adjectives, verbs?
- ▶ How “early” can it be constrained?
 - ▶ e.g. in participles

General issues with large grammars

Constraining
AGReement in the
SRG
for the 20th
DELPH-IN summit
in Olomouc

Intro

- ▶ Treebanking top-1 profiles
 - ▶ incr tsdb()? acetools?..
- ▶ Tracking regressions and versions
- ▶ How much regression testing is enough?
- ▶ Designing the pipeline