Spanish Resource
Grammar updates

for the 20th
DELPH-IN summit
in Olomouc

Intro

non-mal SRG

mal SRG

Conclusion

Spanish Resource Grammar updates

for the 20th DELPH-IN summit in Olomouc

Olga Zamaraeva, Lorena Suárez Allegue, Carlos Gómez-Rodríguez
Department of Informatics/CITIC, Department of Philology
Universidade da Coruña

July 2 2024

Spanish Resource Grammar v. 2023

- LREC/COLING 2024
  https://aclanthology.org/2024.lrec-main.1312/

SRG updates summary: Two parts

- Last year: Freeling v.4.1

SRG updates summary: Two parts

- Last year: Freeling v.4.1
- Agreement in the SRG
  - Identifying underspecified PNG values
  - Constraining PNG (GEN) across the grammar
  - Evaluating the effects

SRG updates summary: Two parts

- Last year: Freeling v.4.1
- Agreement in the SRG
  - Identifying underspecified PNG values
  - Constraining PNG (GEN) across the grammar
  - Evaluating the effects
- SRG-mal
  - Context: MSCA project: SRG for grammar checking

SRG updates summary: Two parts

- Last year: Freeling v.4.1
- Agreement in the SRG
  - Identifying underspecified PNG values
  - Constraining PNG (GEN) across the grammar
  - Evaluating the effects
- SRG-mal
  - Context: MSCA project: SRG for grammar checking
  - Learner treebanks:
    - lead to L2 RQs
    - help find points of overgeneration

Agreement in the (non-mal) SRG

- Discussion on Wednesday
- Setup:

Agreement in the (non-mal) SRG

Spanish Resource
Grammar updates

for the 20th
DELPH-IN summit
in Olomouc

Intro

non-mal SRG

mal SRG

Conclusion

- Discussion on Wednesday
- Setup:
  - Run SRG on learner sentences with NP/AP gender agreement errors

Agreement in the (non-mal) SRG

Spanish Resource
Grammar updates

for the 20th
DELPH-IN summit
in Olomouc

Intro

non-mal SRG

mal SRG

Conclusion

- ▶ Discussion on Wednesday
- ▶ Setup:
  - ▶ Run SRG on learner sentences with NP/AP gender agreement errors
  - ▶ Look at parsed sentences

Agreement in the (non-mal) SRG

- Discussion on Wednesday
- Setup:
  - Run SRG on learner sentences with NP/AP gender agreement errors
  - Look at parsed sentences
  - Identify missing PNG constraints

Agreement in the (non-mal) SRG

▶ Discussion on Wednesday
▶ Setup:
  ▶ Run SRG on learner sentences with NP/AP gender agreement errors
  ▶ Look at parsed sentences
  ▶ Identify missing PNG constraints
  ▶ Add constraints

Agreement in the (non-mal) SRG

- Discussion on Wednesday
- Setup:
  - Run SRG on learner sentences with NP/AP gender agreement errors
  - Look at parsed sentences
  - Identify missing PNG constraints
  - Add constraints
  - Test on AnCoRa/TIBIDABO

Agreement in the (non-mal) SRG

- Discussion on Wednesday
- Setup:
    - Run SRG on learner sentences with NP/AP gender agreement errors
    - Look at parsed sentences
    - Identify missing PNG constraints
    - Add constraints
    - Test on AnCoRa/TIBIDABO
    - (repeat)

non-mal SRG: selected AGR changes

Spanish Resource
Grammar updates

for the 20th
DELPH-IN summit
in Olomouc

Intro

non-mal SRG

mal SRG

Conclusion

▶ Where AGR was missing:
  ▶ optcomp
  ▶ copulas
  ▶ participle inflectional and derivational rules
  ▶ "adjpart" phrase (participle to modify noun)
  ▶ head-specifier rules

$$\begin{bmatrix} optcomp\text{-}n \\ \text{SYNSEM|LOCAL|AGR } \boxed{1} \\ \text{HEAD-DTR|SYNSEM|LOCAL|AGR } \boxed{1} \end{bmatrix}$$

Other changes in the SRG

- ...more Freeling interface tweaks
  - both frustrating and gratifying:
    - flexible
    - ad-hoc
- Some illegal DTR|RELS constraints removed
  - many more remain
  - packing impacted
- A few lexical entries reassigned type

(non-mal) SRG release 0.3.5

Spanish Resource
Grammar updates

for the 20th
DELPH-IN summit
in Olomouc

Intro

non-mal SRG

mal SRG

Conclusion

Overgeneration:

corpus	0.3.4	0.3.5
agr	0.75	0.0

Accuracy:

corpus	0.3.4	0.3.5
agr	1.0	1.0
mrs	0.81	0.95
tbdb01	1.0	1.0
tbdb02	0.93	0.94
tbdb03	0.88	0.91
tbdb04	0.86	0.89
tbdb05	0.86	0.89

corpus	0.3.4	0.3.5
tbdb06	0.82	0.88
tbdb07	0.76	0.86
tbdb08	0.82	0.81
tbdb09	0.77	0.79
tbdb10	0.76	0.75
tbdb11*	0.50	0.53
tbdb12*	0.65	0.64

*After re-verification, 77-79%

(non-mal) SRG release 0.3.5 parsing speed

Performance (assessed with tsdb++, not sure how reliably):

corpus	time compared to 0.3.4	edges compared to 0.3.4
mrs	-31%	-30%
tbdb01	-26%	-15%
tbdb02	-12%	-11%
tbdb03	-31%	-26%
tbdb04	-43%	-34%
tbdb05	-46%	-34%
tbdb06	-68%	-48%
tbdb07	-76%	-59%
tbdb08	-67%	-56%
tbdb09	-75%	-65%
tbdb10	-87%	-68%
tbdb11	-72%	-65%
tbdb12	-204%	-38%

non-mal SRG: Remaining problems

Spanish Resource
Grammar updates

for the 20th
DELPH-IN summit
in Olomouc

Intro

non-mal SRG

mal SRG

Conclusion

- ▶ Parser limitations (RAM)
- ▶ Parse ranking model is old
- ▶ Analyses:
  - ▶ Clitics
  - ▶ Ellipsis
  - ▶ Problems in coordination
  - ▶ Still more issues with Freeling

mal SRG

Spanish Resource
Grammar updates

for the 20th
DELPH-IN summit
in Olomouc

Intro

non-mal SRG

mal SRG

Conclusion

- ▶ Focus: Gender agreement in NP, AP
- ▶ Method: Inflectional rules (Freeling tags)
- ▶ Dev corpus: COWSL2H (Yamada et al. 2020)
- ▶ Research corpus: CEDEL2
- ▶ RQ area:
  - ▶ L1 effects on gender agreement errors (num, type)
- ▶ Demo: iTell (working locally! thanks, Luis!)
- ▶ Funding: MSCA grant No 101063104

UNIVERSIDADE DA CORUÑA   citic

mal SRG: current design

▶ Focus: Gender agreement in NP/AP

mal SRG: current design

Spanish Resource
Grammar updates

for the 20th
DELPH-IN summit
in Olomouc

Intro
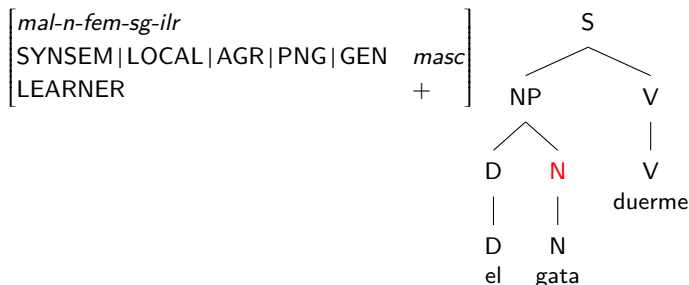
non-mal SRG

mal SRG

Conclusion

- ▶ Focus: Gender agreement in NP/AP
- ▶ Method: Inflectional rules:
  - ▶ In the SRG, gender is not specified in the lexicon
  - ▶ Freeling probabilistically assigns each word a tag
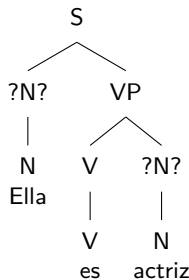 - ▶ The tag corresponds to a lexical rule specifying gender

$$\begin{bmatrix} mal\text{-}n\text{-}fem\text{-}sg\text{-}ilr \\ \text{SYNSEM}|\text{LOCAL}|\text{AGR}|\text{PNG}|\text{GEN} & masc \\ \text{LEARNER} & + \end{bmatrix}$$

```
        S
       / \
      NP    V
     /  \    |
    D    N   V
    |    |  duerme
    D    N
    el  gata
```

mal SRG: current design

Spanish Resource
Grammar updates

for the 20th
DELPH-IN summit
in Olomouc

Intro

non-mal SRG

mal SRG

Conclusion

- ▶ mal SRG: doubles the tags to allow *masc* and *fem* for each noun or adjective or participle
  - ▶ Simple but increases ambiguity/parsing time
  - ▶ Theoretically: which words should have mal-infl-rules? Nouns? Adjectives? Pronouns? Determiners?

$$\begin{bmatrix} \textit{mal-n-fem-sg-ilr} \\ \text{SYNSEM|LOCAL|AGR|PNG|GEN} & \textit{masc} \\ \text{LEARNER} & + \end{bmatrix}$$

$$\begin{bmatrix} \textit{n-fem-sg-ilr} \\ \text{SYNSEM|LOCAL|AGR|PNG|GEN} & \textit{fem} \\ \text{LEARNER} & - \end{bmatrix}$$

```
            S
          /   \
       ?N?     VP
        |     /   \
        N    V    ?N?
       Ella  |     |
             V     N
             es   actriz
```

mal SRG accuracy on COWSL2H

- COWSL2H (Yamada et al. 2020) corpus of written Spanish of L2 and heritage speakers, from UC Davis
- Partially annotated manually for gender agr. errors

mal SRG accuracy on COWSL2H

Spanish Resource
Grammar updates

for the 20th
DELPH-IN summit
in Olomouc

Intro

non-mal SRG

mal SRG

Conclusion

- COWSL2H (Yamada et al. 2020) corpus of written Spanish of L2 and heritage speakers, from UC Davis
- Partially annotated manually for gender agr. errors
- We took the sentences with such manual annotations (that had no other errors)
- ...reserved 66 of them for testing

mal SRG accuracy on COWSL2H

Spanish Resource
Grammar updates

for the 20th
DELPH-IN summit
in Olomouc

Intro

non-mal SRG

mal SRG

Conclusion

- COWSL2H (Yamada et al. 2020) corpus of written Spanish of L2 and heritage speakers, from UC Davis
- Partially annotated manually for gender agr. errors
- We took the sentences with such manual annotations (that had no other errors)
- ...reserved 66 of them for testing
  - **recall:** how many of the annotated errors we catch with the appropriate mal rule
  - **precision:** how many of our use of mal rules are correct

precision	recall (top 1)	recall (top 5)
0.84	0.58	0.72

mal SRG: Precision on well-formed sentences

Spanish Resource
Grammar updates

for the 20th
DELPH-IN summit
in Olomouc

Intro

non-mal SRG

mal SRG

Conclusion

- The grammar should be "precise"
- Yet, without a large treebank and a statistically trained model, mal-rules are happy to apply in various cases
  - *Es actriz* ('is actress')
  - no article or adjective, which gender to use?
- For now:
  - Parse with normal SRG
  - Parse the items with 0 results with mal SRG

parsed with mal SRG   **actual error in the sentence?**
78/7289 (0.0107)   Yes, though not the majority

mal SRG: Research Questions

Spanish Resource Grammar updates

for the 20th DELPH-IN summit in Olomouc

Intro

non-mal SRG

mal SRG

Conclusion

- ▶ **If** we had a highly precise mal SRG, we could run it on CEDEL2 (many L1s) and ask:
    - ▶ How does L1 affect the number and type of mistakes?
        - ▶ E.g. does L1 Russian help with gender compared with L1 English?
        - ▶ Does L1 English help with structures which have an article?
    - ▶ RQs such as above are typically pursued using small samples
    - ▶ Corpus approaches remain shallow
        - ▶ E.g. regex, no syntactic context

mal SRG: Research Questions

Spanish Resource Grammar updates

for the 20th DELPH-IN summit in Olomouc

Intro

non-mal SRG

mal SRG

Conclusion

- **If** we had a highly precise mal SRG, we could run it on CEDEL2 (many L1s) and ask:
  - How does L1 affect the number and type of mistakes?
    - E.g. does L1 Russian help with gender compared with L1 English?
    - Does L1 English help with structures which have an article?
  - RQs such as above are typically pursued using small samples
  - Corpus approaches remain shallow
    - E.g. regex, no syntactic context
- Instead of doing this, I was updating treebanks after AGR modifications...

non-mal and mal SRG updates: Summary

- ▶ Thanks to Montse Marimon for large-scale, careful work
- ▶ Thanks to Lorena S. Allegue for treebanking

non-mal and mal SRG updates: Summary

- Thanks to Montse Marimon for large-scale, careful work
- Thanks to Lorena S. Allegue for treebanking
- Working with learner corpus helped improve the SRG
  - Noticeable gains in performance

non-mal and mal SRG updates: Summary

- ► Thanks to Montse Marimon for large-scale, careful work
- ► Thanks to Lorena S. Allegue for treebanking
- ► Working with learner corpus helped improve the SRG
  - ► Noticeable gains in performance
- ► No "easy" high precision with mal SRG
  - ► prevents pursuing RQs
- ► Treebanking/pipeline remain a bottleneck

non-mal and mal SRG updates: Summary

- Thanks to Montse Marimon for large-scale, careful work
- Thanks to Lorena S. Allegue for treebanking
- Working with learner corpus helped improve the SRG
  - Noticeable gains in performance
- No "easy" high precision with mal SRG
  - prevents pursuing RQs
- Treebanking/pipeline remain a bottleneck
- Dream: Integrated environment autocompleting code and highlighting changes in illustrative treebank along with reasons for the changes