

Factivity of Emotive Predicates in Peninsular Spanish

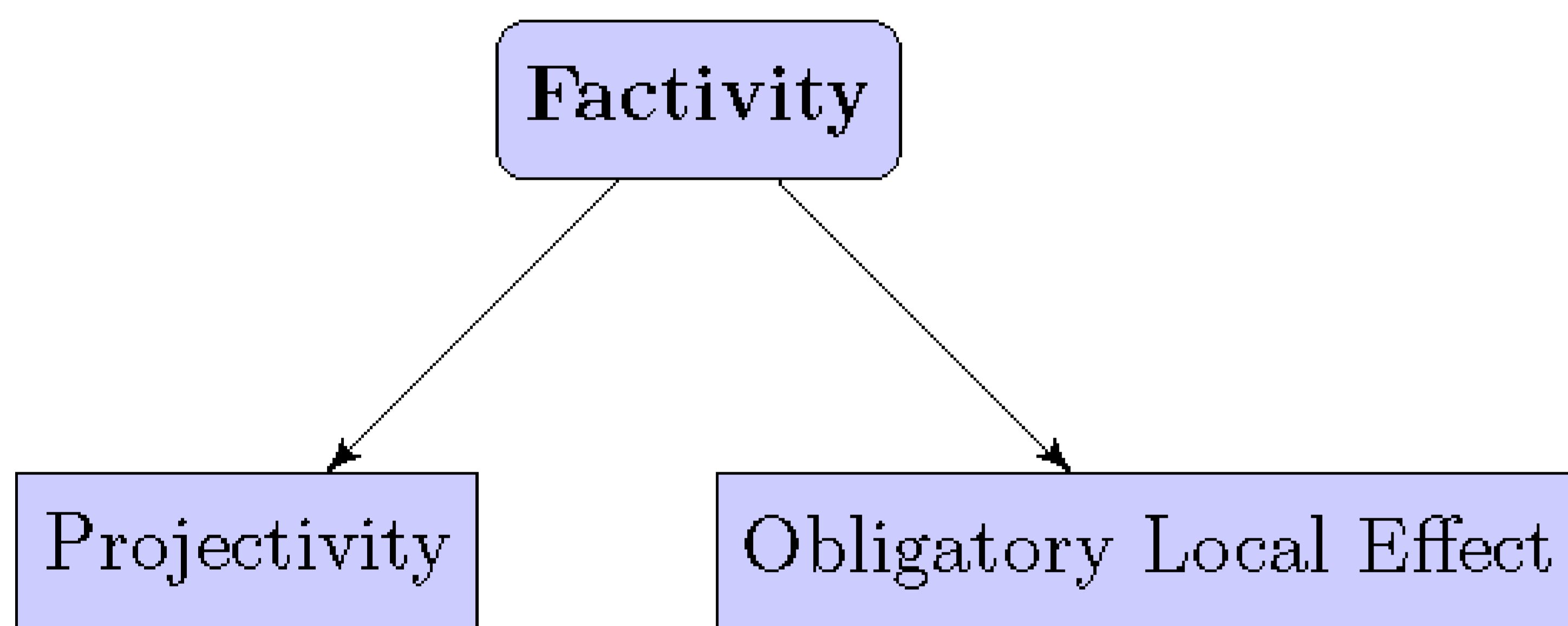
Joseph Larson

Indiana University

What is factivity?

- Semantic categorization of predicates proposed by [3].
- Employment of predicates commits the speaker of utterance to truth value of complement clause.
- In English, these complement clauses begin with subordinating conjunction *that*; in Spanish, the equivalent is *que*.
- Example:
 - *La Ministra de Sanidad, Celia, recalcó ayer que la denuncia partía del PSOE y lamentó que el problema de las listas de espera se utilice políticamente como arma arrojadiza entre partidos.*
- 'The Minister of Health, Celia, emphasized yesterday that the complaint arose from the PSOE [Spanish Political Party] and regretted that the problem with waiting lists is being used political as a throwing weapon between parties.'
- To test if a predicate is factive, we can use two different diagnostics.

Properties of Factive Predicates

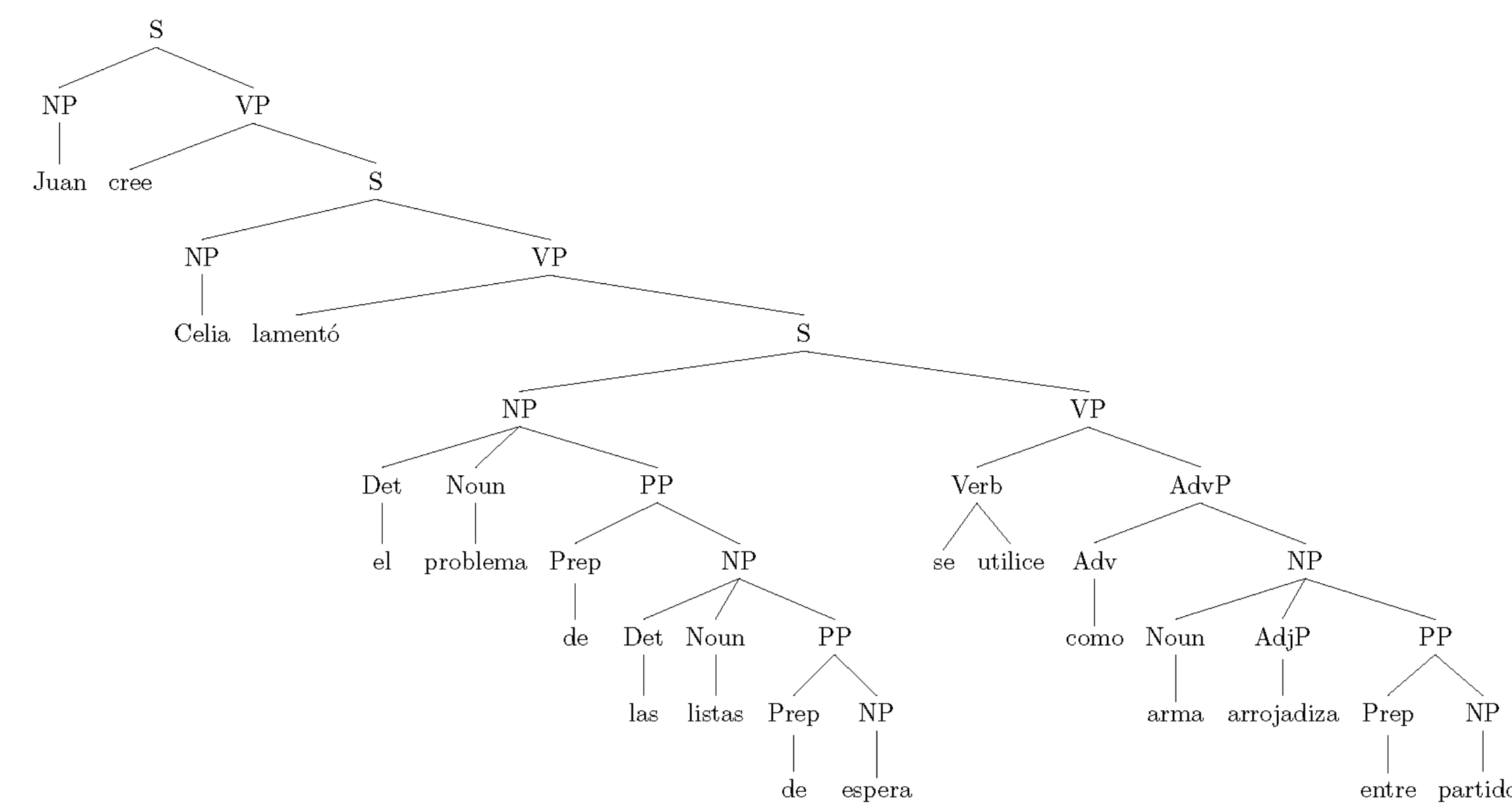


First diagnostic: Projectivity

- Projectivity is a property of certain semantic content. Content that is presupposed under certain entailment canceling operators e.g. negation, interrogation, different modalities is said to 'project.'
- Projection is tested using the 'Family of Sentences' diagnostic. [1] If the content can no longer be presupposed in any of these sentences, then the predicate does *not* project.
 - Celia lamentó que el problema de las listas de espera se utilice políticamente como arma arrojadiza entre partidos.
 - Celia no lamentó que el problema de las listas de espera se utilice políticamente como arma arrojadiza entre partidos.
 - ¿Lamentó Celia que el problema de las listas de espera se utilice políticamente como arma arrojadiza entre partidos?
 - Tal vez Celia lamentó que el problema de las listas de espera se utilice políticamente como arma arrojadiza entre partidos.
 - Si Celia lamentara que el problema de las listas de espera se utilice políticamente como arma arrojadiza entre partidos, se lo diría.
- Since in each sentence the proposition *el problema de las listas de espera se utiliza políticamente como arma arrojadiza entre partidos* can be presupposed, we can say this sentence projects.

Second Diagnostic: Obligatory Local Effect

- A property of some projected content, including complements of factive predicates [5].
- Projected content has an obligatory local effect if a belief predicate e.g. *think* or *believe* can be inserted before and its entailment necessarily carries over within the *think* embedded predicate.
 - #Juan cree que Celia lamentó que el problema de las listas de espera se utilice políticamente como arma arrojadiza entre partidos y que no se utiliza políticamente como arma arrojadiza.
- If the content has obligatory local effect, most sentences like the previous one will be judged infelicitous, if not completely contradictory.



Since *lamentar* applies to everything within the *cree* VP, it would be considered infelicitous to negate its embedded clause within the same VP.

Methodology of Current Study

- I wanted to replicate a similar study by [2]. They created a subcorpus by extracting factive predicates from naturally occurring English data. Whereas they only examined projection, I wanted to examine
- For data, I used the Ancora Spanish Corpus. It is a multi-level annotated (morphologically, syntactically) Spanish language corpus with a total of 525,000 tokens. [4]
- I extracted examples of emotive factives with *que* complements.
- After extracting all examples, I applied both aforementioned tests to evaluate factivity of predicate.

Extracted Examples by Predicate

Predicate	Total Examples	Projection	Obligatory Local Effect
<i>lamentar</i>	17	14	17
<i>alegrarse</i>	3	2	3
<i>ser de extrañar</i>	2	2	2
<i>extrañar</i>	3	2	3
<i>preocupar</i>	2	2	2
<i>sorprender</i>	3	3	3
Total	30	25	30

Table 1. Data by individual predicate.

- Only 30 examples of emotive factive predicates could be extracted from the corpus.
- Of these 30 examples, 25 were found to fully project.
- All examples did have an obligatory local effect per the diagnostic.

Extracted Examples by Grammatical Tense and Aspect

Tense	Total Examples	Projection
Present	14	14
Past preterite	10	10
Past imperfect	3	1
Conditional	3	0

Table 2. Data by grammatical tense and aspect of factive predicate.

- All conditional sentences did not project, demonstrating that within this dataset the conditional negates the factive component of the predicate.
- While all past preterite sentences did project, not all past preterite sentences did, implying that grammatical aspect affects factivity.

Who Cares?

- First, from a general perspective it is important to know speakers' and writers' intentions and beliefs. Knowing what projects can help our understanding of this further.
- Second, from a language teaching perspective, the fact that we found so few examples of factive predicates in the corpus implies such constructions are not that common in Spanish, which in turn implies that second language Spanish courses should not focus on such constructions as much.
- Lastly, from an NLP perspective, if we want to automate language processes that can understand human language, projection is an important aspect of this.

Findings and Next Steps

- With what little data I have been able to find, I can tentatively hypothesize that grammatical aspect affects projection.
- I have already found several different corpora and written scripts to extract more sentences. The next immediate step is going through these sentences and applying the same diagnostics to them.
- After my diagnostics are applied, I will get IRB approval to have native speakers annotate my findings, using a Likert scale 1-5, of how certain the speaker of the sample utterance is of the embedded clause's truth value.
- With the annotated corpus finished, I will test a LLM's e.g. BERT to inference projection.

References

- [1] G. Chierchia and S. McConnell-Ginet. *Meaning and Grammar*. MIT Press, Cambridge, MA, 1990.
- [2] Marie-Catherine de Marneffe, Mandy Simons, and Judith Tonhauser. The commitmentbank: Investigating projection in naturally occurring discourse. *Proceedings of Sinn und Bedeutung*, 23(2):107-124, Jul. 2019.
- [3] Paul Kiparsky and Carol Kiparsky. *Fact.*, page 143-173. Mouton, The Hague, 1970.
- [4] Mariona Taulé, M. Antònia Martí, and Marta Recasens. AnCor: Multilevel annotated corpora for Catalan and Spanish. In Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, and Daniel Tapias, editors, *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco, May 2008. European Language Resources Association (ELRA).
- [5] Judith Tonhauser, David Beaver, Craig Roberts, and Mandy Simons. Toward a taxonomy of projective content. *Language*, 89:66-109, 03 2013.