

The syntax and information structure interface: VS order in Catalan vs Spanish

Núria Barrios-Jurado

19th Workshop on Syntax, Semantics and Phonology – Nantes Université 2023

Outline

- The Phenomenon
- Theoretical Background
- Research Questions
- Corpus data
- Experimental data
- Concluding Remarks

Focus Structure

- Predicate-focus:

What happened to your car?
My car [BROKE DOWN].

- Argument-focus:

I heard your motorcycle broke down?
[My CAR] broke down.

- Sentence-focus:

What happened?
[My CAR broke down].

(Lambrecht 1994:223)

Focus Structure

Sentence-focus: (adapted from Lambrecht 1994:223)

What happened?

English: [My CAR broke down].

Spanish: [Se me ha roto el COCHE].

Catalan: [Se m'ha espallat el COTXE].

- In English, focus is usually marked by means of prosodic stress. In the Romance languages (which are SVO) word-order is a more common mechanism.
- In Spanish and Catalan, final constituents are assigned prosodic stress. It is believed informational focus aligns with sentential stress.

Topic

- Topic is understood as “what the sentence is **about**”.
- All sentences need a topic, even sentence-focus structures.
- Topic = Subject of Predication (SoP)
- Erteschik-Shir’s (1997) **stage topic**: the spatio-temporal parameters of the sentence.
- Topic needs to be *given* information, not *old*.

(Strawson 1964, Gundel 1974, Reinhart 1981, Erteschik-Shir 1997, 2007)

- Some Romance languages are more restrictive than others when it comes to the mapping of IS onto syntax. (Leonetti 2017)

What happened?

Spanish:

- Ha llegado María.*
has arrived Mary
'Mary has arrived.'
- Se ha ido María.*
RFL has left Mary
'Mary has left.'

Catalan:

- Ha arribat la Maria.*
has arrived the Mary
'Mary has arrived.'
- #Se n'ha anat la Maria.*
RFL CL=has left the Mary
'It is Mary who left.'

Theoretical background

- VS is the canonical word-order for unaccusatives in the Romance languages.
=> VS order should be felicitous in sentence-focus structures.
- There's a difference in interpretation between SV/VS order:

Italian (from Pinto 1997)

1. *Dante e entrato.*

Dante is entered

'Dante has entered (into some place).'

2. *E entrato Dante.*

is entered Dante

'Dante has entered (here/into this place).'

Theoretical background

- In absence of an overt locative or temporal phrase, the Romance languages require a (somewhat deictic) null locative in subject position to licence sentence-focus in VS. (Fernández-Soriano 1999, Corr 2016, a.o.)
- Bentley & Cruschina's (2018) BFSI proposal for Italian:
 - covert preverbal **Subject of Predication** (SoP) which can be
 - **Thematic SoP:** a goal argument is available in the thematic grid of the verb (e.g. *arrive*).

Sono arrivati i ragazzi
Are arrived the kids
'The kids have arrived'
[_{subj} (∃loc)][_{pred} sono arrivati i ragazzi]
 - **Situational SoP:** a non-thematic goal argument inferred from the situation or context (e.g. change-of-state verbs such as *die*).

Sono morti i soldati.
are died the soldiers
'The soldiers have died.'
[_{subj} (∃e)(past(e))][_{pred} sono morti i soldati]

Theoretical background: Bentley & Cruschina (2018)

Change-of-state predicates are divided into:

- **Definite change** (which specify an endpoint)
- **Indefinite change** (which encode the attainment of a degree with respect to a given property and only imply the existence of a final state)

Predicate type	SoP type	Example
Quantized change – speaker oriented verbs of inherently directed motion	Thematic	<i>Sono entrati I clienti.</i> 'The clients have come in (here)'
Activities with goal argument	Thematic	<i>Hanno telefonato I ragazzi.</i> 'The kids have phoned (here/us)'
Quantized change – telic verb with a state-goal	Situational	<i>Sono morti i soldati.</i> 'The soldiers have died'
Non-quantized change – atelic verbs with goal inferred from adjectival base	Situational	<i>Si è svuotato il serbatoio.</i> 'The tank has become empty.'
Non-quantized change – atelic verbs with goal inferred from contextual clues	Situational	<i>Si è accorciata la gonna.</i> 'The skirt has become short.'

Research Questions:

- What verbs can appear in a sentence-focus VS construction in Spanish and Catalan? Are Spanish and Catalan different?
- Are the BFSI licensing arguments in Bentley & Cruschina (2018) able to account for the Catalan and Spanish evidence?
 - Data triangulation: corpus study, Production task, and Acceptability Judgement Task

Corpus Study

- Corpora:
 - Spanish: *CORPES*
 - Catalan: *CTILC*
- Written data from 21st century
- Diatopic varieties:
 - Spanish: Peninsular Spanish
 - Catalan: Central Catalan

Corpus study

Aim for 400 hits/verb:

English	Spanish	Catalan	
<i>Arrive</i>	<i>Llegar</i>	<i>Arribar</i>	} Definite result state
<i>Die</i>	<i>Morir</i>	<i>Morir</i>	
<i>Appear</i>	<i>Aparecer</i>	<i>Aparèixer</i>	} Indefinite result state
<i>Grow</i>	<i>Crecer</i>	<i>Créixer</i>	
<i>Increase</i>	<i>Aumentar</i>	<i>Augmentar</i>	
<i>Improve</i>	<i>Mejorar</i>	<i>Millorar</i>	} Other
<i>Belong</i>	<i>Pertenecer</i>	<i>Pertànyer</i>	
<i>Sleep</i>	<i>Dormir</i>	<i>Dormir</i>	

➤ Total hits: 5,627 (Spanish: 3,200; Catalan: 2,427)

Manual correction of hits

- Overt subject
- Declarative sentences
- Indicative mood
- Root clause

Disregarded:

- Causative alternation
- Hits with negation
- Root clause preceded by subordinate clause

Corpus study

Each corpus hit was coded for the following categorical binary variables:

- Word order [**SV/VS**]
- Presence of a sentence-initial temporal or locative adverbial [**bare/loctemp**]
- Definiteness of the subject [**def/indef**]
- Verb class [**def result, indef result, other**]

Catalan examples from CTILC:

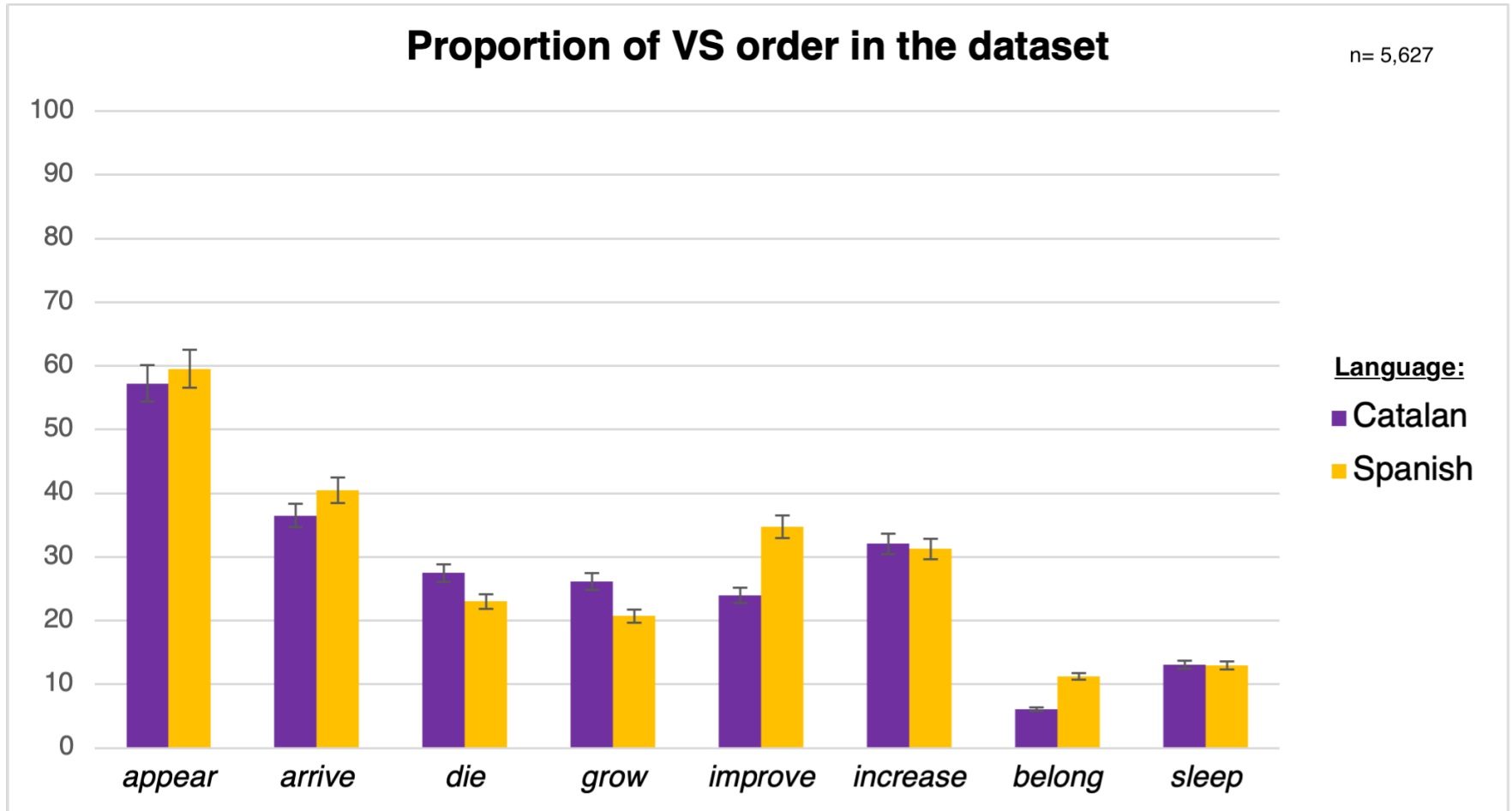
1. [VS, loctemp, def, def result]

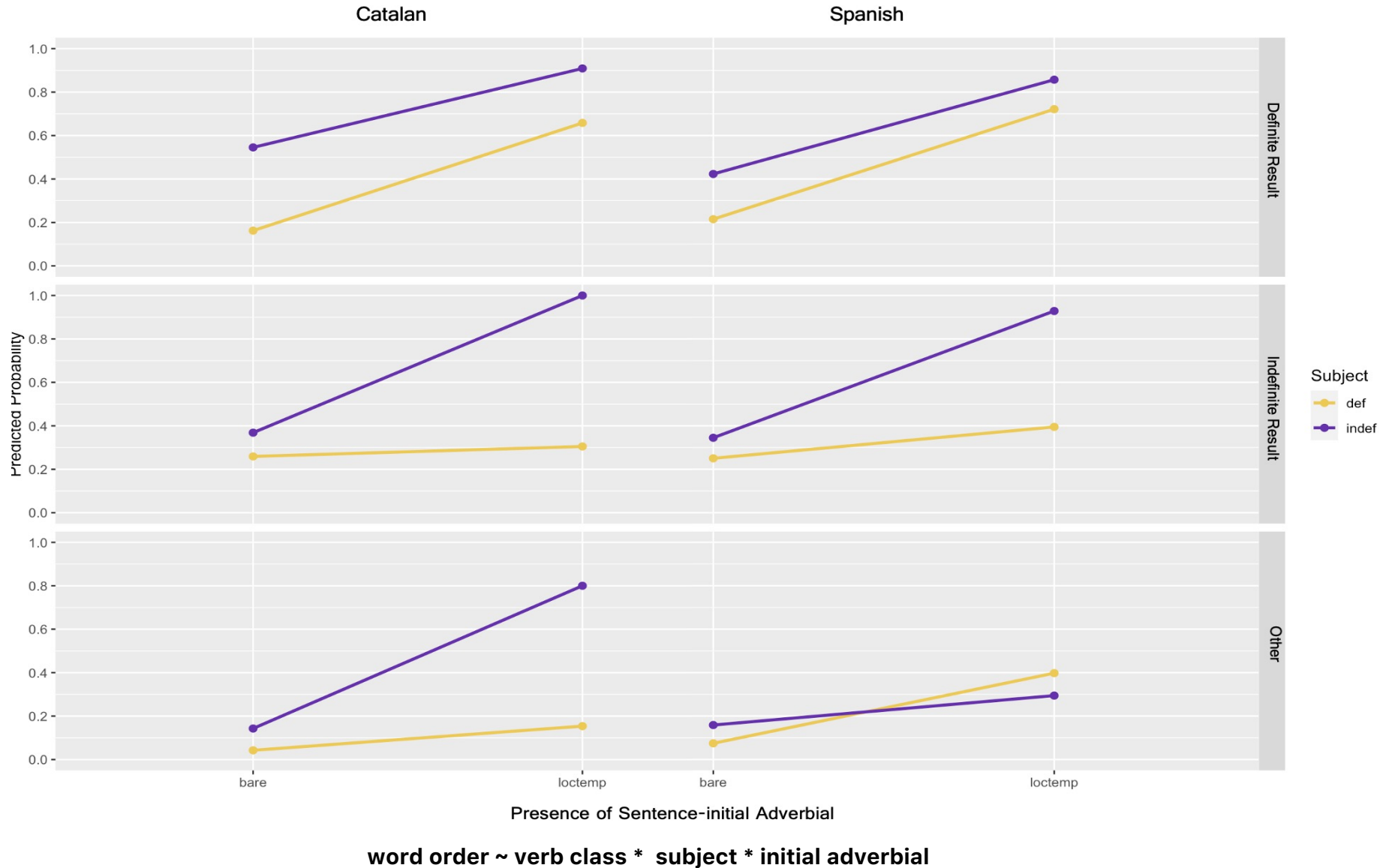
El dia 10 va morir Arturo Nogueira.
 The day 10 AUX.PST die Arturo Nogueira.
 'Arturo Nogueira died on the 10th.'

2. [SV, bare, indef, def result]

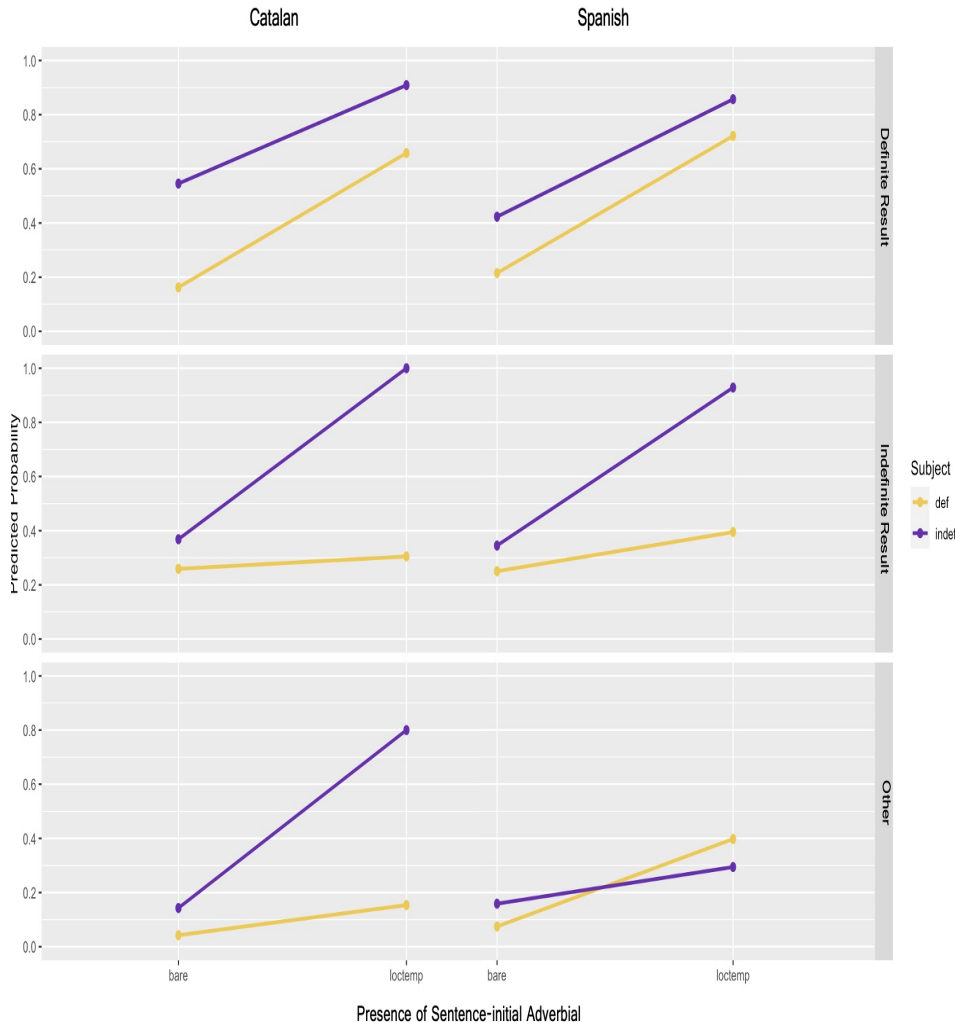
Molts vegetals morien.
 Many vegetables die.IMPF
 'Many vegetables died.'

Corpus study: Results





Corpus study: Results



- Results from fixed-effects regression model show independent variables are statistically significant (p -value < 0.05).
- Catalan and Spanish do not show significant differences.
- Presence of adverbial and indefinite subjects favour VS order.
- There are some similarities between 'indefinite result' verbs and 'other'.

Fieldwork data: AJT / Production Task

- 80 native speakers: 40 Central Catalan; 40 Peninsular Spanish
- Software: Psychopy
- 1:1 filler ratio (transitive verbs in narrow/contrastive focus)
- 12 **verbs** (x3 verb classes) x **word order** (SV/VS) x **definite subject** (def/indef)

English	Spanish	Catalan	
<i>Arrive</i>	<i>Llegar</i>	<i>Arribar</i>	Definite result
<i>Die</i>	<i>Morir</i>	<i>Morir</i>	
<i>Appear</i>	<i>Aparecer</i>	<i>Aparèixer</i>	
<i>Leave</i>	<i>Irse</i>	<i>Anar-se</i>	Indefinite result
<i>Grow</i>	<i>Crecer</i>	<i>Créixer</i>	
<i>Increase</i>	<i>Aumentar</i>	<i>Augmentar</i>	
<i>Improve</i>	<i>Mejorar</i>	<i>Millorar</i>	Unergative
<i>Change</i>	<i>Cambiar</i>	<i>Canviar</i>	
<i>Sneeze</i>	<i>Estornudar</i>	<i>Esternudar</i>	
<i>Walk</i>	<i>Andar</i>	<i>Caminar</i>	Unergative
<i>Sing</i>	<i>Cantar</i>	<i>Cantar</i>	
<i>Talk</i>	<i>Hablar</i>	<i>Parlar</i>	

Fieldwork data: AJT / Production Task

Participants presented with a context ended in “*What happened?*” (written + audio)

- **AJT:**

- 2 sentences appear in SV and VS order.
- Participant listens to them and ranks them from 1 to 5.

Spanish: *El rey ha muerto. /Ha muerto el rey.*

English: *The king has died.*

- **Production Task:**

- 1 infinitive appears on screen.
- Participant produces a sentence using the verb in answer to ‘what happened?’.

Spanish: *LLEGAR*

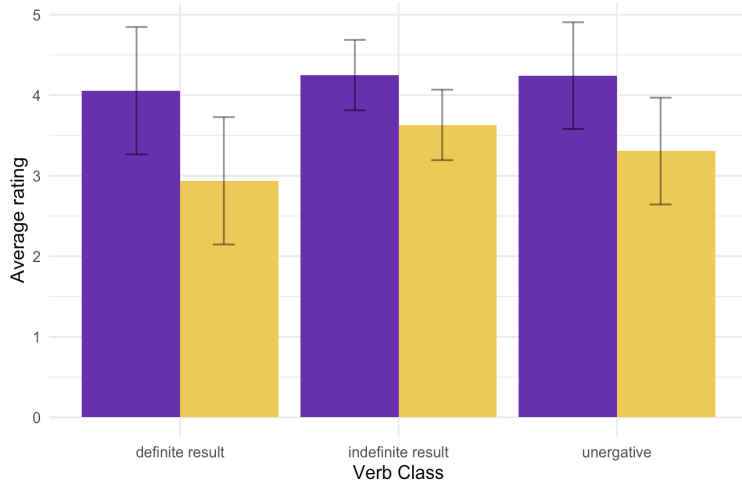
Target sentence: *Han llegado dos famosos.*

English: *ARRIVE*

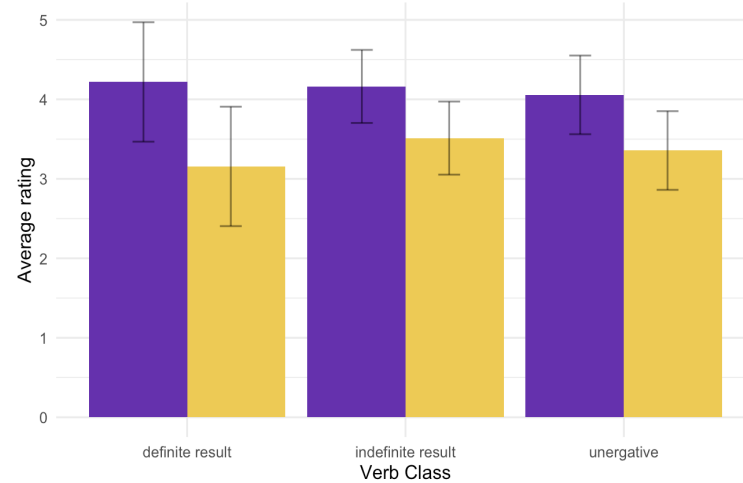
Target sentence: *Two famous people have arrived.*

Acceptability Judgement Task

Average rating in SV order by 'Verb Class' and 'Subject' (Catalan)

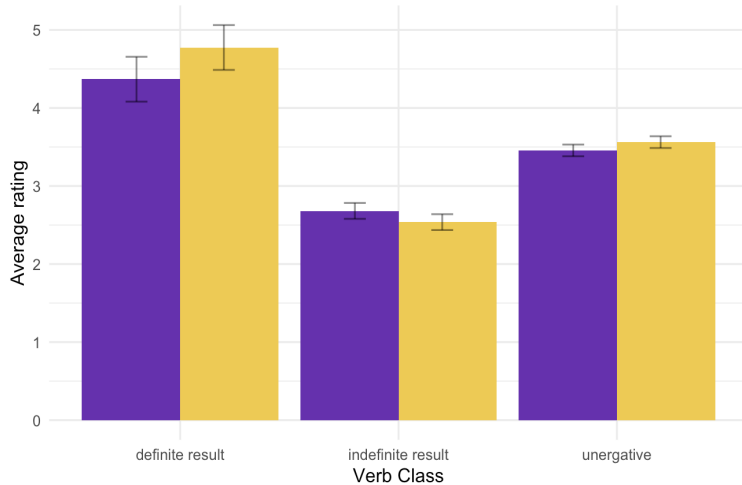


Average rating in SV order by 'Verb Class' and 'Subject'

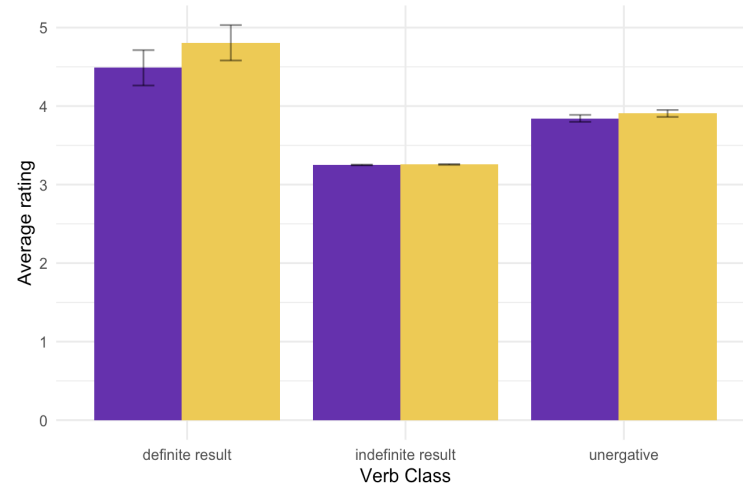


Subject:
■ definite
■ indefinite

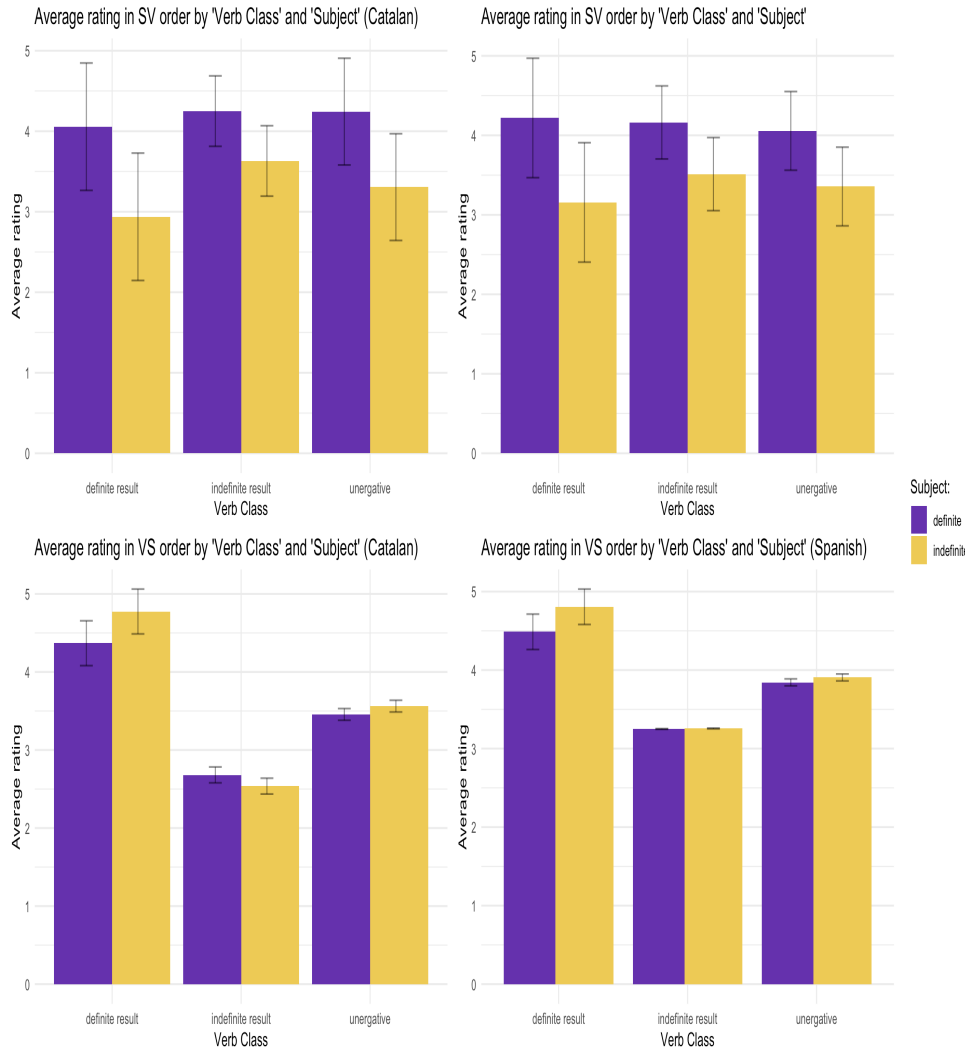
Average rating in VS order by 'Verb Class' and 'Subject' (Catalan)



Average rating in VS order by 'Verb Class' and 'Subject' (Spanish)



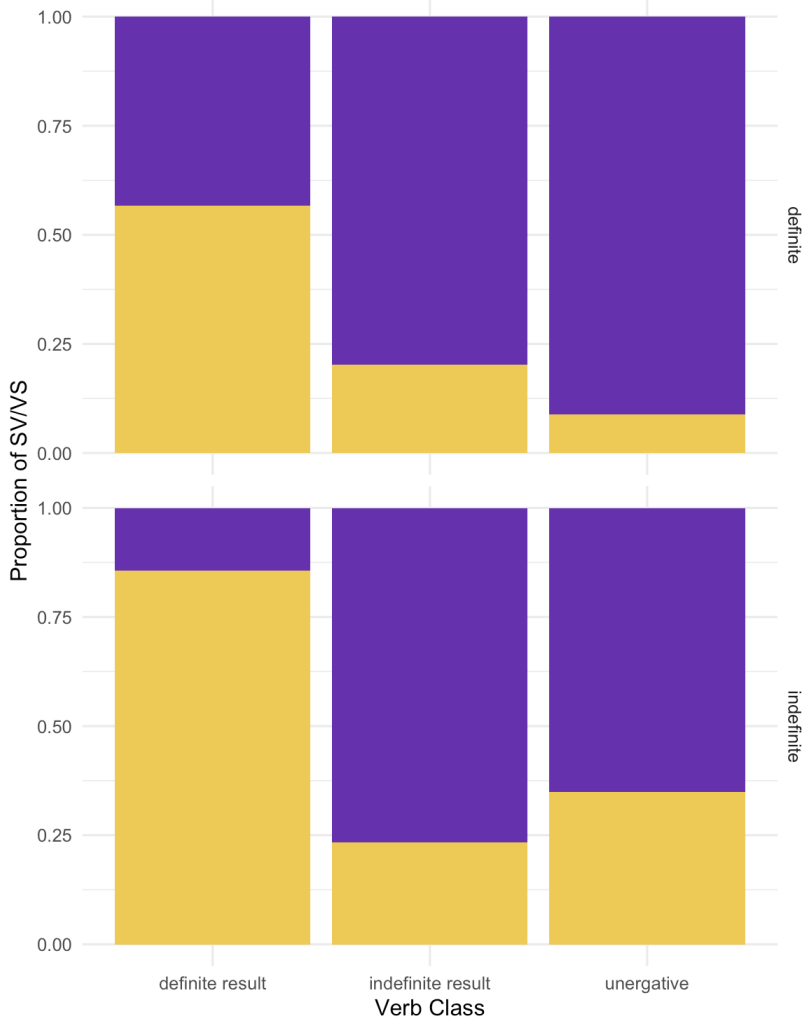
Acceptability Judgement Task



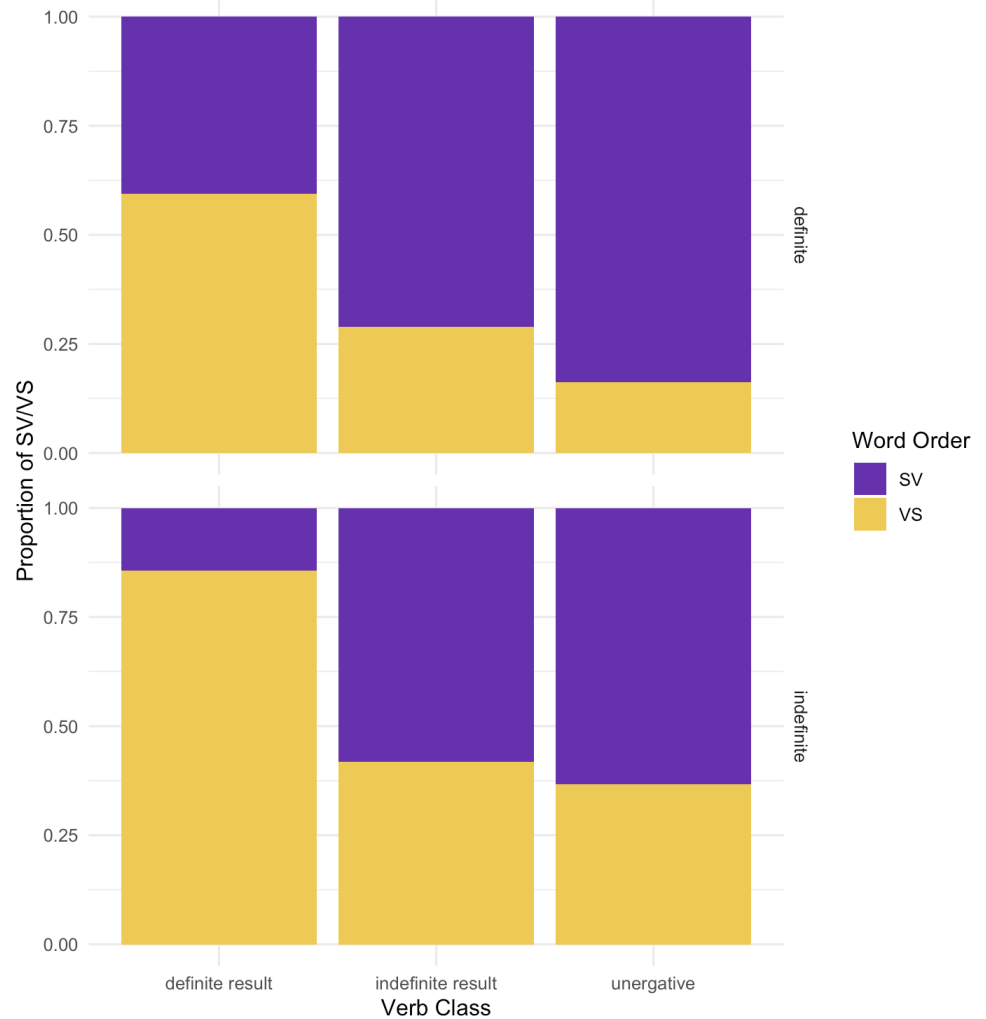
- Catalan and Spanish show no significant differences.
- No significant differences between verb class in SV order.
- Definite subjects are preferred in SV order.
- Some differences between verb class in VS order. Indefinite result and unergatives seem to be more similar than definite and indefinite result.
- Slight preference for indefinite subjects in VS order definite result verbs.

Production Task

Word Order by Verb Class and Subject (Catalan)

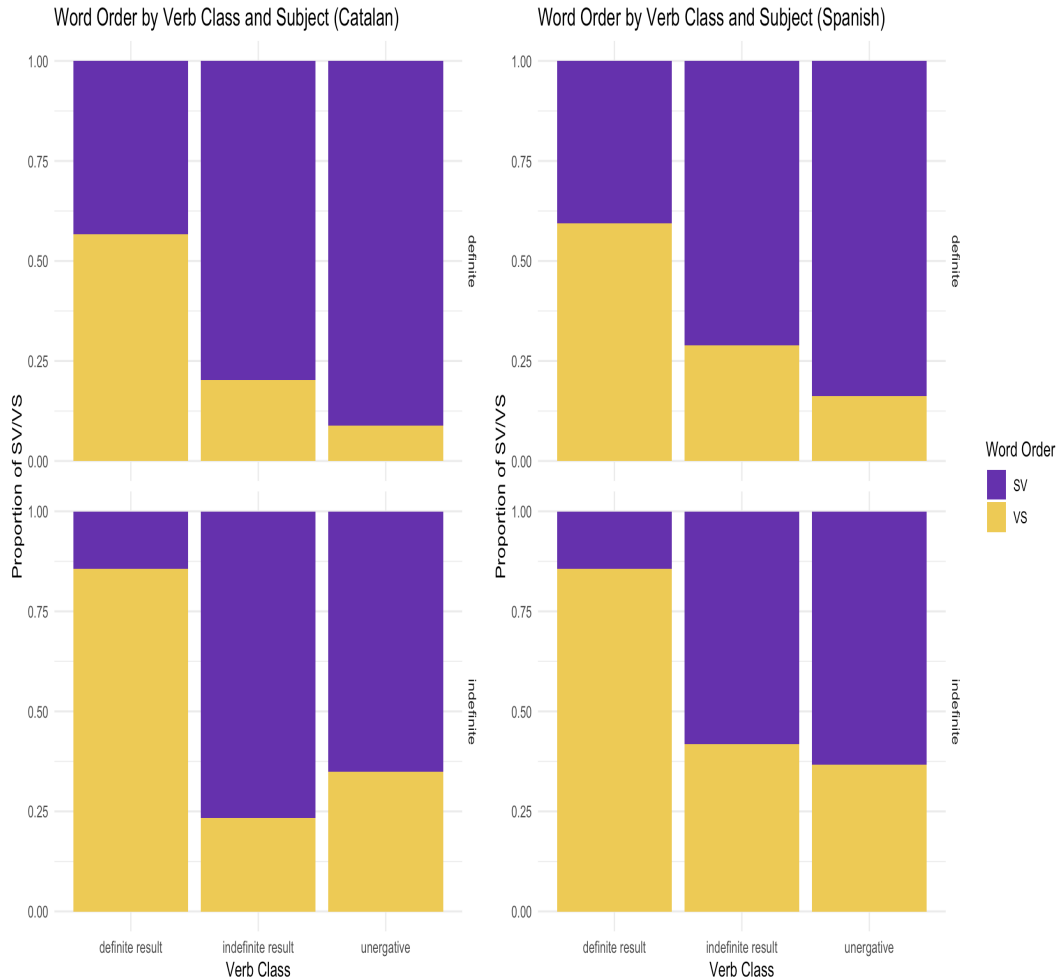


Word Order by Verb Class and Subject (Spanish)



Word Order
■ SV
■ VS

Production Task



- Catalan and Spanish show no significant differences.
- VS is more frequent with definite result verbs, especially with indefinite subjects.
- Some differences between verb class. Indefinite result and unergatives seem to be more similar than definite and indefinite result.

Concluding remarks

- The data seems to indicate there is no significant differences between Catalan and Spanish.
- Bentley & Cruschina's (2018) theoretical proposal seems to be on the right track with regards to *verb class* being crucial in the licensing of BFSI. However, at least for Catalan and Spanish, only verbs with a definite goal might license it.
- Quantitative analysis of fieldwork data will help determine if the predictors are statistically significant (as the plots seem to suggest).

Concluding remarks – Further research

- Qualitative analysis may shed some light onto microvariation between languages.
- Need to explore the differences between categorical andthetic interpretations.

English: *The king has died.*

Spanish: *El rey ha muerto.* SV = categorical?

Ha muerto el rey. VS =thetic?

- Need to explore exhaustive reading in VS order in Catalan vs Spanish:

Catalan: *Ha marxat la Maria.* (Mary left and nobody else did)

Spanish: *Se ha ido Maria.* (Mary left and maybe somebody else did too)

References

- Beavers, J. (2013). Aspectual classes and scales of change. *Linguistics*, 51(4), 681-706.
- Benincà, Paola (1988). 'L'ordine degli elementi della frase e le costruzioni marcate'. In: *Grande grammatica italiana di consultazione alla linguistica del mòcheno*. Ed. by Lorenzo Renzi. Vol. 1. Bologna: Rosenberg & Sellier, 115–225.
- Bentley, Delia & Silvio Cruschina (2018). 'The silent argument of broad focus: Typology and predictions'. *Glossa* 3 (1): 1–37. ISSN: 2397-1835. DOI: 10.5334/gjgl.677.
- Corr, Alice (2016). 'Wide-focus subject-verb inversion in Ibero-Romance: A locative account'. *Glossa: a journal of general linguistics* 1 (1) 1–33. ISSN: 2397-1835. DOI: 10.5334/gjgl.85.
- Erteschik-Shir, Nomi (1997). *The Dynamics of Focus Structure*. Cambridge Studies in Linguistics. Cambridge: Cambridge University Press. ISBN: 978-0-521-59217-8. DOI: 10.1017/CBO9780511519949.
- Erteschik-Shir, Nomi (2007). *Information Structure: The Syntax-Discourse Interface*. Oxford University Press.
- Lambrecht, Knud (1994). *Information structure and sentence form: topic, focus, and the mental representations of discourse referents*. Cambridge studies in linguistics 71. Cambridge: University Press. ISBN: 978-0-521-38056-0.
- Leonetti, Manuel (2017). 'Basic constituent orders'. In: *Manual of Romance Morpho-syntax and Syntax*. De Gruyter, 887–932. ISBN: 978-3-11-037708-8. DOI:10.1515/9783110377088-024.
- Leonetti, Manuel & Victoria Escandell Vidal (2021). *La estructura informativa. spa. Gramática del Español* 18. Madrid: Visor. isbn: 978-84-9895-641-2
- Pinto, Manuela (1997). *Licensing and interpretation of inverted subjects in Italian*. UiL OTS dissertation series. Utrecht: Led. ISBN: 978-90-5434-064-5.
- Sheehan, Michelle (2006). 'The EPP and Null Subjects in Romance'. PhD thesis. Newcastle Upon Tyne: Newcastle University. URL: <http://people.pwf.cam.ac.uk/mtb23/NSP/Sheehan%20dissertation.html>.
- Sheehan, Michelle (2010). 'Free' inversion in Romance and the Null Subject Parameter'. In: *Parametric Variation: Null Subjects in Minimalist Theory*. Ed. By Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan. Cambridge: Cambridge University Press, 231–262.
- Zubizarreta, Maria Luisa (1998). *Prosody, focus, and word order*. Cambridge, Massachussets: MIT Press. ISBN: 978-0-262-74021-0.

Corpora:

- REAL ACADEMIA ESPAÑOLA. Banco de datos (CORPES XXI) [en línea]. Corpus del Español del Siglo XXI (CORPES). URL: <http://www.rae.es>.
- IEC. Corpus textual informatitzat de la llengua catalana [en línea] (CTILC) URL: ctilc.iec.cat.