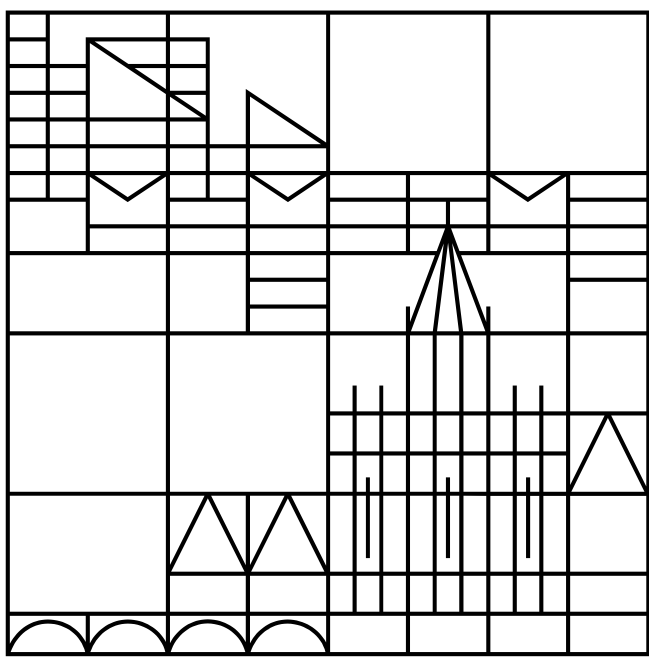# On word order variation and information structure in Spanish and Italian 'why'-interrogatives

Svenja Schmid & Georg A. Kaiser

svenja.schmid@uni-konstanz.de, georg.kaiser@uni-konstanz.de

DFG Deutsche Forschungsgemeinschaft

Universität Konstanz

FOR2111

Background

Word order variation in Italian and Spanish 'why'-interrogatives

– Word order in Italian and Spanish *wh*-interrogatives: *wh*VS → obligatory subject-verb inversion (Rizzi 2001, Suñer 1994, Torrego 1984).

(1) Che (***Luca**) ha comprato (**Luca**)?   (Italian)
what Luca has bought Luca
'What did Juan buy?'

(2) ¿Qué (***Juan**) compró (**Juan**)?   (Spanish)
what Juan bought Juan
'What did Juan buy?'

– 'Why' (Sp. *por qué* and It. *perché*) allows preverbal and postverbal subjects (e.g., Rizzi 2001, Torrego 1984)

(3) Perché (**Maria**) balla (**Maria**)?   (Italian)
why Mary dances Mary
'Why does Mary dance?'

(4) ¿Por qué (**María**) baila (**María**)?   (Spanish)
why Mary dances Mary
'Why does Mary dance?'

– Bianchi, Bocci & Cruschina (2017) and Leonetti (2018) show that information structure affects word order in Italian 'why'-interrogatives
→ Italian: postverbal subjects are interpreted as focal ([+Foc]), preverbal subjects are [-Foc]
→ Spanish: no research on this topic, but crucial difference to Italian concerning the frequency of word order patterns (see Table 1)

RQ 1: Does information structure affect word order in Italian and Spanish 'why'-interrogatives?
RQ 2: Do Spanish and Italian 'why'-interrogatives differ from each other and if yes, why?

	SV	VS
IT	**65** 78.3%	18 21.7%
SP	**25** 21.4%	92 78.6%

Table 1: Word order in 'why'-interrogatives (Parallel Text Corpus)

Empirical studies

Parallel Text Corpus: Method & Results

– 200 matrix 'why'-interrogatives with an overt subject (from detective novels and stories published in 'The Complete Sherlock Holmes' and the Bible (Old Testament: Gen.-2 Kings, New Testament: Matt.-Acts))

– Manual annotation of
  ▪ **Syntactic features**:
  word order (whSV vs. whVS)
  clause type (matrix vs. embedded)
  subject (null vs. overt)
  ▪ **Discourse categories of the subject**: (Heidinger 2018:46)
  Focus: set of relevant alternatives ([+Foc] vs. [-Foc]), IRA: 97.3%, Cohen's κ = 0.73 (substantial)
  Givenness: new vs. given information ([+Given] vs. [-Given]), IRA: 100%, Cohen's κ = 1.0 (perfect)

		unmarked	marked		
		[+Given] [-Foc]	[-Given] [+Foc]	[+Given] [+Foc]	[-Given] [-Foc]
Italian	**whSV**	**51**	1	7	5
	*wh*VS	2	0	13	3
Spanish	**whSV**	**8**	2	9	6
	*wh*VS	73	1	14	4

Table 2: Word order in 'why'-interrogatives across information structure conditions (absolute numbers)

– Word order has an interpretative effect in both languages
Position of unmarked subjects: <u>Italian</u>: 'why'SV,
<u>Spanish</u>: 'why'VS

– Statistical analysis with mixed-effects regression model in R:
  ▪ Significant effect of language (β = 5.48, SE = 0.81, z = 6.74, p < .001)
  ▪ Significant effect of 'markedness' in Italian (β = 3.54, SE = 0.84, z = 4.22, p < .001) and Spanish (β = -2.10, SE = 0.50, z = -4.19, p < .001)
  ▪ No effect of text (β = -0.42, SE = 0.40, z = -1.05, p > .05)

Example of annotations

Example of an item

Forced-choice experiment: Method & Results

Experimental design:
– Web-based questionnaire in a Latin square design
– Three independent variables with two levels each:
  ▪ Language: Italian vs. Spanish
  Concerning the subject:
  ▪ Focus: [+Foc] vs. [-Foc]
  ▪ Status of information: [-Given] vs. [+Given]
– 24 items, resulting in 96 experimental stimuli
– 4 lists, every participant receives 24 experimental stimuli and 24 fillers in a pseudo-randomized order

Participants:
– 372 participants (recruited via Facebook)
– 157 Italian monolinguals (mean age = 32.3, range = 20-63, 111 female, 46 male)
– 215 speakers of Spanish from Spain (mean age = 34.3, range = 20-78, 169 female, 45 male, 1 other)
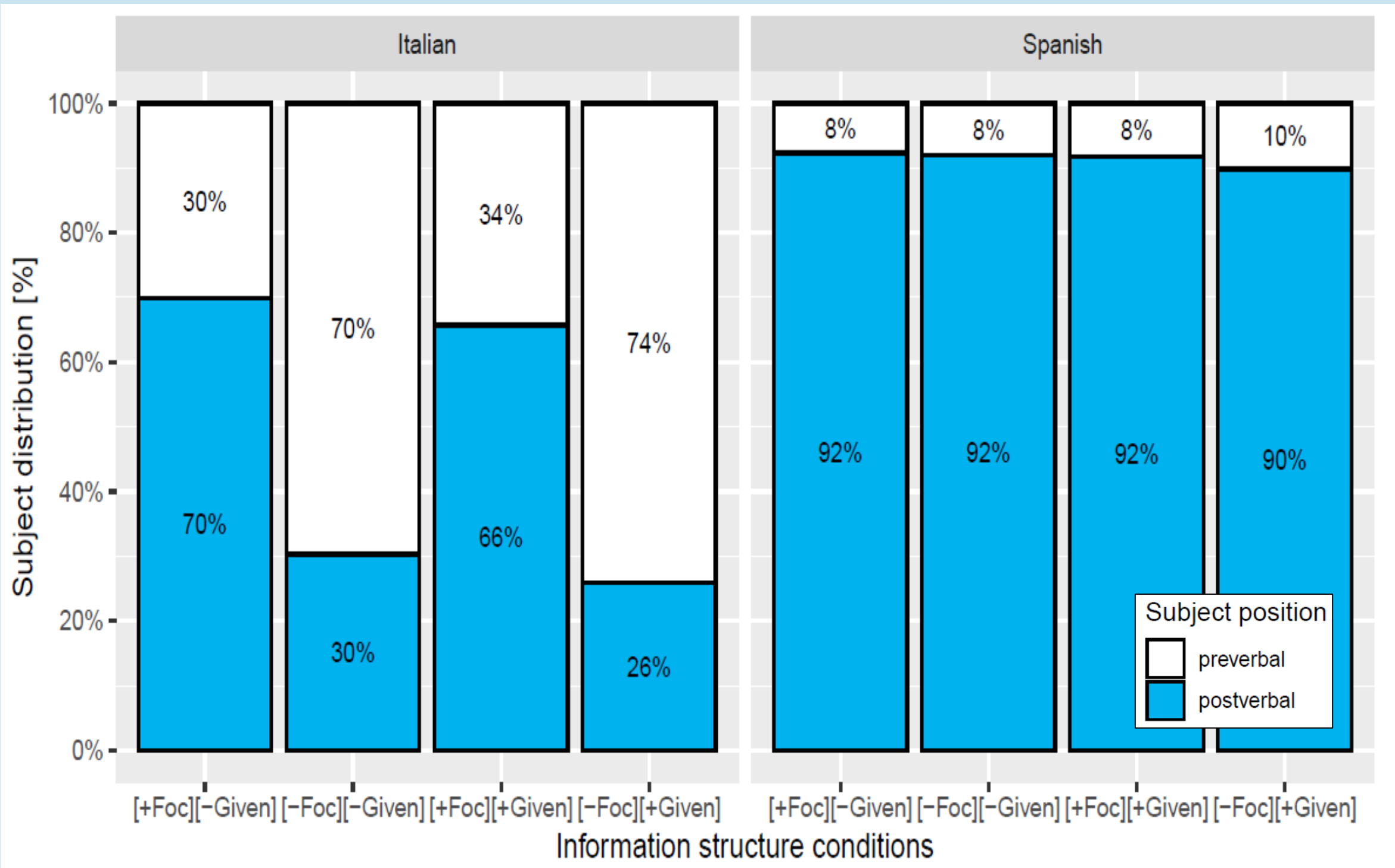– 3 Italian and 3 Spanish participants were excluded

Figure 1: Distribution of pre- and postverbal subjects in Italian and Spanish across information structure conditions.

Statistical analysis with mixed-effects regression model in R:
– Significant effect of language (β = -2.8, SE = 0.07, z = -39.05, p < .001)
– Significant effect of focus in Italian (β = -2.1, SE = 0.12, z = -17.51, p < .001)
– no effect of focus in Spanish (β = -0.06, SE = 0.16, z = -0.35, p = .72)

Discussion

RQ 1: Does information structure affect word order in 'why'-interrogatives?

Italian: YES
Especially focus has important impact on word order in 'why'-interrogatives

Spanish:
– Intransitive verbs: NO
  general preference for postverbal subjects regardless of the information structure
– Transitive verbs: more research needed

RQ 2: Do Spanish and Italian 'why'-interrogatives differ from each other and if yes, why?

YES → differences with respect to the frequency of word order patterns:
– Italian: subject position is triggered by information structure (preverbal = [-Foc], postverbal = [+Foc])
– Spanish: subject position only marginally affected by information structure

A first syntactic account for 'why'-interrogatives within the cartographic framework:
– **wh-word**: 'why' is directly merged in IntP (Rizzi 2001, Shlonsky 2011, Stepanov & Tsai 2008)
– **Verb**: the verb does not move to C in neither of the languages (Goodall 1993, Suñer 1994, Rizzi 2001)
– **Subject**: difference in subject position is due to different properties regarding information structure.
Fine structure of the split-IP (as in Belletti 2004, Cardinaletti 2004)
  ▪ <u>Italian</u>: [CP [SubjP subject[-Foc] [TP . . . [TopP [FocP subject[+Foc] [TopP [vP]]]]]]]
  ▪ <u>Spanish</u>: [CP [TP subject [T . . . [TopP [FocP subject[+Foc] [TopP [vP subject[-Foc] [vP]]]]]]]]

REFERENCES: • Belletti, A. 2004. Aspects of the low IP area. In: L. Rizzi (ed.) *The structure of CP and IP*. Oxford, 16-51. • Bianchi, V., Bocci, G., & Cruschina, S. 2017. Two types of subject inversion in Italian wh-questions. *Revue roumaine de linguistique* 62, 233-252. • Cardinaletti, A. 2004. Towards a cartography of subject positions. In: L. Rizzi (ed.) *The structure of CP and IP*. Oxford, 115-165. • Goodall, G.1993. Spec of IP and Spec of CP in Spanish wh-questions. In: W.J. Ashby et al. (eds), *Linguistic Perspectives on the Romance Languages. Selected Papers from the 21st Linguistic Symposium on Romance Languages (LSRL XXI). Santa Barbara, California, 21-24 February 1991*. Amsterdam, 199–209. • Kaiser, G. A., Heusinger, K. v. & Schmid, Svenja. 2019. Word order variation in Spanish and Italian interrogatives. The role of the subject in 'why'-interrogatives. In: N. Pomino (ed.). Proceedings of the IX Nereus Workshop (=Arbeitspapiere 131). Universität Konstanz, 69-90. • Leonetti, M. 2017. Basic constituent orders. In: A. Dufter & E. Stark (eds.), *Manual of Romance Morphosyntax and Syntax*. Berlin, 887-932. • Leonetti, M. 2018. Two types of postverbal subject. *Italian Journal of Linguistics* 30, 11-36. • Rizzi, L. 1997. The fine structure of the left periphery. In: L Haegman (ed.) *Elements of Grammar. Handbook in Generative Syntax*. Dodrecht, 281-337. • Rizzi, L. 2001. On the position 'Int(errogative)' in the left periphery of the clause. In: Cinque, G & Salvi, G. (eds.) *Current Studies in Italian Syntax. Essays offered to Lorenzo Renzi*. Amsterdam, 287-296. • Stepanov, A & Tsai, W,-T. D. 2008. Cartography and licensing of wh-adjuncts: a cross-linguistic perspective. *Natural Language Linguistic & Theory* 26, 589-638. • Suñer, M. 1994. V-movement and the licensing of argumental wh-phrases in Spanish. *Natural Language & Linguistic Theory* 12(2), 335-372. • Torrego, E. 1984. Algunas observaciones sobre las oraciones existenciales con haber en español. *Estudis gramaticals 1*, 329-339.