Embedded questions in Romance: Nominal, not clausal

Carla Bombi and Jeanne Lecavelier (University of Potsdam)

Spanish and French can express the equivalent of *'I know [what/who p]'* with a definite construction, introduced by a definite article and a demonstrative pronoun respectively:

(1) WHAT-CORRELATES

a. Sé [**lo** **que** hizo ruido].
I.know DEF.N COMP made noise
'I know what made noise.'

b. Je sais [**ce** **qui** a fait du bruit].
I know DEM REL made noise
'I know what made noise.'

(2) WHO-CORRELATES

a. Sé [**el** **que** hizo ruido].
I.know DEF.M.SG COMP made noise
'I know who made noise.' (Spanish)

b. Je sais [**celui** **qui** a fait du bruit].
I know DEM.M.SG REL made noise
'I know who made noise.' (French)

The status of (1)–(2) is unclear: Spanish *what*-correlates are sometimes described as concealed questions, but recent approaches suggest an interrogative analysis (Suñer 1999 vs. Kellert 2018); meanwhile, French constructions are treated as a syntactic variant of wh-words (Sportiche 2008, Konrad 2019). In both languages, *who*-correlates are barely discussed. There is also an unresolved empirical puzzle: *what*-correlates have a distribution similar to embedded *wh-*, but *who*-correlates are more restricted, and it is unclear what licenses them.

Proposal: 1/ We propose a unified analysis of (1)–(2) as concealed questions (CQs, Heim 1979), treated as individual concepts with maximality semantics; **2/** the distribution of *who*-correlates is constrained by how ϕ-features interact with maximality and domain restriction.

1/ They are CQs (individual concepts). ***a) Analysis.*** (1)–(2) involve intensionalised headless RCs. We illustrate with Spanish, but extend to French in our talk. Based on the structure in (3), all constructions have the same basic denotation in (4) (Caponigro 2003, Hinterwimmer 2008), but they differ in that only *who*-correlates are specified for gender/number, cf. (5): *what*-correlates are number-neutral, but *who*-correlates refer to humans.

(3) Structure: $[_{\text{DP}}$ lo/el $[_{n\text{P}}$ *n* $[_{CP}$ OP_1 que t_1 hizo ruido]]] (adapted from Saab 2008)

(4) Denotation: $[\![lo/el\ n{+}RC]\!] = \sigma x[P(x) \wedge Q(x)]$
→ P ≈ THING (*what*) / PERSON (*who*)
→ Q ≈ RC-property

(5) Features:
a. *lo* ↔ [D]
b. *el* ↔ [D, #: SG, γ: M, human]

(6) Derivation: $[\![\text{DP in (1)}]\!] = [\![\text{lo}]\!]\ ([\![n\ \text{Op}_1\ \text{que}\ t_1\ \text{hizo ruido}]\!])$
$= \lambda P.\sigma x[P(x)]\ (\lambda x.\text{THING}(x) \wedge \text{made}(\text{noise}, x)) = \sigma x[\text{THING}(x) \wedge \text{made}(\text{noise}, x)]$

In (1)–(2), the DP in (6) combines with *know* similar to regular CQs (*'I know the price of milk'*). Assuming CQs are individual concepts (Romero 2005, 2007), we get the following denotation:

(7) $[\![(1)]\!] = 1$ iff for every world w' compatible with the speaker's knowledge:
$\sigma x[\text{THING/PERSON}(x,w') \wedge \text{make}(\text{noise},x,w')] = \sigma x[\text{THING/PERSON}(x,w_0) \wedge \text{make}(\text{noise},x,w_0)]$
≈ *the speaker knows the value of "the thing(s)/person that made noise" in the actual world*

b) Evidence. Syntactically, our constructions behave like regular DPs, both under question-embedding verbs (*know*) and extensional ones (*buy*). They are compatible with *'all'*, which only takes DEF-DPs, cf (8), and they license superlatives in Sp., which obligatorily involve a definite DP, cf. (9) (Bosque & Brucart 1992). In these cases, interrogatives are ruled out.

(8) J'achèterai / Je sais [{tout **ce** **que**} / {*tout **quoi**} tu veux pour Noël].
I=will.buy / I know all DEM REL / all what you want for Christmas
'I will buy / I know everything that you want for Christmas.' (French)

(9) Compré / sé [{**lo que**} / {**#qué**} más me ha gustado].
I.bought / I.know DEF.N COMP / what more me has pleased (Spanish)
'I bought / I know what I liked the most.' ⇝ *superlative reading only with* **lo que**

Semantically, our constructions also show properties of CQs: specificational readings (Frana 2020) and Heim's ambiguity, shown in (10) (Heim 1979), which does not appear with unambiguously interrogative pronouns such as English *which* or Spanish *qué*, cf. (11).

(10) I know the thing that Ana knows.
Reading A: *I know the same thing Ana knows.*
Reading B: *I know which thing Ana knows.*

(11) Sé [**lo que/qué** sabe A.]
I.know DEF.N C/what knows A
→***lo que:*** *reading A & B;* ***qué:*** *only B*

2/ Differences in distribution. But if all the constructions are CQs, why do *what*-correlates have a wider distribution than *who* ones? Two factors explain their differences.

a) Interaction of ϕ-features and maximality. The constructions are introduced by σ, which presupposes a maximal individual satisfying the NP-description. This plays out differently for *what* and for *who*. Since *what*-correlates are number-neutral, the maximal individual can be atomic or plural. But *el que/celui que* are M.SG: its domain contains only atomic entities, and maximality thus translates into uniqueness (Sharvy 1980). Taking gender into account, they presuppose that *'there is a unique male individual with the RC description'*. This imposes stronger conditions on the context for the SG version of *who*-correlates:

(12) *Context: there is a birthday party.*

a. #Je sais [**celui qui** viendra].
I know DEM.M.SG REL will.come
'I know who will come.'
(there is no unique male individual coming: many people (F/M) attend parties)

b. Je sais [**ce qu'**ils lui offriront].
I know DEM REL=3PL DAT will.gift
'I know what they will gift her.'
(there is a maximal plurality of presents)
(French)

b) Condition on domain restriction. *Who*-correlates are restricted by a further, previously unidentified factor: they are only possible in contexts involving groups, lists, teams, etc., cf. (13). We capture this as a condition on the shape of the domain: it must have a finite, pre-specified cardinality, akin to d-linking in *which*-questions (Pesetsky 1987).

(13) Je sais [**celui qui** va gagner la compétition de ce soir].
I know DEM.M.SG REL will win the competition of this evening
'I know who will win the competition tonight.' (French)
(✓: if there is a participant list. ✗*: if the competition registration hasn't opened yet.)*

The condition also applies to *what*-correlates, but is masked by their restrictor, 'THING'–a higher order concept which easily has a kind reading. The domain of kinds is already specified as containing existing kinds, and so *what*-correlates may occur without a list context.

(14) Sé [**lo que** está comiendo Ana].
I.know DEF.N COMP is eating Ana
'I know what Ana is eating.' (Spanish)

KIND: I know the *kind of thing* A. is eating
$D_k = \{$FISH, LETTUCE, CARROT, ...$\}$
✓*no specific context needed*

Conclusion. The Spanish and French data add to the growing evidence that, cross-linguistically, CQs are a widespread alternative to (embedded) interrogatives (Arkadiev & Caponigro 2021, Li 2024). The paper contributes to the study of these alternative structures by identifying specific features regulating their distribution, as well as criteria to tease them apart from true embedded interrogatives.

Arkadiev, P. & I. Caponigro. 2021. *Conveying content questions without wh-words.* Proceedings of SuB 25. **Caponigro, I. 2003.** *Free Not to Ask.* PhD Thesis. **Frana, I. 2020.** Concealed questions. **Heim, I. 1979.** Concealed questions. **Kellert, O. 2017.** Interrogatives. **Konrad, I. 2019.** *Quand la nominalisation envahit la subordination.* PhD Thesis. **Li, H. & J. Tamura 2024.** Embedded questions as def. descriptions. SuB Poster presentation. **Romero, M. 2005.** Concealed questions and specificational subjects. **Saab, A. 2008**.*Hacia una teoría de la identidad parcial en la elipsis.* PhD Thesis. **Sharvy, R. 1980.** A more general theory of definite descriptions. **Sportiche, D. 2008.** Inward bound. Splitting the wh-paradigm and French relative qui. *Ms. UCLA.* **Suñer, M. 1999.** La subordinación sustantiva: La interrogación indirecta.