

Definite but definitely different: Headless RCs in Spanish and Brazilian Portuguese

Carla Bombi (University of Potsdam) & Giuseppe Varaschin (Humboldt University of Berlin)

The paper deals with the typology of headless relative clauses (RCs) and the provenance of matching effects cross-linguistically. We study the microvariation of *what*-headless RCs in Brazilian Portuguese (BP) and (European) Spanish, which are both introduced by a definite article (*o / lo*).

- | | |
|---|--|
| <p>(1) Eu como o que vejo.
 I eat.1SG D.M what see.1SG
 ‘I eat what I see.’ (BP)</p> | <p>(2) Yo como lo que veo.
 I eat.1SG D.N COMP see.1SG
 ‘I eat what I see.’ (Spanish)</p> |
|---|--|

Despite their superficial similarity, (1) and (2) are fundamentally distinct structures. BP (1) is a true FR introduced by a [WH +, D +] element (Marchesan 2012, Brito 2024). Spanish (2), in turn, challenges typologies of headless RCs on two counts: 1/, it instantiates an empty *n* construction (Plann 1980, Saab 2008), a headless RC not contemplated by current typologies; 2/ its matching effects cannot be accounted for by usual Agree mechanisms or relabeling, both of which rely on structures like the ones we see in BP (Himmelreich 2017, Cecchetto & Donati 2015). This comparison allows us to refine both the inventory of headless RCs and the conditions that give rise to matching effects.

1. Different structure. a) **BP** *o que* patterns as a *wh*-form (Šimík t.a.): (i) it coordinates with other *wh*-pronouns, *lo que* does not, (3)-(4); (ii) it can appear in interrogatives, *lo que* cannot, (5)-(6); (iii) it combines with *ever*-morphemes, *lo que* does not, (7)-(8). For **BP**, these diagnostics support the COMP-type FR analysis in (11), where *o que* is a complex form bearing D and *wh* features (Groos & van Riemsdijk 1981; Grosu 2003; Himmelreich 2017). Further motivation comes from two new observations: *o que* can co-occur with domain-restricting *wh*-modifiers (9), and with overt Cs in some dialects (10). In contrast, **Spanish** *lo que*’s behavior supports the empty *n* structure in (12) (Plann 1980; Saab 2008; contra e.g. Ojea 2013). This is also motivated by the use of *lo* in nominalizations (e.g. *lo n verde* ‘the green’, *lo n de siempre* ‘the of always’, which also involve *lo* + empty *n*). Alternative analyses of *lo que* as a light-headed RC (Citko 2004) are untenable because the latter do not exhibit matching effects, (20), and *lo* is not a pronoun/demonstrative.

- | | |
|--|--|
| <p>(3) Sei o que e quem fez barulho.
 I.know D.M what and who made noise</p> | <p>(4)*Sé lo que y quién hizo ruido.
 I.know D.N C and who made noise</p> |
| <p>(5)*(O que a Ana viu?
 D.M what the Ana saw</p> | <p>(6) ¿(Lo que vio Ana?
 D.N C saw Ana</p> |
| <p>(7) Leio o que quer que ela escreva.
 read.1sg D.M what ever that she writes
 ‘I read whatever she writes.’</p> | <p>(8)*Leo lo que-quiera que escriba.
 read.1sg D.N QUE-ever that writes
 Intended: ‘I read whatever (she) writes.’</p> |
| <p>(9) Levei o que de bonito Ana comprou.
 I.took D.M what of pretty Ana bought
 ‘I took what was beautiful that Ana bought.’</p> | <p>(10) Comprei o que que ela pediu.
 I.bought D.M what that she asked
 ‘I bought what she asked for.’</p> |
| <p>(11)[DP \emptyset_D [CP <i>o que</i>_i [CP (<i>que</i>) ... <i>t_i</i> ...]]]</p> | <p>(12)[DP <i>lo</i> [_{nP} \emptyset_n [CP <i>OP_i</i> <i>que</i>_{C⁰} ... <i>t_i</i> ...]]]</p> |

2. Matching. A hallmark of FRs are matching effects, whereby the *wh*-pronoun has to satisfy the selectional restrictions of both matrix and embedded verb. Both BP and Spanish *lo que* display matching effects, (13)-(16), but the latter challenges Agree-based accounts of the phenomenon.

- | | |
|---|--|
| <p>(13) <i>Match (BP)</i>
 Eu falo <u>de</u> d-[o que Ana fala <u>de</u>]
 I talk about-what Ana talks</p> | <p>(14) <i>Mismatch (BP)</i>
 *Eu amo <u>DP</u> [do que Ana esta rindo <u>de</u>]
 I love of-what Ana is laughing</p> |
| <p>(15) <i>Match (Sp)</i>
 Hablo <u>de</u> de [lo que habla <u>de</u> Ana]
 speak.1SG of DEF.N C speaks Ana
 ‘I talk about what Ana talks’ (e.g. a cake)</p> | <p>(16) <i>Mismatch (Sp)</i>
 *Amo <u>DP</u> [<u>lo</u> de (lo) que se <u>ría</u> <u>de</u> Ana]
 I.love D.N of REL C REFL laughs Ana
 ‘I love about what Anna is laughing.’</p> |

a) BP. On Agree approaches (Caponigro 2002 et seq.), matching arises from a dependency between the FR’s Op and the null D, triggered by one of them bearing an unvalued feature, (17). Building on Brito (2024), we argue that this dependency involves a categorial feature: the null D that gives the FR its D properties is a weak head (Oda 2024) that must be licensed via Agree with an Op with a valued D feature (a DP), capturing (13) and ruling out pied-piping in (14), where Op is a PP. The embedded P in (13) is deleted when stranded in PF – a very productive process in BP (Kato 2010).

(17) $[DP \emptyset_{[D: _]}] [CP \text{ o que}_{[WH: _, D: +]}] [CP C_{[WH: +]}] \dots \Rightarrow [DP \emptyset_{[D: +]}] [CP \text{ o que}_{[WH: +, D: +]}] [CP C_{[WH: +]}] \dots$

b) Spanish. While matching effects look similar in Spanish, they cannot easily be attributed to structural mechanisms like Agree, due to the different structure of the *lo que* construction. First, the Agree-based analysis of matching relies on agreement between Op and a null D – but Spanish *lo que* does not contain a null D, but a null *n*, cf. (12). Second, the mismatching structure to be ruled out is unlike the BP one: a sequence like (18a) is independently excluded, since it involves the movement of the embedded preposition beyond SpecCP to SpecDP across an *n* layer – which is unavailable in Spanish. Instead, we want to rule out (18b), where a P+RelOp has been pied-piped to SpecCP – a configuration not contemplated by previous Agree-based accounts of matching. Third, adapting the Agree-based mechanism to tackle these two issues both under- and overgenerates. The adaptation would involve stipulating category Agree between the empty *n* and Op, as in (19). This stipulation overgenerates because light-headed relatives, which also involve an empty *n*, do not show matching effects, cf. (21). On the other hand, it undergenerates because it fails to capture that mismatches like (18b) are also ungrammatical with noun phrase ellipsis (NPE), cf. (21). As NPE results from deletion of full *n*Ps in PF, (21) is structurally identical to headed RCs – which don’t show matching. Considering this, we propose that the source for matching in Spanish is a PF ban on $K_1 X K_2$ structures, where *K* is the case-receiving head in nominals (Norris t.a.) and *X* is phonologically light (e.g. *lo*, but not *aquello*). We assume Ps like *de* are spell-outs of K_{obl} and that *K*s always pied-pipe, yielding $[K_1 [lo \ n \ [[K_2 \ Op]_i \ que \ \dots \ t_i]]]$. $K_1 X K_2$ violations are circumvented by deletion of K_2 , subject to recoverability. Matching is, thus, a byproduct of the identity requirement on the deletion of the pied-piped *K*. This works for K_{obl} in (15). Overt sequences like D- \emptyset_n -P-RC in (18b) are out because D is preceded by *K* (*nom* or *acc*), which does not license deletion of P/ K_{obl} .

(18) a. *Mismatch 1:*

b. *Mismatch 2:*

* $[DP [de \ Op]_1 \ lo_D \ [nP \ \emptyset_n \ [CP \ que \ A. \ laugh_de \ t_1]]]$ * $[DP \ lo_D \ [nP \ \emptyset_n \ [CP [de \ (Op)]_1 \ que \ A. \ laugh_de \ t_1]]]$

(19) $[DP \ lo_D \ [nP \ \emptyset_n[N: \ _]] \ [CP \ Op_1[N: \ +] \ que \ Ana \ cooked \ t_1]]] \Rightarrow [DP \ lo_D \ [nP \ \emptyset_n[N: \ +]] \ [CP \ Op_1[N: \ +] \ \dots]$

(20) What do you love? *Amo_DP* (‘I love...’)

aquello **de** lo que se ríe **de** A.

DIST.N of REL.N C REFL laughs A.

‘I love that about which A. laughs.’ (LHR)

(21) What thing do you love? *Amo_DP* (‘I love...’)

*la eosa **de** la que se ríe **de** A.

D.F thing.F of REL.F C REFL laugh A

‘I love the one A. laughs about.’ (NPE)

Implications. The above facts have wider typological implications. First, the typology of headless RCs should be expanded to include an empty *n* option to capture Spanish *lo que*, enriching Caponigro (2023), who only proposes FRs, LHRs, and superfree RCs. Second, Spanish is problematic for theories of matching (e.g. be they based on Agree or relabeling) because the structural conditions required for matching in such theories are not satisfied by *lo que*. While phonological approaches to matching exist (Fuß & Grewendorf 2014), Spanish *lo que* constitutes a case where matching can *only* be captured morphophonologically, (i.e. via deletion, to avoid the ban on D - \emptyset_n - P-RC), and not with purely syntactic means. This provides grounds for reconsidering morphophonological analyses of matching more generally, including for languages where structural approaches work.

References: Brito, A. (2024). *The syntax of FRs with invariable “o que” revisited*. *RAPL*, (11), 88-108. Citko, B. (2004). *On headed, headless, and light-headed relatives*. *NLLT*, 22(1), 95-126. Gutiérrez-Rexach, J. (2014). *Interfaces and Domains of Quantification*. OSUP. Caponigro, I. (2002). Free Relatives as DPs with Silent D and a CP Complement. In: V. Samiian, ed., *Proceedings of WECOL 2000*. California State University, 140–150. Caponigro, I. (2023). Referring and quantifying without nominals: Headless Relative Clauses across languages. In J. Kim, B. Öney, Y. Zhang, and L. Zhao (eds), *Proceedings of the 33rd Semantics and Linguistic Theory Conference (SALT 33)*, 744-774. Cecchetto, C & C. Donati (2015). *(Re)labeling*. MIT Press. Fuß, E., & G. Grewendorf (2014). Freie Relativsätze mit d-Pronomen. *Zeitschrift für Sprachwissenschaft*, 33(2), 165-214. Groos, A. and H. van Riemsdijk (1981). Matching Effects with Free Relatives: A Parameter of Core Grammar. In: A. Belletti, L. Brandi and L. Rizzi, (eds.), *Proceedings of GLOW*. Scuola Normale Superiore, Pisa, 171–216. Grosu, A. (2003). A Unified Theory of Standard and Transparent Free Relatives. *NLLT*, 21(2), 247-331. Himmelreich, A. (2017) Case Matching Effects in FRs and Parasitic Gaps. PhD, U.Lepizig. Kato, M. 2010. Optional prepositions in Brazilian Portuguese. In K. Arregi et al. (eds.), *Romance Linguistic: Interactions in Romance*, 171-184. John Benjamins. Marchesan, A. (2012). *As relativas livres no PB*. PhD, UFSC. Norris, M. (t.a.), Nominal inflection in Distributed Morphology, *Cambridge Handbook of Distributed Morphology*. Oda, H., 2024. Large-scale pied-piping in the labeling theory and conditions on weak heads. *The Linguistic Review*, 41(1), 153-185. Ojea, A. (2013). *A uniform account of headless relatives in Sp*. *Lang. Sciences*, 40, 200-211 Plann, S. (1980). *Rel. clauses in Sp without overt antecedents*. PhD, UCal. Saab, A. (2008) *Hacia una teoría de la identidad en la elipsis*. PhD, UBA. Šimík, R. t.a. *From interrogatives to relatives*. In *Cambridge Handbook of Compar. Syntax*.