Claus D. Pusch (Freiburg im Breisgau)

Relative pronoun reduction and resumptive pronouns in spoken Catalan. A corpus-based study

Relative clauses and relative clause constructions have been attracting the interest of linguists for a long time, and the relevant bibliography available on this subject is immense.[1] The attraction of the 'marvelous world of relative clauses' ("el maravilloso mundo de los relativos", Osuna García, 2005: 19) is certainly due to the variability and complexity that relative constructions in the better described (European) languages are subject to, with the particularly elaborate Latin relative-clause syntax serving as a model for traditional grammaticography as well as for many modern approaches to the typology and to the structural and functional description of relative clauses (cf. Touratier, 1980; Lehmann, 1984). As Osuna García (2005: 19s.) notes, there have been linguists, like Rodolfo Lenz, who considered the existence of relative clauses and their syntactic integration through a paradigm of relative pronouns (in the sense of traditional Latin grammar) as a sign of developmental 'maturity' of this language and of the ethnolinguistic community who speaks it, an idea that the quoted author – Osuna – seemingly still adheres to to a certain degree.

However, recent typologic studies based on balanced world-wide samples of languages have clearly shown that the relativization strategy used in Latin and many other European languages is a rather exotic morphosyntactic feature, and that it is so densely concentrated among European tongues that the hypothesis of an areal ("Sprachbund") phenomenon seems more than plausible. This areal concentration becomes manifest when, for instance, the data of the "World Atlas of Language Structures" (Haspelmath *et al.*, 2005) is taken into account. Maps 122 (cf. fig. 1) and 123 (cf. fig. 2) of this atlas, which include data from 112 languages, leave us

---

with no doubt that the 'relative pronoun strategy' familiar to us and described by the authors of the maps (Comrie / Kuteva, 2005: 494) in terms of a strategy where "the position relativized is indicated inside the relative clause by means of a clause-initial pronominal element, and this pronominal element is case-marked (by case or by an adposition) to indicate the role of the head noun within the relative clause", is, on a global level, a very minoritarian one, whereas alternative strategies such as Comrie / Kuteva's 'non-reduction strategy' ("the head noun appears as a full-fledged noun-phrase within the relative clause" [*op.cit.*: 495], hence doing without pronominalization) or especially the 'gap strategy', which "involves cases where there is no overt case-marked reference to the head noun within the relative clause" (*op.cit.*: 495), are by far more frequent.
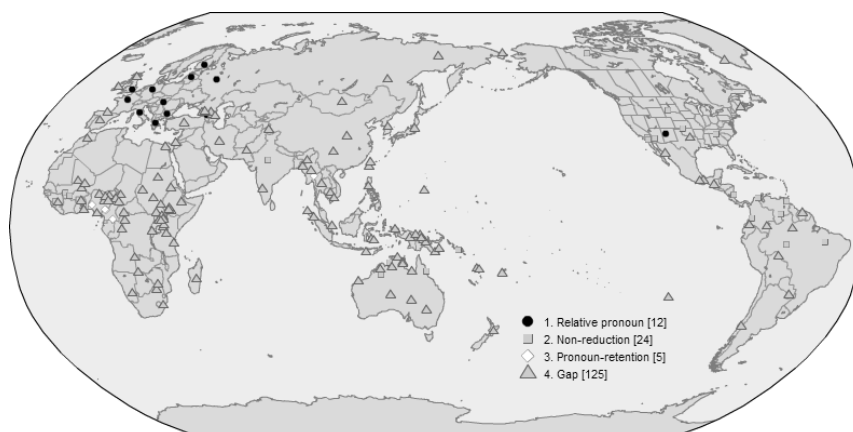

Fig. 1. Strategies used to relativize on the subject in the languages of the world (WALS; Comrie / Kuteva, 2005)

The scarcity of the relative pronoun strategy among the languages of the world may be attributed to the structural complexity and the heavy functional load of the relative pronoun in a traditional (Latinist) perspective, characteristics of European-style relativization positively described by Osuna García (2005: 20) as a means of extraordinary grammatical efficiency ("una forma gramatical [...] de una extraordinaria rentabilidad") and briefly summarized, on the basis of Lehmann's (1984; 1995) fundamental account, in chap. 1 of this paper. However, one has to bear in mind that most descriptions of relativization in Latin are based on data from the
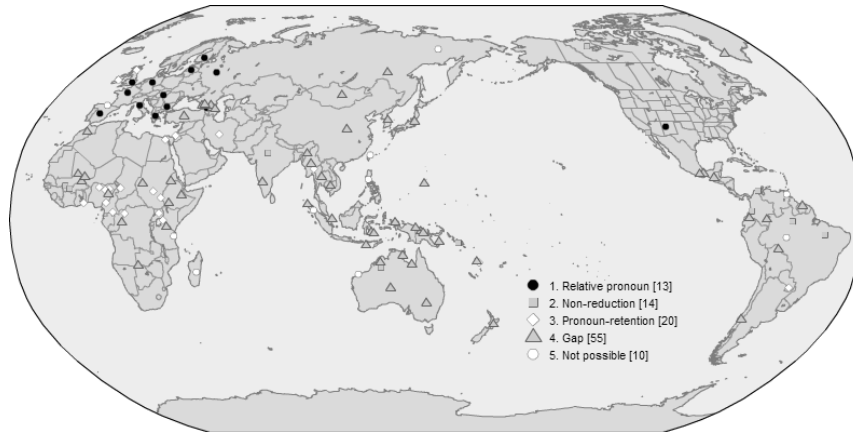
Fig. 2. Strategies used to relativize on oblique constituents in the languages of the world (WALS; Comrie / Kuteva, 2005)

sometimes excessively elaborate and highly artificial written form of that language, known as Classical Latin, whereas less elaborate vernacular varieties and, more particularly, spoken Latin had a system of relativization strategies perhaps not 'simpler' but structured in a different way. The same holds for the diachronic successors of Latin, i.e. Romance languages, where there is a relevant divergence between relativization strategies put to use in the written vs. the oral varieties, with spoken language strategies coming closer to the more frequent techniques found in the world-wide samples of typological research, e.g. the above-mentioned gap strategy or Comrie / Kuteva's 'pronoun-retention strategy', not commented on hitherto, where "the position relativized is explicitly indicated by means of a resumptive personal pronoun" (*op.cit.*: 495).

As mentioned before, the relative clause and the variability of relativization strategies has attracted continuous interest from general linguistics and namely from formally oriented approaches to syntax (cf. Alexiadou *et al.*, 2000, and Bianchi, 2002, for an overview). In the realm of Romance, both comparative accounts of relative-clause formation (e.g. Cid Abasolo, 1999; Fiorentino, 1999; Fiorentino, 1998, on oral strategies; and, especially, Schafroth, 1993, still, in my view, the most complete account of oral and written varieties in all Romance languages from both a synchronic and diachronic perspective) and approaches to the relativization systems of indi-

vidual languages are available: Cinque (2001 [1988]), Fiorentino (1999) and Scarano (2002) describe the Italian system, the latter two works including extensive non-standard data. As for French, Riegel / Pellat / Rioul (2001 [1994]: 479–489) give a survey of the relative clause in the standard variety (cf. aussi Godard, 1988), and the semantic distinction between restrictive and appositive relative clauses has been widely discussed on the basis of this language (cf. Kleiber, 1987; articles in Fuchs (ed.), 1987), but the interest for non-standard relativization strategies has been particularly strong in French language studies (cf., apart from the short but fundamental paper by Guiraud (1966), numerous contributions by Gadet (1997 [1988]: 115ss.; 1995; 2003), and Gapany (2004)). For Spanish, apart from the already mentioned monograph by Osuna García (2005), Brucart (1999) has to be mentioned as the most comprehensive presentation of relativization in this language; finally, for Catalan as the language that is in the center of interest of the present contribution, we owe the most detailed and best documented analyses of the phenomenon of relative-clause formation and variation to Solà (e.g. 1972 and, most notably, 2002).

In the light of this large body of literature on the subject, the objective of the present contribution can only be a rather modest one: After a brief overview of the functions and ensuing structural properties of relative clauses and relative pronouns according to Lehmann (1984) (chap. 1) and an equally brief description of general tendencies observable in the diachronic development of relativization strategies from Latin to Modern Romance, with a focus on so-called 'deviant' strategies typical for the oral language (chap. 2), the main types of non-norm-compliant, 'deviant' relativization patterns, found in Romance in general, will be illustrated and discussed on the basis of data from recent corpora of spoken Catalan (chap. 3). For the time being, only relative clauses involving the (unstressed) relative pronoun / particle *que* (< Lat. QUE(M)) will be taken into account. A complete description of the system of relative clause formation in spoken Catalan, both in quantitative and qualitative terms, is beyond the scope of the present article, but it may constitute a first element of such a description.

1 Relativization and the functions of relative pronouns

Relative clauses are subordinate clauses. The aim of syntactic subordination is to integrate a sentence into another sentence, for which a kind of rank shift (Lehmann, 1995: 1200) is necessary: the sentence to be inte-

grated is transformed in order to function like a nominal constituent of the other sentence into which it is supposed to be inserted. Therefore, subordination is a process of nominalization. The nominal character of a relative clause becomes obvious through its ability to commute with a prototypically nominal element as in (1):

(1)   a.   "El Buli" is a restaurant that everybody knows
      b.   "El Buli" is a restaurant known by everybody
      c.   "El Buli" is a renowned restaurant
      d.   "El Buli" is a famous restaurant

As can be seen from these invented examples, the degree of nominalization or 'nominality' is a scalar phenomenon, with the relative clause in (1a) representing the less nominalized (and most sentential) case, the participial constructions in (1b–c) intermediate cases and the attributive adjective in (1d) being the less sentence-like and thus most nominal(ized) variant. These examples also make clear that in most cases the aim of the relative clause construction is to integrate a sentence into another sentence as an adnominal attributive. In these instances, the term 'attributive clause' instead of the somehow vague term 'relative clause' might be a good alternative (Lehmann, 1995: 1201). Adnominal relative clauses are characterized by the presence of a head noun in the superordinate (matrix) sentence that also appears in some function or position in the subordinate clause, therefore constituting the 'hinge' that allows the relativized segment to be integrated attributively into the matrix. This hinge element or head noun is called 'nucleus' in Lehmann's (1984; 1995) description of relative clauses.

    The nominal that forms the nucleus of the relative construction has a semantic and syntactic role in both the matrix and the subordinate clause, and this role may be morphologically manifest through case-marking and agreement. The peculiarity of 'European-style' relativization is the fact that in the relative clause a relative pronoun takes the place of the nucleus noun and may also 'take over' some of the morphological marking of the syntactic function of this noun, as can be seen in (2):

(2)   a.   Ferran Adrià is a great cook. I really admire this cook.
      b.   Ferran Adrià is a great cook. I really admire him.
      c.   Ferran Adrià is a great cook, whom I really admire.

English has no inflectional case marking on nouns, but variant (2b), equivalent to (2a) but involving an object pronoun, makes obvious how the relative pronoun in (2c) 'inherits' some morphological case-marking from the nucleus constituent that it pronominalizes in the subordinate clause. The relative pronoun is a real pronoun in the sense that it actually represents (structurally and functionally) the nucleus within the relative clause, i.e. it opens a syntactic gap at the position otherwise occupied by the nominal nucleus and, at the same time, fills that gap left by the nominal nucleus, whose lexical expression subsists in the matrix sentence, to which the relative pronoun is anaphorically related. The complexity of the 'European-style' relative construction is therefore due to the multifunctional character (Lehmann, 1984: 248) of the relative pronoun, which combines the three basic operations of relative-clause formation according to Lehmann: subordination, identification of the syntactic gap, and attribution ("die für die R[elativ]S[atz]bildung konstitutiven Operationen [...]: Nominalisierung (Subordination), Attribution (Nukleusbildung) und Leerstellenbildung" [Lehmann, 1984: 246]).[2] With this triple mission, the relative pronoun bears a heavy functional load or – expressed in a positive way, as mentioned above – reveals itself to be of an 'extraordinary grammatical efficiency'. However, structural efficiency does not necessarily correlate with functional efficiency, as far as language production and processing is concerned, and for this reason many European languages and, more specifically, their non-standard vernacular varieties have resorted to alternative ways of handling these basic operations of relative clause formation.

2 Strategies of relative-clause formation in spoken Romance

As is well known, Classical Latin (CL) had a full-fledged paradigm of case-marked relative pronouns that furthermore marked the inflectional categories of gender and number in a straightforward way. The case-marking allowed the relative pronoun to identify the syntactic role of the head noun

---

2   It must be emphasized that Lehmann (1984; 1995) normally speaks of 'formation of a syntactic gap' ("Leerstellenbildung"), whereas my summary of his approach has insisted on the 'filling' of the syntactic gap ('Leerstellenbindung') operated through the relative pronoun. This does not seem a contradiction to me: the head noun is deleted from the subordinate clause in order to allow this subordinate clause to be hitched up to a co-referential noun occurring in the matrix sentence; with a saturated argumental structure in the clause, such a process of clause conjunction would not be possible. At the same time, through its morphological apparatus, the relative pronoun is able to replace the deleted head noun syntactically and, in doing this, to fill the gap to a certain extent.

in the subordinated clause ("Leerstellenbildung"), whereas number and gender agreement features established the attributive relation with the head noun in the matrix. The subordinating capacity of the CL relative pronouns may be attributed to the nominalizing function of the Indo-European roots *kwo- / *kwi- that they derived from (Schafroth, 1993: 59s.). The Spoken ("Vulgar") Latin (SL) system, in comparison with this elaborate written paradigm, was somehow reduced and open to restructuring, which actually occurred on the way from SL via Proto-Romance to Early Romance (cf. Schafroth, 1993: 60ss., with further references) and resulted in a tripartite paradigm including the forms QUI, QUE(M) and CUI. This pronominal paradigm was still able to mark subordination / nominalization, but could no longer mark attribution through nominal agreement; it was able, though, to identify the syntactic role of the head noun, at least to a large extent, as QUI was associated with subject case, QUE with direct object case, and CUI with indirect object (oblique) case; however, the oblique case function seems to have been expressed from a very early stage on through combinations of CUI and preceding prepositions (Schafroth, 1993: 72s.). In Late Latin, the locative adverbs UBI and UNDE appear in relativizing contexts and take over, in some Romance languages, functions of the pronominal elements of the tripartite paradigm mentioned before. In Medieval Romance, a new Latinizant type of relative pronouns on the basis of the definite article + QUALIS was created, irradiating from scripta traditions mainly in Western (Gallo-)Romance (Kunstmann, 1990). This new learned paradigm allowed, once again, to straightforwardly mark the three basic operations of relative-clause formation according to Lehmann in one (albeit complex) morpheme, but it remained restricted to written and formal registers. The less formal registers of spoken Romance gave preference to an alternative strategy of indicating these operations, i.e. to split up the subordinating, the attributive, and the syntactically identifying functions of the relative pronoun and to express these functions analytically through separate markers. This strategy of relative-clause formation is known in Romance linguistics under the French label "décumul", introduced by Guiraud (1966: 41), who illustrates it among others with the following example (3b) in contrast to (norm-compliant) (3a):

(3) a. Je viens te donner de nos nouvelles$_i$ qui$_i$ sont très bonnes
     b. Je viens te donner de nos nouvelles$_i$ qu'elles$_i$ sont très bonnes

In this decumulative relative-clause formation strategy, the conjunctive element *que* only carries out the operation of subordination / nominalization. That – out of the three relative pronouns that the Romance system inherited from CL / SL – it was QUE(M) that acquired this function was facilitated by its formal convergence with the most frequent conjunction used to introduce complement clauses in Romance, i.e. the *que* which had developed out of QUOD / QUIA. Obviously, from this moment on it does not really make sense to call relative-clause initial *que* a relative pronoun, as it doesn't agree with nor identify, through agreement features, any noun; many scholars prefer to speak of *que* as a relative particle. The operation of identifying the syntactic function of the head noun / nucleus in the relative clause is taken on, in the decumulative construction, by a (generally un-stressed) anaphoric pronoun or pronominal adverb which has been called, since Lehmann (1984), a resumptive pronoun.[3] If the pronoun used for resumptive purposes carries gender and / or number features, it takes on both the function of marking the syntactic gap and that of attribution; if it does not carry agreement features, it only marks the gap by identifying the syntactic function of the nucleus.

The decumulative construction, which is akin to Comrie / Kuteva's (2005) pronoun-retention strategy mentioned above, is only one alternative strategy found in spoken Romance languages to reduce the functional load of the relative-clause-initial conjunctive element. A more radical strategy consists in simply marking neither the operation of syntactically identifying the nucleus' position / function nor that of attributing the relative clause to it, as (4c) in the following group of examples, taken again from Gui-raud's (1966: 40) paper on spoken French, illustrates:

(4)   a.   l'homme$_i$ dont$_i$ je vous parle (norm-compliant)
      b.   l'homme$_i$ que je vous en$_i$ parle (decumulative)
      c.   l'homme$_i$ que je vous parle

This strategy, where only the operation of subordination is marked through the (homonymous) relative/complementizer particle *que*, is de-

---

3   Actually, Lehmann (1984: 97ss.) presents the notion of resumptivity slightly differently: for him, a resumptive pronoun in a relative construction must have nominal features and has to represent the nucleus (sc. its syntactic function) in the relative clause (cf. *op.cit.*: 97). In this sense, relative pronouns like those in the CL paradigm are also resumptive pronouns. In Romance linguistics, however, the term is generally used only for non-subordinating pronouns in decumulative or similar constructions.

scribed as "reduction of inflection" ("réduction de la flexion") by Guiraud (1966: 72) and, more handily but ambiguously, called "defective relative-clause formation" elsewhere (Gapany, 2004: 126ss.). It corresponds to Comrie / Kuteva's (2005) gap strategy.

Guiraud (1966: 42s.) mentions one more strategy of relative-clause formation typically found in oral varieties, described as pleonastic relative clauses (cf. also Gapany, 2004: 128ss.) and explained by Guiraud as hyper-corrections; (4d) exemplifies this type of relative clause:

(4)   d. l'homme$_i$ dont$_i$ je vous en$_i$ parle

In cases like this, there is still separation of the basic operations of relative-clause formation *à la* Lehmann in the sense that the relative pronoun indicates subordination, and the – again – resumptive pronoun within the relative clause may indicate attribution, but the operation of marking the syntactic function of the nucleus is taken on by both the conjunctive and the resumptive pronominal element. As Guiraud's judgment of 'hyper-correction' suggests, this double marking is not generally considered as comparable to the decumulative and the defective types of relative clauses, as no effect of 'simplification' is involved, but as an interference of the (written) norm. According to the WALS maps, pleonastic relative clauses do not seem to be a relevant 'default' type of relative-clause formation in the languages of the world.

It must be emphasized that, although the 'deviant' types of relative clauses (deviant from a Euro-centric scripturality-biased point of view) have been studied mainly in various modern varieties of Romance, these alternative strategies are attested in SL from the Classical period of Latin onward, as Lehmann's (1984: 389ss.) data prove.

3 Strategies of relative-clause formation in spoken Catalan

The fact that relativization strategies in oral Catalan differ from those used in written Catalan is taken into account in all recent grammars of that language which take a descriptive stance. Badia i Margarit (1994), who treats relative clauses under the heading "adjectival clauses" ("oracions adjectives", *op.cit.*: 357ss.), mentions the 'deviant' forms of relative clauses as strategies put to use in spontaneous speech ("[a] la llengua més espontània", *op.cit.*: 369) in order to avoid the use of the complex learned relative pronouns of the ART + QUALIS type (*el / la qual, els / les quals* in Catalan).

Whereas Badia i Margarit, at this point, concedes large social spreading of these 'deviant' forms – namely the decumulative one – ("aquesta via és la més usada en la llengua col·loquial (i àdhuc en la parlada per persones cultes)", *op.cit.*: 369), at another place he calls them outright incorrect (cf. *op.cit.*: 166). Wheeler / Yates / Dols (1999: 536–538) dedicate a paragraph within the chapter on relative clauses to the "Non-standard/colloquial constructions with *que*" and mention all the patterns of oral relativization techniques outlined above with reference to Guiraud; "[a] grasp of these patterns", they emphasize, "is important in order to understand everyday spoken Catalan" (*op.cit.*: 536). The authors insist on the diamesic dualism and avoid designating the 'deviant' strategies as incorrect or explicitly discouraging their use.

The most comprehensive and most differentiated treatment of the relativization strategies prevailing in spoken Catalan, however, is to be found, as mentioned above, in Solà (2002: 2512–2533), who under the title "Relative clauses with pronominal duplication" ("Relatives amb duplicació pronominal") treats the subject in great detail, with a huge number of examples at hand. This thorough treatment may be surprising, if the reader of Solà (2002) recalls the opening sentence of the large chapter dedicated to relative clauses, where the author announces that he will analyze the standardized uses and forms ("En aquest capítol s'estudien els usos estàndard de les construccions relatives", *op.cit.*: 2459). However, what Solà tries to make clear is that this field of Catalan syntax remains to be fully explored, that things are less than clear, and that therefore there is no really established standard yet; at more than one point he describes the area of relative-clause formation as a frontier zone of linguistic description of Catalan ("un terreny fronterer", *op.cit.*: 2526). And, above all, he insists on the repercussions of 'oral' strategies in writing (and vice versa, although less prominently).

The authors of the three cited grammars rely, in their description of spoken / colloquial / spontaneous styles of relative-clause formation, on examples that are invented or randomly selected from different sources and, in most cases, not referenced. This does not necessarily weaken their analyses. However, as anyone familiar with spoken Catalan is left with the impression that 'deviant' (reduced / 'defective', decumulative and pleonastic / 'hypercorrect') forms of relative clauses are a frequent phenomenon, it seems justified to approach this area on the basis of attested oral uses and to subject the different strategies to a closer, corpus-based examina-

tion in order to evaluate their status in contemporary Catalan at least in qualitative terms. This is the aim of the present chapter.

3.1 The data

The data used for this analysis comes from two recently published corpora of Catalonian Catalan, both of which form part of the large corpus project carried out at the University of Barcelona under the abbreviation CUB ("Corpus de Català Contemporani de la Universitat de Barcelona"; see Alturo / Boix / Perea, 2002, for a general presentation). The first of these (Payrató / Alturo, 2002) contains transcriptions of selected texts from the Corpus of Colloquial Conversation ("Corpus oral de conversa col·loquial" (COC); see Oller *et al.*, 2000, for details), amounting to some 70.500 words (out of 357.500 of all the texts recorded in this corpus). The second resource used in this study belongs to the Corpus of Oral Registers (COR) of the CUB project and contains transcriptions from a large array of oral text genres, ranging from more formal registers such as media talks or court or civil services encounters to informal ones such as private conversations (cf. Alturo / Boix / Perea, 2002: 161ss.). The published selection (Alturo *et al.*, 2004) contains some 154.000 words (out of 347.000 in the entire corpus). The total amount of spoken language data used here corresponds therefore to a mere 225.000 words, which is obviously a rather modest quantity but still, for the time being, the most important published oral resource available for Catalan.[4] The transcriptions of both published selections come in machine-readable form on CD-ROMs, which also contain the corresponding audio data.

For the present study, the transcription files of the CD-ROMs have been converted to plain text format and queried with the MonoConc Pro 2.0 concordance tool (Barlow, 2000). Unfortunately, as neither of the two corpora, referred to from now on as COC and COR, is morphologically tagged or syntactically annotated, purely lexical searches turned out to be

---

4   One has to bear in mind that the recently published multimedia corpus C-ORAL-ROM (Cresti / Moneglia, 2005), which claims the status of a spoken reference corpus for the four major Romance languages (French, Italian, Portuguese, and Spanish), contains not more than 300.000 words for each individual language either. To the regret of the corpus designers of C-ORAL-ROM, Catalan could not be included in this project due to a lack of funding (*op.cit.*: xiii). – The third component of CUB, the Corpus of Dialects ("Corpus Oral Dialectal", COD; Viaplana / Perea, 2003), a small part of which has become publicly available, has not been taken into account in this study.

the only reliable search strategy. For the present study, only occurrences of the relative pronoun / particle *que* have been searched. The query produced a tremendous number of hits (some 2200 in COC and 3600 in COR), which obviously included a huge amount of non-relevant occurrences such as *que* complementizer for complement clauses, *que* as part of complex adverbial conjunctions (*abans que* 'before', *encara que* 'although' etc.), *que* in comparative constructions (equivalent to English *than*) etc. After sorting out manually these non-relative uses of *que*, only a few hundred occurrences remained, from which the 'ordinary' (and totally norm-compliant) cases such as (5–6), where *que* introduces a relative clause, the head noun of which occupies the subject or direct object (DO) position in the relative clause, were eliminated and 'deviant' and otherwise more or less remarkable cases were – again manually – extracted.

(5)   e:ls ajuts e: directes (... 1.28) m:: (. 0.25) comporten el perill (.. 0.30)
       de: (.. 0.42) un (. 0.17) intervencionisme (.. 0.61) o d'un d'un risc un
       risc de dirigisme o de intervencionisme (.. 0.46) per part (... 1.38) de
       l'administració *que* atorga els diners (COR)[5]
(6)   doncs no ho sé\ com per exemple els els Continente:s o les Glòries
       que tenim ara aquí molt a la vora (... 1.18) i això fa que el:: (... 1.23) que
       sigui un_ un difíci- un difícil competidor per les botigues (COR)

The resulting working corpus of this study consists of 74 relative clauses introduced by *que* in a non-subject and non-DI gap position. Obviously, from such a small data-base, no quantitative conclusions can be derived, with the following analyses indicating mere tendencies or being limited to case studies only.[6] The figure suggests that the overall frequency of such

---

5   The authors of the CUB corpora make use of a quite differentiated transcription
     scheme based on a proposal by John DuBois *et al.* developed in the 1990-s, which is
     detailed in the published corpora. Most of the transcription features have been main-
     tained in the examples from COC and COR quoted in this paper, including the fol-
     lowing: \ = descending final tone; / = ascending final tone; _ = maintaining final tone;
     — = truncated tonal group; : :: ::: = (variable) lengthening of a sound; (.) (..) (...) =
     breaks (of variable length, with exact length in seconds); (e) = skipped sound; (?) = un-
     sure transcription; - = truncated word.
6   It would be a challenge to submit the complex relative-clause system of Catalan to real
     quantitative analyses and statistical tests, of the type applied by Biber *et al.* (1999: 608–
     630) to the (British and American) English relative clause. Biber *et al.*'s study – as their
     whole grammar – is based on a 40 million word corpus, out of which some 12 million
     words belong to conversational and other oral data. For Catalan, the 52 million word

remarkable or 'deviant' instances of relative clauses in spoken Catalan seems to be considerably lower than one would suppose at first glance.

3.2 Decumulative relative clauses

Decumulative relative clause formation, where the different operations involved in relativization are distributed between separate morphemes but without any redundant marking, is rather scarce in our corpus data. Some instances of *que* (marking subordination) + clitic pronoun (marking attribution and gap position) were found where the head noun occupied the indirect object (IO) position in the subordinated clause, as in (7–9):

(7) lo que s'està ficant molt és que van (.. 0.66) a un local gai van (.. 0.56) els gais p(e)rò també van molt molt tio$_i$ *que li$_i$* agrada la música m:aquinera que foten (COC)

(8) és gent que beu molts cubates és gent$_i$ *que* no *li$_i$* importa gastar-se cinc mil peles en una nit sis mil nou mil (COC)

(9) és perquè és un nen$_i$ (.. 0.76) *que li$_i$* falta el llenguatge (.. 0.38) p(e)rò que té una actitud dins de: el grup (. 0.20) per dir-ho (ai)xí\ que jo intueixo intueixo (.. 0.43) que: (.. 0.66) afavorirà el funcionament\ i clar (COR)

The head-noun / nucleus in (7), *tio* 'guy', is used here as a collective noun in the sense of 'many people (of that kind)' comparable with *gent* (used in (8)); there is no number agreement between *tio* and the matrix verb, something not uncommon with this kind of – grammatically singular but logically and semantically plural – noun and frequently found with *gent* also, as in (10; without 'deviant' relative clause):

(10) p(e)rò vull dir *gent* que no *fan* ostentació\ bueno\ (COC)

but number agreement holds between *tio* and the verb in the relative clause, leading to the use of *li* as resumptive pronoun. No kind of agreement inconsistency occurs in (8) with the collective noun *gent*.

"Corpus Textual Informatitzat de la Llengua Catalana" accessible on the Institut d'Estudis Catalans' web-site <http://www.iec.cat> may be a suitable (although not entirely comparable) source for written data, but the lack of equivalent oral data would lead to serious problems.

Despite the low number of examples, it might be worth pointing out that in all cases of decumulative relative clauses with the clitic in IO gap position, the head-noun is [+human]; sentences such as (invented) (11) do not occur in our data:

(11) ens han donat un cotxe$_i$ que li$_i$ faltava gasolina

This is in accordance with Solà's (2002: 2523) finding (referring back to studies by C. Silva-Corvalán) that animacy of the head-noun favors pro-nominal resumption, and with a general repugnance in spoken language of using personal pronouns (namely agreement-sensitive ones) for non-ani-mate reference (Thun, 1986).

Only one example of a partitive adverbial pronoun *en* in a decumulative relative clause was detected in the corpora. The example is complex, how-ever, and does not represent a clear-cut case of decumulation:

(12) moltes vegades la presència de la flor és un record de Xile\ (. 0.24)
       m/ del seu Xile natal\ del país d'origen\ (. 0.27) un país *que* ell sem-
       pre que en parla (. 0.12) *en* parla com un país de grans contrastos (.
       0.18) eh\ de una exuberància floral (. 0.19) molt gran\ (COR)

In this example, an adverbial clause (AC) introduced by *sempre que* 'always that, whenever' is inserted into the relative clause (RC), which depends on the matrix (M) *un país*, leading to the following structural bracketing:

(12) a. [$_M$un país$_i$ [$_{RC}$que ell [$_{AC}$sempre que en$_i$ parla] en$_i$ parla com un país
          de grans contrastos]]

Represented like this, the relative clause *que ell en parla...* really looks like being of the decumulative type, with subordinating *que* and gap-filling (but not attributing) *en* replacing an oblique relative pronoun *de què* or *del qual.* However, the resumptive *en* – as noted, the only example in the corpus – may also be explained as a parallelism to the *en parla* found in the adverbial clause.

The corpus data contain a certain (but, again, rather low) number of decumulative relative clauses with *que* + locative adverbial pronoun *hi.* This pronominal element, just as the aforementioned *en*, has had a very eventful life in Ibero-Romance and, contrary to its cognate in Castilian, remains in frequent use in modern Catalan (cf., e.g., the still fundamental work of

Badia Margarit, 1947). The high recurrence of *hi* and *en* in the modern language is partly due to the fact that these pronouns have developed lexicalized uses in relation with certain verbs, where the pronominal value of the pronoun has weakened considerably but where an aspectually relevant adverbial meaning subsists. This is the case with *anar-se'n* 'to leave, to go away', *sortir-se'n* 'to cope with sth., to solve (a problem)', *entendre-hi* 'to have a knowledge of sth., to understand sth.', *veure-hi* 'to have sight', and the very frequent existential *haver-hi* in its impersonal 3S form *hi ha* 'there is sth., sth. exists'. Solà (2002: 2513) reminds us that in the case of a relative clause with such a lexicalized 'incorporated' pronoun *hi* (or *en*), one has to carefully check whether the pronominal element is really resumptive or not.

In practice, however, this decision is not an easy one. *Que + hi* might be an alternative decumulative construction for the adverbial relative pronoun *on* 'where', which is replaceable by PREP + *què* or PREP + ART + *qual* (cf. Solà, 2002: 2553s.). If the relative clause contains highly lexicalized verb + *hi* combinations like *haver-hi*, as in (13–14), one would have to classify these cases as reduced relative clauses (cf. 3.4) and not as decumulative ones:

(13) A: perquè: és la_ (. 0.27) una de les obres (.. 0.46) e:_
          ex(.0.16)pressives_ (. 0.22) del (.. 0.32) quincipient gòtic\
      B: (.. 0.74) Jordi\ (ai)xò ho hem d'haver vist\ (.. 0.95) és en el
          claustre$_i$ aquell *que hi$_i$* han tants_ (COC)
(14) p(e)rò hi han llocs$_i$ *que* sí que *hi$_i$* ha padrí i padrina no\ (COC)[7]

The following example would be a better candidate for a decumulative locative relative construction, as *anar-hi* seems far less lexicalized than *haver-hi* (and also less lexicalized than the aspectually opposed *anar-se'n*), but here the absence of the auxiliary verb in the relative clause complicates the matter:

(15) p(e)rò que sigui una: una cosa: diferent a: anar a sopar i dallò_ perquè
      jo a l'única despedida de soltera *que hi* (a)nat_ (.. 0.63) que ha sigut la
      de l'Angelita_ (.. 0.52) va ser de pena\ (COC)

Apart from the (rather unappealing) hypothesis that this example is due to a transcription error, the form *hi* might be a kind of 'conflation' of the

---

7   Note the agreement pattern of 'impersonal' *hi ha* with the subsequent noun and pronominal adjective, respectively, in these two examples.

auxiliary *he* (< *haver*) and the pronominal adverb *hi*. This does not really affect the status of *hi* as a resumptive pronoun but makes this example all the more intriguing.

The main problem with decumulative locative relative clauses is in fact the evaluation of the degree of lexicalization attained by verbs frequently accompanied by *hi*. Whereas *haver-hi* will be considered by any grammar or dictionary of contemporary Catalan as lexically fixed and *anar-hi*, as in (15), as probably not, the status of *guanyar-hi*, as in (16), is difficult to determine:

(16) vaig entendre_ (... 1.25) que sobre el preu de venta_ (.. 0.45) hi
      guanyava el trenta per cent\ (... 1.55) i no hi guanya un trenta per cent
      (??) sobre-- per lo que tu expliques\ és sobre el preu de compra *que* hi
      guanya un trenta per cent\ (COC)

Guanyar-hi 'to make a profit, to earn (money)' is not among the verbs listed by Solà (2002: 2513) which "porten més o menys gramaticalitzats els pronoms *en* o *hi*". However, the presence of 'redundant' *hi* also in the segments of (16) that are not relativized speaks in favor of lexicalization. The fact that (16) is a cleft construction (cf. Solà, 2002: 2540ss. on clefts in Catalan), where the relativized segment is introduced by mere *que* and not by adverbial or complex relative pronouns, advocates for the same analysis as a non-decumulative construction.

An unambiguous case of decumulation with a verb incorporating (partly or fully) lexicalized *hi* would be a relative clause where *hi* appears twice. Unfortunately, no example of this type is found in the corpus data. The only occurrence of double *hi* has been detected in a clearly non-locative relative clause, so that this pronominal doubling must be considered a redundancy phenomenon independent from decumulative relative clause formation:

(17) A: p(e)rò vull dir que:_ (.. 0.38) aquí:\
      B: home sí sí\
      A: amb els pocs$_i$ *que*$_i$ *hi* devien haver-*hi*_
      B: molt molts castells-- molts castells de per (a)quí:_ (COC)

Rather surprisingly, examples of decumulative constructions involving possessive pronouns in cases where the gap position in the subordinate clause corresponds to the possessor, such as (18–19), have not been detected in the corpus data of COR or COC either.

(18) són amics$_i$ de tota la vida *que* coneixem massa bé les *seves*$_i$ debilitats
                                                   (Wheeler / Yates / Dols, 1999: 538)
(19) el noi$_i$ *que* el *seu*$_i$ pare és metge [...] (oral; overheard)

3.3 Pleonastic relative clauses

As only (unstressed) *que* has been taken into account in this study, oppor-
tunities for pleonastic relative clauses to appear are limited. As mentioned
above, a pleonastic relative clause is characterized by the fact that the syn-
tactic function of the relative gap is expressed twice, through an inflected
relative pronoun and a clitic pronoun in the subordinate clause. It is a very
controversial issue, though, whether *que* can be considered an inflected
relativizer at all. As has been pointed out already, many grammarians claim
it to be a general, non function-specific relative particle.[8] If one attributes
gap-filling functions to *que* in the sense that "it can function as either sub-
ject or object inside the relative clause, which itself may be restrictive or
non-restrictive" (Wheeler / Yates / Dols, 1999: 535), it remains question-
able if *que* then is to be considered as case-marked, which would imply to
postulate the existence of two homophonous relativizers *que*, one inflected
relative pronoun and one non-inflected particle, or if *que* actually does not
identify the subject or direct object (DO) position in these cases. Only on
the basis of the first assumption may pleonastic relative clauses with *que* be
conceivable. The controversial status of 'inflected' *que* will not be further
discussed in this study.

   If we accept, for the time being, that *que* is able to identify subject and
DO gap positions (in the sense of *Lehrstellenbildung* according to Lehmann),
pleonastic pronouns are to be much more expected in the case of an object
gap than in that of a subject gap, as Catalan, as a pro-drop language, usu-
ally makes use of subject pronouns in marked contexts of emphasis and
contrast only – and then, the pronominal elements used are stressed pro-
nouns but not clitics. This is borne out by the data: in all the occurrences
of pleonastic relative clauses (some 10 examples in the working corpus) the
relative clause contains *que* + a DO clitic. (20–21) contain paradigmatic
cases:

---

8   As Solà does, when he describes *que* as "[r]elatiu àton proclític sense significat ni flexió:
     els seus trets lèxics són els de nexe i anàfora" (Solà, 2002: 2552).

(20) s:i tinc temps-- xx bueno\ entre demà i demà passat_ lo que volia fer_
     era a baix al traster$_i$ aquell que vam-- *que*$_i$ vam mig ocupar-*lo*$_i$_ (.. 0.44)
     és posar una mica amb ordre_ bueno\ (COC)

(21) i:_ el altre_ són els:_ passports$_i$ amb Micom_ *que*$_i$ ara *els*$_i$ han ficat dins
     de Bay\ (COR)

Note the variable DO clitic position (proclitic in (21), enclitic in (20)). In
(22), *ho*, a clitic pronoun that normally "represents the direct object 'it'
when the direct object complement cannot be identified as a specific
noun" (Wheeler / Yates / Dols, 1999: 184), is used; but this example also
permits a reading in which the clitic refers to the specific referent the wolf
("el llop"), in which case *ho* would be used abusively and pleonastically:

(22) ai no\ la mare no pot ser\ (.. 0.51) que la mare no té aquesta veu\ (...
     1.64) deu ser el llop$_{?i}$\ *que* $_{?i}$ ens *ho* $_{?i}$ va dir la mare\ no obrirem eh\
     no obrirem\ que la mare ens va dir que no obríssim\ (COR)

What makes this example unclear is, again, the polyfunctionality of *que*, a
generalized phenomenon in Romance: *que* may be analyzed, in the
sequence under observation, as a relative pronoun / relative particle but
also – along with the *que*s in the preceding and subsequent text segments –
as a non-relative conjunctive *que* in a causal reading.

   (23) contains an interesting example of number-feature incongruence
between the head noun and the 'redundant' DO clitic – the relativizer *que*
being obviously unable to carry number features –, but this non-agreement
concerns the attributive function of the relative clause and does not affect
the gap-filling function of the clitic *els*:

(23) que llavòrens_ a sobre hi posaves_ un:_ (.. 0.31) un mòdul
     d'expansió$_i$_ (.. 0.44) *que*$_i$ *els*$_i$ enllaçaves no sé per on_ (.. 0.71) et sona_
     (. 0.19) no serà un expansion module_ no\ (COR)

3.4 Reduced relative clauses

The most frequent form of 'deviant' relative clause formation in our
working corpus, comprising almost 50% of the examples, is the "defec-
tive" or inflectionally reduced relative clause, where out of the three basic
operations of relative-clause formation, only subordination is morphologi-
cally marked through the relative particle *que* but where neither the syntac-

tic gap is identified nor the attribution of the clause to the head-noun / nucleus is indicated. This comparatively high frequency of reduced relative clauses is in accordance with Solà's estimation (based on Cid Abasolo, 1999) that this type of relativization is a maximally economic strategy ("una estratègia *màximament econòmica*" [Solà, 2002: 2532; highlighted as in orig.]) appropriate for the oral language system. Provided that we accept *que* to syntactically identify the subject and DO gap, as outlined in the previous paragraph, reduced relative clauses with *que* necessarily involve cases where the head noun occupies an oblique position in the relative clause, i.e. where the gap is governed by a preposition.[9] Normatively speaking, such cases require complex relative pronouns with PREP + stressed pronouns *qui* or *què*, as in (24–25), or PREP + ART + *qual*, as in (26):

(24) tu que vius a recer de l'Altíssim_ (.. 0.47) i passes la nit a l'ombra del Totpoderós_ (.. 0.68) digues al Senyor_ (.. 0.63) sóc la muralla on m'emparo\ (.. 0.47) el meu Déu *en qui* confio\ (COR)

(25) per tant_ ((writes on the blackboard)) quin és_ e:_ quins són els termes_ *en què* es produeix aquesta crisi_ concretament_ analitzarem_ intentarem analitzar les causes\ (COR)

(26) és un llibre *en el qual* m'he basat_ (.. 0.77) força_ per_ per_ preparar aquest tema\ (COR)

Oblique positions figure low in the accessibility hierarchy for relativization both in terms of Keenan / Comrie's (1977)[10] and Lehmann's (1984) approach, and oblique positions governed by nouns are even less accessible for relative clauses than oblique positions governed by verbs. It comes as no surprise, then, that oblique relative clauses are scarce in oral corpora and occurrences restricted to formal speech, as the examples above from religious (24) and academic contexts (25–26) illustrate, and that norm-compliant complex relativizers are extremely infrequent.

In our working corpus, no clear-cut example of a reduced relative clause with an IO in the gap position could be identified. As an inflectional dative case survives in Catalan (as in most Romance languages) in the pro-

---

9   The fact that Catalan, as do Castilian and several other Romance languages, allows or calls for the encoding of highly individuated DOs by means of a preposition – the so-called differential object marking or 'prepositional accusative' – is not taken into account here.

10 Keenan / Comrie (1977: 66 and *passim*) define obliqueness in a more restrictive way, not including IO and genitive (possessor) positions under this heading.

nominal system but not in the nominal one, speakers seem to prefer in these situations the decumulative construction, as described in 3.2. The oblique gap positions expressed through reduced relative clauses in our data include prepositional objects of different governing verbs: we find, for instance, occurrences of *pensar* (*en*) + NP 'think of so. or sth., remember so. or sth.' ((27–28); in order to show clearly the oblique relation the reduced relative clauses are repeated in the b versions with oblique relativizers, but these forms are often questionable or felt as ungrammatical or pragmatically inadequate by native speakers, which is indicated by the question mark(s) or the asterisk):

(27) a. mira\ amb anglès_ hi ha una cosa *que* has de: *pensar*\ (.... 8.35) amb
           anglès sempre_ sempre has de tindre en compte una cosa\ (COC)
      b. hi ha una cosa *en què* / *en la qual* has de pensar
(28) a. la mare pot dir_ (.. 0.45) al marit_ escolta'm\ (... 1.08) és la dona
           que tu pensaves\ (COR)
      b. és la dona ˀ*en qui* / ˀ*en la qual* tu pensaves

There are two reduced relative clauses in the data which are prepositional objects governed by *parlar* (*de*) 'talk about':

(29) a. lo altre_ e::_ òbviament és exactament lo mateix_ *que* hem *parlat*_
           am e:_ am l'altre:_ projecte\ (COR)
(30) a. necessiten_ aquest- -- aquests serveis\ (. 0.18) un és el *que* hem
           parlat abans del comerç\ (.. 0.58) necessiten un comerç_ que_
           estigui adequat_ (.. 0.38) a::_ a les seves necessitats_ (COR)

The reconstructed non-reduced relative constructions corresponding to these example, contrary to (27–28), are dubious or even straightforwardly unacceptable, probably due to the fact that the anteceding nucleus is an unstressed pronoun (30) or a neuter pronominal expression (29) (cf. Solà, 2002: 2484s.):

(29) b. és exactament lo / el mateix ˀ*de què* / ˀ*del qual* hem parlat
(30) b. *un és el *de què* / *del qual* hem parlat abans

Therefore, the 'reduced' versions (29a–30a) seem to be the only viable ones in this type of construction.

More reduced relative clauses with an oblique gap governed by verbs include comitative *quedar* (*amb*) '(to) arrange to meet so.' and *connectar* (*amb*) 'establish a relationship with so.' (32–33), and instrumental *comprar* (*per / a*) / *vendre* (*per / a*) 'buy / sell at (a certain price), buy / sell for (a certain amount of money)' (31):

(31) a. A: de la diferència entre el preu *que compra* i el preu *que ven_*
          B: no\ hi ha un—
          A: hi ha un trenta per cent de benefici\
          B: no\ (COC)
      b. de la diferència entre el preu *per què / a què / pel qual / al qual compra* i el preu *per què / pel qual ven* hi ha un trenta per cent de benefici

(32) a. allò de què_ hi ha persona *que connectes* i persona *que* no\ que això passa sempre\ (COC)
      b. hi ha persona *amb qui / amb la qual connectes* i persona *amb qui / amb la qual* no

(33) a. vaja\ feu la vostra_ nosaltre:s_ contactarem amb la gent *que* havíem *quedat_* (.. 0.33) una gent d'aquí Esparreguera\ (COR)
      b. nosaltres contactarem amb la gent *amb qui / amb la qual* havíem *quedat*

Oblique gaps may also be the result of a nucleus that occupies, in the relative clause, a position governed by a preposition which depends on a noun or a NP; according to Lehmann (1984: 213), these adnominal prepositional phrases are, again, organized in an accessibility hierarchy, with possessor phrases being more accessible to relativization than standards of comparison and these being more easily relativized than other adnominal prepositional phrases; all these functions, however, rank lowest on the overall accessibility scale for relative clause formation. Not surprisingly, no example of this kind was found among the reduced relative clauses of our data.

Wheeler / Yates / Dols (1999: 541) mention that the combination of *el / la / els / les* + *que* "is frequently heard after a preposition (instead of stressed *què/qui* or compound *el qual*, etc.), but is condemned as non-standard, in particular, as a Castilianism." In our working corpus, there was only one occurrence of this oblique relativizing structure in a reduced relative clause (with an unexpected preposition, too):

(34) a. la:_ dimensió_ e:_ ideològica_ (. 0.25) (??) evidentment\ *de la que*
          m'he referit abans_ (.. 0.34) i la dimensió econòmica_ (.. 0.65) de la
          premsa_ (COR)
      b. la dimensió ideològica *a què / a la qual* m'he referit abans

This construction has to be distinguished from a situation where a definite
article is adjacent to the relativizer *que* due to ellipsis or specific word-order
phenomena (Badia i Margarit, 1994: 364; Wheeler / Yates / Dols, 1999:
544s.). (36) is a candidate for this category, where the recovery of the ellip-
sis calls for a quite distant element in the preceding context. The reduced
gap position, in this case, is a locative prepositional phrase depending on
the verb *anar* (*a*) 'go (to)':[11]

(35) a. (a)nàvem a sopar_ a una pizzeria_ no havien ni reservat taula ni re
          sats/ (.. 0.38) bah\ (.. 0.53) (a)nàvem catorze o quinze_ un
          dissabte a la nit_ (.. 0.41) a sopar\ (.. 0.98) i més a més vem marxar
          tard\ (.. 0.79) am(b) el am(b) una guagua d'aquestes_ (a)nàvem
          am(b) això i lo que—l'únic que va estar bé\ (.. 0.42) au\ (.. 0.82) a
          a *la que* anàvem no hi (ha)via taula_ (. 0.15) després no sabíem on
          (a)nar\ (COC)
      b. a la (pizzeria) *on / a què / a la qual* anàvem no hi havia taula

As inferable from the non-reduced variant (35b), the locative relative
clauses differ from the oblique relative clauses mentioned hitherto in that,
in addition to the complex relativizers PREP + stressed *què* and PREP +
ART + *qual*, they allow to be introduced by a compact relative pronoun *on*,
homonymous with the question-word *on* 'where?'. Locative relative clauses
share this characteristic with certain relative clauses where the nucleus
denotes time (introduced by *quan*; see below 3.5) or manner (introduced,
albeit seldom, by *com*), characteristics according to which they constitute a

---

11 There is yet another construction that has to be singled out in this respect, and this is
    the *el que* used as a neuter relativizer in headless relative clauses. Neuter *el que* is fre-
    quently replaced, in spoken Catalan, by castilianizing *lo que* (and therefore criticized by
    normative grammarians), maybe in order to keep the neuter relativizer distinct from the
    el (ART) + *que* in the context of the elliptic structures under scrutiny here (Badia i Mar-
    garit, 1994: 362). Our corpus provided us with a particularly intriguing example of this
    category in the form of a reduced relative clause, (i), which, for reasons of space, will
    not be further discussed here:

    (i) potser de totes les feines que he fet_ és *lo que:_* (.. 0.35) veies més el resultat (COR)

specific sub-group of 'adverbial' relative clauses (cf. Lehmann, 1984: 318ss.; Wheeler / Yates / Dols, 1999: 547s.; Solà, 2002: 2553ss.); like their English equivalents, these adverbial relative pronouns "do not need to occur with a preposition, since they substitute for an entire adverbial (while the other relativizers substitute only for a noun phrase)" (Biber *et al.*, 1999: 624). What is remarkable in the context of the present chapter is that reduced relative clause formation is particularly frequent with gap positions corresponding to adverbial expressions of place. A dozen examples have been detected in our data; (36–38) give some illustration for this phenomenon:

(36) a. tots crèiem_ que era un barri_ *que* més aviat la gent gran_ és la que predominava_ (COR)

   b. era un barri *on / en què / en el qual* més aviat la gent gran és la que predominava

(37) a. aquell lloc *que* vem anar e:l dia de:_ les de Salio:ns_ o:_ (COC)

   b. aquell lloc *on / a què / al qual* vam anar

(38) a. el seu propi programa electoral de_ eliminació del doble finançament\ (.. 0.70) no hi ha (??) enlloc *que* s'ha eliminat el doble finançament_ (. 0.23) de la televisió pública\ (.. 0.48) a: Espanya\ (COR)

   b. no hi ha enlloc [??]*on* / *a què* / **al qual* s'ha eliminat el doble finançament

The presence of the head-noun as a locative adverbial phrase in the matrix clause, containing the same preposition that would introduce the PREP + *què* or PREP + ART + *qual* relativizer in the subordinate clause, seems to encourage the formation of a reduced relative clause, avoiding thereby the repetition of the preposition (but this could also be achieved, without resorting to a reduced relative clause, through the use of *on*):

(39) a. A: [jo] no hi he estat a un casament *que* aplaudeixin\
   B: jo tampoc\ (COC)

   b. jo no hi he estat a / en un casament *a què / en què / al qual / en el qual / on* aplaudeixin

(40) a. A: això són les pàgines web\
   B: però_ jo mai he arribat a cap lloc *que* volia anar\ (COC)

   b. jo mai he arribat a cap lloc *a què / al qual / on* volia anar

In the case of toponyms, the PREP + *què* and PREP + ART + *qual* strategy is, generally speaking, not available (Solà, 2002: 2553), but a non-reduced relative clause introduced by *on* would still be possible:

(41) a. m'agrada més fer-ho a Les Preses *que* tinc el forn aquell que--\

                                                                        (COC)

      b. m'agrada més fer-ho a Les Preses, *on / *a què / *a les quals* tinc el forn aquell

In any case, relative clauses where the gap position is identified as a locative expression seem to be particularly susceptible to inflectionally reduced relativization.

Another correlation that the quantitatively reduced working corpus allows for is that between reduced relative clauses and multi-level relativization (*coniunctio relativa*; cf. Solà, 1972: 134ss.). By this term we mean a case in which a relative construction contains, apart from the matrix with the nucleus, a relative clause which itself functions as the matrix for a further subordinate clause, with the nucleus assuming some grammatical function in both subordinate clauses. Structures like these range from a minimum of three syntactic levels (M + RC + XC, such as (12) above or (42)) to more complex constructions, like (43):[12]

(42) aquest és el taxista amb qui la Carme creu que anirem

                                                        (Solà, 2002: 2528)

(43) hem vist l'artista del qual hem sentit que la Carme deia que en Pau es burlava (*op.cit.*: 2528s.)

---

12 This kind of complex relativization pattern has to be distinguished from the so-called 'stacking' of relative clauses (cf. Lehmann, 1984: 197ss.), which is a recursive application of relativization with the same attributive features, leading to constructions such as (ii). These 'stacked' relative clauses may – with contextual, i.e. pragmatic restrictions – be altered in their linearization, at least to a certain extent (cf. (iib)), something that would lead to ungrammatical results in the case of multi-level relativization as analyzed here:

   (ii) a. the only artists I have ever known who are personally delightful are bad artists

                                                  (O. Wilde, *apud* Lehmann, 1984: 198)

         b. the only artists who are personally delightful whom I have ever known are bad artists (*op.cit.*: 199)

Such constructions are problematic in more than one respect. This becomes obvious when one tries to explain the syntactic structure of an example like (42) through bracketing. (42a) is inadequate, as the oblique relativizer obviously pertains to the segment *anirem* and not to *la Carme creu*, which is proven by the possibility to eliminate the latter segment but not the former one (42b–c):

(42) a.   *[_Maquest és el taxista [_{RC}amb qui la Carme creu [_{XC}que anirem]]]
     b.   aquest és el taxista amb qui anirem
     c.   *aquest és el taxista amb qui la Carme creu

However, bracketing as in (42d) is not satisfying either, as the segment *la Carme creu que* is no reasonable clausal unit and, furthermore, the relation of relativization that certainly holds (although maybe implicitly) between this segment and the nucleus is totally blurred:

(42) d.   *[_Maquest és el taxista [_{RC}amb qui [_{XC}la Carme creu que] anirem]]

Another problematic point, related to that of explicating the syntactic structure of such constructions through bracketing, concerns the attribution of a syntactic status to the second subordinate clause. Whereas in (12) it seems manifest that the clause introduced by *sempre que* is an adverbial clause, in examples with simple *que* such as (42) the answer is less easy to provide. If one tries to translate this sentence into German, where the relativizer and the complementizer corresponding to *que* may be distinguished phonologically (and orthographically), the second clause would appear to be a complement clause (CC):

(42) e.   das ist der Taxifahrer mit dem Carme glaubt dass wir fahren[13]

It is not within the scope of the present article to further discuss the problematic status of these multi-level relative constructions, which certainly deserve closer examination (but cf. Solà, 2002: 2527ss. for a more

---

13 This translation will sound hardly acceptable to most speakers of German, who would prefer to translate (42) in a way that makes manifest the relation of relativization between the nucleus and the *la Carme creu* segment (and which is absolutely possible in Catalan, too):

    (iii) [_Mdas ist der Taxifahrer [_{RC}von dem Carme glaubt [_{CC}dass wir mit ihm fahren]]]

detailed description and analysis). What seems significant, however, is to emphasize that the corpus data from COC and COR suggest that these multi-level relative constructions favor the use of the reductive relativization strategy. Eight examples of this kind have been detected in the corpora. Their internal structure is more transparent than that of examples like (42–43), and most of them follow a 'M + RC + CC' scheme. (44–45) are paradigmatic cases:

(44) a. jo en allà hi tinc un cel·lo *que* no sé qui l'hi ha portat\ (COC)
     b. hi tinc un cel·lo **de què* / ??*del qual* no sé qui l'hi ha portat
(45) a. referent a:_ les associacions_ *que* dius que n'hi han moltes_ jo crec que hi han menos\ (COR)
     b. les associacions ??*de què* / ?*de les quals* dius que n'hi ha moltes

The following example (46) is similar to (45); however, the presence of the relative particle *que* seems to have induced the elimination of the complementizer *que* of the second subordinate clause:

(46) a. A: no sé\ es monten coses a vegades *que:_*
         B: (. 0.13) mhm\
         A: (. 0.22) *que* dius no va bé\ no va bé\ (COR)
     b. es monten coses ??*de què* / ?*de les quals* dius *que* no va(n) bé

As becomes obvious from these examples, the reconstructed 'norm-compliant' b variants, with a complex subordinator making explicit the oblique gap position in the relative clause, are not always as felicitous as in (47), where the oblique gap is a locative expression:

(47) a. ((talking about castles in France)) han procurat conservà'ls e una mica\ n'hi han molts *que* ja es veu que està molt tros afegit\ (COC)
     b. n'hi ha molts *on* ja es veu que està molt tros afegit

There are even instances of multi-level relativization where such a reconstruction leads to results that are, if not entirely impossible, at least very cumbersome, as in (48):

(48) a. que he tornat de:_ Menorca_ (. 0.11) i al: la casa aquella *que::_* vem quedar que et preguntaria_ la tenen llogada\ la setmana última de:_ de juny\ (COR)

    b. ??la casa aquella *en relació amb què / la qual / referent a què / a la qual*
       vam quedar que et preguntaria...

As the reconstructed version (48b) makes clear, the relativization in cases like (48a) seems to indicate no more than the fact that the content of the relative clause (and the complement clause that depends on it) has something to do with the head-noun, but no convincing grammatical gap position exists for the head-noun in the relative clause. These examples are similar to those (written) examples cited by Solà,

(49) li han regalat una pintura *que* quan entrarà al despatx quedarà
     bocabadat (Solà, 2002: 2530)
(50) la Verònica [...] tenia un carnet que si el presentaves el dimarts i el
     divendres podies entrar al Goya pagant només mitja entrada
                                        (T. Moix, *apud ibid.*)

which he considers most uncomfortable for traditional grammar ("un dels casos més coneguts de construcció incòmoda per a les gramàtiques lògiques" [*op.cit.*: 2529]), because there is no gap position at all that the head noun could occupy, which therefore has no function within the relative clause that is attributed to it ("el relatiu *no fa cap funció* dins la seva pròpia oració" [*ibid.*; highlighted as in orig.]).[14] The hypothesis that in these cases of multi-level relativization, the relativizer does not simply not mark the syntactic gap but that there is, structurally speaking, no such gap, is supported by the following example, which would not be classified among the reduced relative clauses, because the relativizer seems to identify the position of the nucleus as DO (*saber* [*alguna cosa*] 'know sth.'):

(51) l'artista\ (. 0.17) en aquell moment diu_ (. 0.16) vaig deixar de ser_
     (. 0.15) racionalista_ (. 0.14) per ser alguna cosa *que* encara no sabia
     què seria\ (. 0.20) p(e)rò en aquell moment_ ell deixa de ser
     racionalista_ (COR)

However, the DO position is actually occupied by the second subordinate clause (here: a headless relative clause functioning as a noun; cf. Lehmann, 1984: 316ss.). Thus, two explicative solutions remain: either that *sabia* 'knew'

---

14 And Solà goes on to emphasize that these constructions are limited to the oral system and that the normative language seems unable to produce equivalent constructions: "les llengües estàndard no hi troben solució" (Solà, 2002: 2529, note 53).

governs two direct objects, a rather unattractive hypothesis with the verb *saber*;[15] or that the head-noun of the relative clause introduced by *que* has no precise grammatical function in this clause.

3.5 Relative clauses with a temporal gap position

Relative constructions where the gap position in the subordinate clause corresponds to an adverbial expression of time follow some specific behavior in most Romance languages that clearly diverges from Germanic languages such as German or English, where an adverbial relative pronoun unambiguously identifying the temporal gap or a complex relative pronoun with a preposition can be used in most cases, as in the (invented) examples (52a–b):

(52) a. the day *when* Hugo and Marta marry
      b. der Tag *an dem* Hugo und Marta heiraten

In our corpus, this relativization surfaces as in (52c), which, at first sight, looks like a reduced relative clause:

(52) c. avui_ [...] és el dia *que* es casen l'Hugo i la Marta\ (.. 0.37) i si passa
          alguna cosa més al món_ a nosaltres no ens interessa\ (COR)

As Solà rightly emphasizes, temporal relative clauses are on the borderline of relativization, sharing many formal and functional characteristics with adverbial clauses (and interrogative sentences) ("aquestes formes assenyalen un límit entre les relatives, les interrogatives i les circumstancials" [Solà, 2002: 2475]), and (therefore?) the relativization patterns involved show a high degree of idiosyncrasy (*ibid.*). Romance languages in general make a much more restrictive use of the adverbial relativizer homophonous with the interrogative pronoun of time and give preference to the use of PREP + (stressed) relative pronoun constructions or of the bare relative particle *que* (or equivalent forms) even in the standard language, as the Italian (53a) or Castilian examples (53c–d) from learners' grammars show:

---

15 This hypothesis is more convincing with the verb *creure* 'believe (so. to be sth. or so.)' as in (iv), which otherwise is very similar to (51):

   (iv) i finalment_ am(b) el Rafael Pérez de Estrada_ *que* creiem que és un dels-- dels
        poetes_ andalusos_ més importants_ qui hi ha:_ actualment (COR)

(53) a. era la sera *che* (*in cui / nella qualè*) l'ho conosciuto
                                                    (Kirsten / Mack, 1976: 109)
     b. una noche *en que* iba a buscarla
                                    (F. Umbral, *apud* Butt / Benjamin, 2000: 498)
     c. el único día *que* se produjeron diferencias de importancia fue el
         jueves (*apud ibid.*)

French behaves slightly differently, in that the norm discards PREP + stressed relative pronouns but accepts, alongside bare *que*, adverbial locative *où* 'where' in temporal contexts:

(54) a. la première fois *que* je l'ai vu [...] (Riegel / Pellat / Rioul, 2001, 483)
     b. l'époque *où* j'allais à l'école [...] (*ibid.*)

Standard Catalan behaves similar to Spanish, in that it allows adverbial *quan* (analogous to Castilian *cuando*) in relative clauses, but only if they are non-restrictive (appositive); in restrictive relative clauses, only PREP + *què*, PREP + ART + *qual* and unstressed *que* are tolerated (Solà, 2002: 2475), but due to the numerous idiosyncrasies noted by Solà (2002) and alluded to above, there is not equal choice in all contexts:

(55) a. [...] també dels 90, època *en què* ha predominat [...] la funció lúdica
                                                    (*apud* Solà, 2002: 2476)
     b. tant a les dècades de 1640 i 1650, *quan* els comtats de Roselló i
         Cerdanya estaven ocupats (*apud ibid.*)
     c. ara, *que* / **en què* / *quan* tothom dorm, podem entrar a Internet
                                                                    (*ibid.*)

In this perspective, the examples of temporal relative clauses found in the corpora are unspectacular: *que* is used in all instances, with all the occurrences corresponding to restrictive relative clauses. (55–56) give some illustration of these relative clauses as found in the data:

(55) perquè aquí també hi havien èpoques_ *que* hi havia una mica de
     tensió_ p(e)rò al final la tensió se'n (a)nava rient\ (COR)
(56) A: a les set em despe:rto_ p(e)rò em desperto—
     B: jo també em desperto a l'hora *que* ell se'n va\ p(e)rò:_
     A: jo em desperto_ per exemple_ què sé jo\
     B: p(e)rò m'adormo a dos quarts de nou o així\ (COC)

4 Conclusion

The present study has focussed on some aspects of relative clause forma-
tion in spoken Catalan as documented in recent corpora. Although the
limited size of the working corpus did not allow for a far-reaching quanti-
fied analysis, the data showed that the three types of 'deviant' relative
clause formation strategies described, from Guiraud (1966) on, as typical
for spoken Romance – decumulative, reduced and pleonastic relative
clauses – are also found in oral Catalan, although maybe less frequently
than one would first estimate impressionistically. A characteristic feature of
these alternative relativization strategies is the replacement of inflectionally
elaborate relative pronouns (baring case and agreement features) by the
morphologically uniform relative particle *que*, which, apart from the cases
where the gap position in the relative clause is subject or DO, is unable to
perform the operations of identification of the syntactic gap and that of
attributing the relative clause to the head noun / nucleus, but is a simple
subordinator (i.e. a complementizer). Reduced relative clauses, where only
this inflectionally poor pronominal element *que* subsists, constitute the
quantitatively most important sub-group of 'deviant' relative clauses in our
corpus. Decumulative relative clauses, where the identification of the
grammatical gap and the attribution of the clause to the nucleus is carried
out by a separate resumptive pronoun, follow next in terms of frequency
but seem to be far less prominent in spoken Catalan than reduced relative
clauses.

   This study obviously does not give a full account of relative-clause
formation in spoken Catalan; such a – certainly desirable – undertaking will
be left for future work on the subject.

Bibliography

Alexiadou, Artemis *et al.* (2000): "Introduction"; in: Alexiadou, Artemis e*t
     al.* (eds.): *The syntax of relative clauses*, Amsterdam / Philadelphia: Benja-
     mins (Linguistik Aktuell / Linguistics Today; 32), 1–51.
Alturo, Núria / Boix, Emili / Perea, Ma. Pilar (2002): "*Corpus del Català
     Contemporani de la Universitat de Barcelona* (CUB). A general presentation",
     in: Pusch, Claus D. / Raible, Wolfgang (eds.): *Romanistische Korpuslinguis-
     tik – Korpora und gesprochene Sprache / Romance Corpus Linguistics – Corpora
     and Spoken Language*, Tübingen: Narr (ScriptOralia; 126), 155–170.

— *et al.* (eds. 2004): *Corpus oral de registres. Materials de treball*, Barcelona: Publicacions de la Universitat de Barcelona (Universitat; 13) [book and CD-ROM].

Badia Margarit, Antonio M. (1947): *Los complementos pronominalo-adverbiales derivados de* IBI *e* INDE *en la península ibérica*, Madrid: CSIC (Revista de Filología Española; Anejo 88).

Badia i Margarit, Antoni M. (1994): *Gramàtica de la llengua catalana. Descriptiva, normativa, diatòpica, diastràtica*, Barcelona: Proa (Biblioteca universitària; 22).

Barlow, Michael (2000): *MonoConc Pro 2.0*, Houston, TX: Athelstan [concordance program].

Bianchi, Valentina (2002): "Headed relative clauses in generative syntax", *Glot International* 6:7, 197–204 and 6:8, 235–247.

Biber, Douglas *et al.* (1999): *Longman grammar of spoken and written English*, Harlow: Pearson Eduction.

Butt, John / Benjamin, Carmen (³2000): *A new reference grammar of Modern Spanish*, London: Arnold.

Brucart, José Ma. (1999): "La estructura del sintagma nominal: Las oraciones de relativo"; in: Bosque, Ignacio / Demonte, Violeta (eds.): *Gramática descriptiva de la lengua española. Vol. 1: Sintaxis básica de las clases de palabras*, Madrid: Espasa, 395–522.

Cid Abasolo, Carlos (1999): "Cláusulas relativas con conjunción en las lenguas románicas, eslavas y vasca", *Revista de Filología Románica* 16, 267–288.

Cinque, Guglielmo (2001): "La frase relativa"; in: Renzi, Lorenzo / Salvi, Giampaolo / Cardinaletti, Anna (eds.): *Grande grammatica italiana di consultazione. Vol. 1: La frase. I sintagmi nominale e preposizionale*, Bologna: Il Mulino, 457–517.

Comrie, Bernard / Kuteva, Tania (2005): "Relativization strategies", in Haspelmath *et al.* (eds.), 494–497.

Cresti, Emanuela / Moneglia, Massimo (eds. 2005): *C-ORAL-ROM. Integrated reference corpora for spoken Romance languages*, Amsterdam / Philadelphia: Benjamins (Studies in Corpus Linguistics; 15) [book and CD-ROM].

Fiorentino, Giuliana (1998): "La clausola relativa nel parlato delle lingue romanze", in: Ruffino, Giovanni (ed.): *Atti del XXI Congresso Internazio-*

nale di Linguistica e Filologia Romanza, Palermo 1995. Sezione 2: Morfologia e sintassi delle lingue romanze, Tübingen: Niemeyer, 297–308.

— (1999): "Clausole relative romanze tra innovazione e conservazione", *Revue Romane* 34, 25–60.

Fuchs, Catherine (ed. 1987): *Les types de relatives*, Paris: Larousse (*Langages*; 88).

Gadet, Françoise (1988): *Le français ordinaire*, Paris: Armand Colin.

— (1995): "Les relatives non standard en français parlé : le système et l'usage", in: Andersen, Hanne Leth / Skytte, Gunnar (eds.): *La subordination dans les langues romanes. Actes du colloque international Copenhague 1994*, Copenhague: Université de Copenhague / Institut d'Etudes Romanes (*Etudes Romanes*; 34), 141–162.

— (2003): "La relative française, difficile et complexe", in: Kriegel, Sibylle (ed.): *Grammaticalisation et réanalyse. Approches de la variation créole et française*, Paris: CNRS Editions, 251–268.

Gapany, Joël (2004): *Formes et fonctions des relatives en français. Etude syntaxique et sémantique*, Bern *et al.*: Peter Lang (Sciences pour la communication; 73).

Godard, Danièle (1988): *La syntaxe des relatives en français*, Paris: Editions du CNRS.

Guiraud, Pierre (1966): "Le système du relatif en français populaire", *Langages* 3, 40–48.

Haspelmath, Martin *et al.* (eds. 2005): *The World Atlas of Language Structures*, Oxford *et al.*: Oxford University Press.

Keenan, Edward L. / Comrie, Bernard (1977): "Noun Phrase Accessibility and Universal Grammar", *Linguistic Inquiry* 8:1, 63–99.

Kirsten, Gerhard / Mack, Barbara (1976): *Grammatica italiana per tutti*, Stuttgart: Klett.

Kleiber, Georges (1987): *Relatives restrictives et relatives appositives : une opposition "introuvable" ?*, Tübingen: Niemeyer (Beihefte zur ZRPh ; 216).

Kunstmann, Pierre (1990): *Le relatif-interrogatif en ancien français*, Genève: Droz.

Lehmann, Christian (1984): *Der Relativsatz. Typologie seiner Strukturen, Theorie seiner Funktionen, Kompendium seiner Grammatik*, Tübingen: Narr (Language Universals Series; 3).

— (1995): "Relativsätze"; in: Jacobs, Joachim *et al.* (eds.): *Syntax*, Berlin / New York: De Gruyter (Handbuch der Sprach- und Kommunikationswissenschaft; 9:2), 1199–1216.

Oller, Anna *et al.* (2000): "El COC del CUB: un corpus per a l'estudi de la conversa col·loquial", *Zeitschrift für Katalanistik* 13, 58–91.

Osuna García, Francisco (2005): *Las construcciones de relativo*, Córdoba: Universidad de Córdoba / Servicio de Publicaciones.

Payrató, Lluís / Alturo, Núria (eds. 2002): *Corpus oral de conversa col·loquial. Materials de treball*, Barcelona: Publicacions de la Universitat de Barcelona (Universitat; 11) [book and CD-ROM].

Riegel, Martin / Pellat, Jean-Christophe / Rioul, René (2001): *Grammaire méthodique du français*, Paris: Presses Universitaires de France.

Solà, Joan (1972): "El relatiu a l'actual català literari"; in: *id.*: *Estudis de sintaxi catalana 1*, Barcelona: Edicions 62 (Llibres a l'abast; 104), 103–138.

— (2002): "Les subordinades de relatiu"; in: *id. et al.* (eds.): *Gramàtica del català contemporani. Vol. 3*, Barcelona: Empúries, 2455–2565.

Schafroth, Elmar (1993): *Zur Entstehung und vergleichenden Typologie der Relativpronomina in den romanischen Sprachen. Mit besonderer Berücksichtigung des Substandards*, Tübingen: Niemeyer (Beihefte zur ZRPh; 246).

Thun, Harald (1986): *Personalpronomina für Sachen. Ein Beitrag zur romanischen Syntax und Textlinguistik*, Tübingen: Narr (Tübinger Beiträge zur Linguistik; 262).

Touratier, Christian (1980): *La relative. Essai de théorie syntaxique (à partir de faits latins, français, allemands, anglais, grecs, hébreux, etc.)*, Paris: Klincksieck.

Viaplana, Joaquim / Perea, Ma. Pilar (2003): *Textos orals dialectals del català sincronitzats. Una selecció*, Barcelona: PPU Promociones y Publicaciones Universitarias [book and CD-ROM].

Wheeler, Max W. / Yates, Alan / Dols, Nicolau (1999): *Catalan. A comprehensive grammar*, London / New York: Routledge.