The development of locative relative markers: from typology to sociolinguistics (and back)

Silvia Ballarè (University of Bologna) and Guglielmo Inglese (KU Leuven)

Abstract:

The *accessibility hierarchy* was first proposed by Keenan & Comrie (1977) to describe the cross-linguistic distribution of relative markers in terms of likelihood of relativization of different syntactic roles. The hierarchy is also commonly believed to reflect constraints on possible changes in the domain of relativization. For example, the hierarchy predicts that locative relatives that develop into general relativizers should expand their functional range in a step-by-step fashion from lower to higher roles. In this paper, we revise existing claims about the diachrony of locative relatives. In doing so, we survey known cases of locative relatives that develop into general relativizers, and we also discuss data from linguistic variation in non-standard varieties in European languages, with a focus on social variation in Italian. As we argue, data from Italian suggests that another possible cline of development of locative relatives should be acknowledged, that is, LOCATIVE > CONCERN > SUBJECT.

Keywords: diachronic typology, relative clause, locative relative, language variation, sociolinguistic variation, accessibility hierarchy

1. Introduction

Relative clauses have been a key topic in typological research. A substantial effort has been put into exploring the possible range of variation in relativization strategies and their distribution in the languages of the world. Among proposed typological generalizations on relative clauses, perhaps the most influential in typological studies has been Keenan and Comrie's *noun phrase accessibility hierarchy* (1977, 1979), which essentially concerns the existence of constraints on the availability of individual syntactic roles for relativization within and across languages.

Comparatively less attention has been paid to the diachronic typology of relative clauses. Work by Lehmann (1986), Heine & Kuteva (2002 [now Kuteva *et al.* 2019]), and Hendery (2012), Givón (2012), to name a few, has progressively filled this gap, so that we now have a fairly good understanding of the possible sources and processes whereby relative clause constructions come about and how they can further develop. Specifically, in connection with the accessibility hierarchy, it has been observed that relative markers that originally relativize only one syntactic role may extend to the encoding of other roles. When this is the case, a strong claim has been formulated that extensions to new syntactic roles must follow the accessibility hierarchy, with markers spreading either up- or downwards the scale without skipping any intermediate position (see Section 2.2).

Locative relatives constitute an interesting but, in several respects, still understudied example of such a historical process. These are relative markers meaning ‘where’ that are in origin restricted to the relativization of oblique locatives and subsequently start being employed for the relativization of other syntactic roles. In some languages, this development reaches the stage whereby locative relative markers can even relativize subjects. Such polyfunctionality of locative relatives has so far been documented in a number of languages, and it is commonly believed that locative relatives that expand their syntactic range do so in compliance with the accessibility hierarchy (e.g. Romaine 1984, Fleischer 2004).

In this paper, we discuss possible counterevidence to this generalization, and investigate selected known cases in which relative markers with a locative origin start encoding core argument positions but do not show signs of being used for the intermediate roles on the hierarchy. One of such languages is Italian, whose locative relative marker *dove* ‘where’ can sporadically encode subjects, especially in sub-standard varieties (see Alfonzetti 2002 among others). This is by no means a peculiarity of Italian, as a similar ‘jump’ across the hierarchy has also been observed for English *where* (see Brook 2011). Based on evidence from these languages, we set out to assess the reason why one finds a development that at face value goes against the accessibility hierarchy. As we argue, there are precise motivations that account for this change, and that can only be unveiled if one takes a closer look at sociolinguistic variation and real corpus data. Specifically, we claim that once properly understood, these changes do not in fact contradict the accessibility hierarchy, but simply show that there are different pathways of change from locative to core-argument relatives than the ones commonly discussed in the literature. In our view, in the change from locative to core-argument relatives, a crucial role is played by the intermediate step whereby location marker develops a discourse-oriented function to link the main and relative clause in a topic – comment relationship.

The paper is structured as follows. In section 2 we briefly survey the existing typology of relative clause constructions (section 2.1) and discuss the accessibility hierarchy and its synchronic and diachronic implications (section 2.2). We then turn to an overview of the typology of locative relatives. We first illustrate the ways in which location and other spatial roles can be relativized in the world’s languages (section 3.1), and we then discuss on the origin of specialized locative relative markers (section 3.2). In section 3.3, we focus on the possible processes of functional extension undergone by locative relatives. In section 4, we shift our attention to the development of locative relatives in the languages of Europe. After a general presentation of the evidence from standard and non-standard varieties of European languages (section 4.1), we zoom in on the situation of Italian *dove* and discuss its behavior based on corpus data (section 4.2). In section 5, we discuss a possible diachronic scenario that accounts for the emergence of the new functions of *dove*, which has

important bearings on the general diachronic typology of locative relatives. Section 6 recapitulates the findings of the paper.

2. The typology of relative clauses

2.1. Definition and parameters of variation

Ever since Keenan & Comrie's seminal work (1977), relative clause (RC) constructions have been a core topic of research in typology, so that we now know a lot about the cross-linguistic variation in the domain of relativization strategies in terms of structures and functions (see e.g. Keenan 1985, Lehmann 1986, de Vries 2002, 2005, Andrews 2007). In what follows, we briefly recall the main parameters usually take into account in the typologization of RCs.

Different definitions of what cross-linguistically counts as a RC have been proposed over the years. Even though the details of individual proposals may differ, there is a common agreement that cross-linguistically one can define RCs as a type of "subordinate clause which delimits the reference of a noun phrase (NP) by specifying the role of the referent of that NP in the situation described by the RC." (Andrews 2007: 206). This definition successfully captures some defining morphosyntactic and semantic features of RCs. Semantically, RCs have the purpose of delimiting the reference of NPs in discourse (clearly, this definition only applies to restrictive RCs, and not to appositive ones; see also Shagal 2019: 22-23). Moreover, finite RCs differ from other potentially functionally equivalent constructions, such as adjectives, participles, nominalized form, in that they are clauses, that is, they feature a finite verbal form and are linked to other clauses in a subordination relationship (on nominalization and relativization see Yap *et al.* 2011: 27-33; on participles as relativization strategies see Shagal 2019: 21-30, Chap. 6).

RC constructions as defined above show a remarkable degree of variation across languages. RCs may be classified according to both semantic and structural parameters into several different types (de Vries 2002, 2005, Cristofaro 2003, Andrews 2007, Hendery 2012). The main parameters of variation, as discussed by de Vries (2002: 233), are shown in (1):

(1) a. kind of modification/relation:	restrictive, appositive, degree
b. hierarchical status of RC:	embedded within DP, correlative
c. presence of head:	headed/free relatives
d. presence of relative pronoun:	yes/no
e. presence of complementizer:	yes/no
f. presence of resumptive pronoun:	yes/no
g. hierarchical position of head:	externally/internally headed RCs

h. linear order of head and RC:	head-initial/final relatives
i. inflectional completeness of RC:	finite/participial relatives
j. position of Det with respect to N and RC:	initial/middle/final
k. position of (Case) markers, if any:	on N, on N and RC

Among the various criteria in (1), we would like to focus on the possible marking of RCs. As is well-known, RC constructions can be variously classified also based on the type of relative marker or lack thereof (see Comrie & Kuteva 2013 for an overview). Since in this paper we are especially concerned with the development of dedicated relative markers, that is, locative relatives, we operate a more coarse-grained distinction between overt vs. non-overt strategies. The former features an explicit marker of the RC (e.g. *The man **that** I saw*), while in the latter type, no dedicated marking occurs (e.g. *The man I saw*). Zooming in on overtly marked RCs, there are two main types that fall within this category: RCs that make use of relative pronouns and those that make use of general relativizer.

A relative pronoun can be defined as a “pronominal element [that] is case-marked (by case or by an adposition) to indicate the role of the head noun within the relative clause.” (Comrie & Kuteva 2013). Relative pronouns are generally considered a typical feature of the languages of Europe (see Cristofaro & Giacalone Ramat 2007 for discussion). As an example, consider the German relative pronoun in (2):

- (2) *Der Mann, [den du kennengelernt hast]_{REL}, kommt aus Griechenland*
 ART.F.SG man(M) REL.ACC.M 2SG.NOM know.PTCP have.PRS.2SG come.PRS.3SG
 from Greece
 “The man that you met comes from Greece.”

In (2), the RC (in square brackets) is introduced by the relative pronoun *den* that inflects for gender, case, and number, and agrees with the head noun *der Mann* ‘the man’ in the main clause. The syntactic function of the head noun within the RC is explicitly indicated by the accusative case marking on the pronoun, which indicates that the relativized item is the direct object of the verb of the RC.

Alternatively, overt RCs may feature a *general relativizer*. We define a general relativizer as a marker of relativization that does not feature any indication as to the role that the relativized item

performs within the RC. An example of a general relativizer is Matigsalug Manobo *ne* in (3).¹ In this case, the element that indicates the RC does not bear any case marking, so that the syntactic role of the head must be retrieved from other co(n)textual cues.

(3) Matigsalug Manobo (Austronesian, Malayo-Polynesian; Wang *et al.* 2006: 140)

b. *ne sapulu ka etew ne mignangenan ni Ame*
 now ten the people REL informed by dad
 “Now there are ten people who were informed by dad.”

2.2. The accessibility hierarchy

As anticipated in Section 1, Keenan & Comrie’s (1977) *noun phrase accessibility hierarchy* (henceforth, AH) is perhaps one of the most influential cross-linguistic generalizations on the distribution of RCs (see also Keenan & Comrie 1979). The hierarchy is represented in (4).

(4) *subject > direct object > indirect object > oblique > genitive > object of comparison*

In its original formulation, the AH served as an indication of the likelihood of the relativization of individual syntactic roles. Following the AH, the subject is most easily relativized role, as opposed to object of comparison, which is the least likely. Keenan & Comrie (1977: 67) further proposed that three constraints operate on the AH:

- “1. A language must be able to relativize subjects.
2. Any Relative Clause-forming strategy must apply to a continuous segment of the Accessibility Hierarchy.
3. Strategies that apply at one point of the Accessibility Hierarchy may in principle cease to apply at any lower point.” (Keenan & Comrie 1977: 67)

In practice, the AH makes strong predictions as to the number and types of syntactic roles that individual RC markers can be associated with, and it has proved a remarkably useful tool to describe the distribution of RC markers both cross-linguistically and within individual languages (see e.g. Cristofaro & Giacalone Ramat 2007 on European languages). The second constraint will be of special

¹ General relativizers are often polyfunctional, and besides being used as RC markers, they frequently function as more general linking elements, subordinators and/or complementizers (on the historical reasons behind such a polyfunctionality see e.g. Hendery 2012).

relevance in the remainder of this paper. Since it refers to the application of relativization markers to contiguous regions on the AH, we refer to it as the *contiguity constraint*. As Keenan & Comrie (1977: 67) phrase it, the contiguity constraint entails that “a language is free to treat adjacent positions on the AH as the same, but it cannot “skip” positions”.

Keenan & Comrie’s work has prompted an intensive line of research on the typology of RCs, and over the decades the study of several languages not included in Keenan & Comrie’s sample has eventually offered an increasing body of evidence in support of the generalizations originally advanced by the AH. Languages whose RCs deviate from the AH constraints remain to our knowledge few in number and the existence of a few language-specific exceptions is not generally considered a sufficient counterargument to the strong validity of the AH (see already Keenan & Comrie 1977: 85-87 for the discussion of possible counterexamples from Hausa and Tongan; see also e.g. Cinque 1981, Fox 1987).

To illustrate the type of counterevidence that has been adduced against the AH, consider as an example the inventory of RC markers in Kambera (Klamer 1998: 318-338). This language features two main verbal prefixes that function as relativizers, that is, *pa-* and *ma-*. The former is used to relativize direct objects, indirect objects, as well as several obliques, so that its distribution complies with the AH. By contrast, the relative prefix *ma-* is the main strategy to relativize subjects, both in A and S position, but it can also relativize possessors. This means that in this language “possessors may be relativized using the same strategy as subjects” (Klamer 1998: 321), but the same strategy is not used for other intermediate positions on the AH, thus openly violating the contiguity constraint. The distribution of relativizers in Kambera is summarized in Table 1.²

	Subj	DO	IO	Obl	Gen
<i>pa-</i>	-	+	+	+	-
<i>ma-</i>	+	-	-	-	+

Table 1: *Relativizers in Kambera and the AH (Klamer 1998)*

² Interestingly, historical considerations may shed light on the reason why Kambera *ma-* seemingly fails to comply with the AH, as the prefix originated as a relational noun marker in possessive and attributive constructions (see Holton 2006). This further supports the idea that exceptions to the AH are best understood when one considers the historical processes that give rise to specific RC constructions (see Section 4).

As is the case of other implicational hierarchies that have been proposed in linguistic typology (see Cristofaro & Zúñiga 2018 with references and discussion), the AH was originally meant to capture the synchronic distribution of a specific phenomenon across and within languages. The arrangement of the different syntactic roles on the AH, backed up by cross-linguistic data, was explained as reflecting a synchronic functional motivation, in this case “psychological ease of comprehension” (Keenan & Comrie 1977: 88). However, it has soon been observed that the AH also makes implicit predictions (and constraints) on the possible direction and progress of changes in the (re)distribution of RC markers (*i.a.* Maxwell 1982, Romaine 1984; see discussion in Hendery 2012: 133-136). Maxwell (1982) makes several predictions on the possible changes in the distribution of relativization strategies in time. Among these, he specifically observes that the synchronic contiguity constraints observed by Keenan and Comrie is the outcome of a diachronic constraint whereby markers that extend their syntactic scope to the encoding of other roles on the AH cannot skip any position on the hierarchy. Concerning the direction of the spread of markers across the AH, two different hypotheses have been formulated. According to Maxwell (1982), non-overt strategies tend to start on the highest positions on the AH, that is, they are first used to relativize subjects, and eventually spread downwards. Romaine (1984) makes the opposite claim. Based on a data from RCs in Germanic languages as well as in Konkani and Tok Pisin, Romaine suggests that new relative markers start from below, that is, they are first used on the lower positions on the hierarchy, and only later they spread to roles higher on the AH (note that the two claims are not contradictory, as the Maxwell’s concerns non-overt RCs and Romaine’s overt RCs). As we discuss in the next section, Romaine’s claim fits well the general trend in the development of locative relatives.

3. Locative relatives cross-linguistically

3.1. The encoding of locative relatives and the AH

Locative RCs are a subtype of RCs that can be defined as the relative construction in which the head within the RC expresses a spatial configuration of some sorts. As is expected from the general complexity of the semantic domain of space, locative relatives may differ as to the type of spatial configuration that they can express. Consider for instance the use of English *where* as attested in the Santa Barbara Corpus of spoken American English in (5).

(5) a. <i>I guess he was trying to like, lure us into a place where they would sell, like, imitation cologne</i> (SBC013)	LOCATION
b. <i>I don’t know where I drew it out</i> (SBC008)	SOURCE
c. <i>traditions tent where everyone can come</i> (SBC054)	GOAL

As the sentences in (5) show, English *where* can express either LOCATION, SOURCE, or GOAL, which are widely considered the basic spatial relations commonly expressed by spatial grams (see Stoltz *et al.* 2014, Haspelmath 2019a with ref.). The encoding of different spatial relations through relative markers remains an underresearched topic, so that we still lack a comprehensive view on the possible different patterns of marking spatial roles in RCs. If we focus on the English data in (5), it turns out that bare *where* can express all spatial roles, even though for SOURCE and GOAL also complex locative relatives involving spatial prepositions such as *from* and *to* are available. This is not the only possible pattern, since for instance both bare Italian *dove* and French *où* can express both LOCATION and GOAL and need extra marking when expressing SOURCE or other more complex spatial notions such as perlocative ‘through’. More research is needed to fully explore the patterns marking of spatial relations in RCs (see Stoltz *et al.* 2014). For simplicity’s sake, in the remainder of this paper we will treat locative relatives as a single class of relative markers that can encode at least LOCATION, without focusing on the expression of other spatial roles.

A final important note is in order on the position of locative relatives on the AH. In its original formulation, the AH does not explicitly feature locative relative markers. In principle, locative relatives fall within the *oblique* group, but Keenan & Comrie (1977: 66) explicitly state that they consider as belonging to the oblique category only those NPs that can be considered arguments of main predicates and not those that have a purely adverbial role. Compare the sentences in (6):

- (6) a. *The city **where** I am going (to) is far away*
- b. *The club **where** the girl was singing*

In (6), the NP *where (to)* is commonly considered as an argument of the two-place motion verb *go*, which features a GOAL in its argument structure. By contrast, in (6), *where* performs an adverbial role, as LOCATION is not part of the valency of the monovalent verb *sing*. In this respect, locative relatives differ from temporal relatives, which always behave as circumstantials (see Cristofaro & Giacalone Ramat 2007 for discussion). While we do acknowledge the reasons for operating such a syntactic distinction between argument and adverbial spatial NPs, for the purpose of this paper we consider as belonging to the oblique relative group all locative relative markers, irrespective of their argumenthood, since we have not found evidence that in relativizing location languages resort to different strategies for argumental vs. non-argumental spatial roles.

Turning to the relativization of location, three main patterns can be observed.³ In the first place, languages that have a general relativizer may express all syntactic roles, including location, by means of the same marker. This is the case of e.g. Akkadian *ša* (Huehnergard 1997: 185-188) and Matigsalug Manobo *ne* (Wang *et al.* 2006: 140-141). In the second place, languages may operate a distinction between strategies employed for core arguments, that is, for the upper end of the AH, and those employed for lower roles, including obliques. For example, in Semelai one finds a distribution whereby the relativizer *=mə* is used for subjects and direct and indirect objects, whereas juxtaposition can be used with direct and indirect object and for obliques as well (cf. Kruspe 2004: 339-346). Finally, several languages exhibit specialized relative markers that can only (or primarily) encode location.⁴ In the remainder of this paper when talking about locative relatives we only refer to such specialized locative relative markers. Examples are numerous from several unrelated languages and include (the list is by no mean exhaustive) Italian *dove*, English *where*, Catalan *on* (Wheeler *et al.* 1999: 532), Lumun *na* (Smits 2017: 232-236), Chamorro *nai* (Topping 1973: 280-281), Gumuz /gw/ (Ahland 2012: 420-421), Breton *eman* (Crahe 2013: 327), etc.

3.2. The origin of locative relatives

While sources of relative markers and the rise of RC constructions in general constitute well-studied topics (see e.g. Lehmann 1986, Heine & Kuteva 2002 [now Kuteva *et al.* 2019], Hendery 2012, Givón 2012), not much is known about the origin of specialized locative relative markers, and a detailed diachronic typology of this domain remains a desideratum. In general, two appear to be the main sources of locative relatives. In the first place, relative pronoun meaning ‘where’ can develop out of nouns meaning ‘place, earth’. This process has been documented for a number of relative markers, including Tibetan *sa* (DeLancey 1986), Thai *thîi* (Kullavanijaya 2008), Lao *thii* (2007: 113), Yongning Na *-di33* (Lidz 2010: 558-562), and possibly Hebrew *‘asher* (Holmstedt 2002; Rubin 2005: 49-50 reconstructs the form as Proto-Semitic **’aθar* and argues that the lexical meaning ‘place’ is still visible in Akkadian *ašrum*). In general, this shift complies with the well-known path of grammaticalization that leads from spatial nouns to more abstract markers of location such as spatial

³ We do not discuss here detailed quantitative data on the distribution of the three patterns in the world’s languages, also because finding information on the relativization of locations is not an easy task, as several grammars do not give details about relativization strategies besides the subject/object roles.

⁴ The picture is further complicated by the existence of different possible strategies to relativize spatial roles in RCs. For example, while both Italian *dove* and English *where* are specialized locative relatives, in these languages relativization of spatial roles is also possible with the general relative pronoun together with spatial prepositions, e.g. It. *da cui* ‘from which’, *in cui* ‘in which’. In this paper, we disregard such alternative ways of relativizing spatial roles, and focus on specialized locative relatives only. See also fn. 11.

prepositions or case markers (see several examples in Kuteva *et al.* 2019). The shift PLACE > WHERE possibly starts in contexts in which a generic noun meaning ‘place’ is used in apposition to place names and is subsequently reanalyzed as a marker of a RC headed by the place name. This development is reconstructed for Thai *thîi* by Kullavanijaya (2008). The type of bridging context in which appositional *thîi* ‘place’ was reanalyzed as relative ‘where’ is exemplified in (7):

- (7) *khùt ɲaw dîn thîi máʔhǎasǐisàtthaarâatcùʔnii yìap*
 dig take soil place M. step.on
 “Dig the soil, the **place** Mahaasrisattharaatculamunee steps on > Dig the soil **where** M. steps on.” (Kullavanijaya 2008: 453)

Another known source of locative relatives are locative interrogative adverbs ‘where?’. This is in its turn part of a wider well-known tendency for interrogative markers to develop into relative markers, which is a shift especially common in the languages of Europe (see Kuteva *et al.* 2019: 356-357 for discussion and references). A well-known example of this type of change is the history of Modern Greek *pou*, which ultimately goes back to the interrogative adverb *pou* ‘where?’ (Nicholas 1998). According to Nicholas (1998), the development semantically runs as follows: first the locative INTERROGATIVE ‘where?’ undergoes a semantic shift to INDEFINITE locative ‘everywhere’ and from this a restrictive RELATIVE locative meaning ‘where’ emerges. Most locative relatives of modern Indo-European languages of Europe ultimately have an interrogative origin (e.g. Poletto & Sanfelici 2017: 818).

3.3. The functional extension of locative relatives

In several languages, locative relative markers as defined in the previous sections are specialized for the locative function and cannot be used in other contexts. However, there is a large body of evidence suggesting that markers that originate as locative relatives may extend their functional range to the relativization of other semantic and/or syntactic roles. Sometimes, the extension is limited to other semantic roles (which syntactically can still be regarded as belonging to the oblique/adverbial class, see Section 3.1). For instance, the Lhasa Tibetan *=sa*, originally locative in function, also extended to the relativization of the semantic role of beneficiary (cf. DeLancey 1986), and similarly in Yongning Na the locative relative *-di33* came also to express the semantic role of purpose (Lidz 2010: 560). We refer to these as cases of *partial extension*. Similar extensions are unproblematic, as they can be understood as instances of a more general tendency for locative markers to be the source of other more abstract case markers, based on the metaphoric mapping of non-spatial semantic roles

onto the spatial domain (see the entries LOCATIVE > in Kuteva *et al.* 2019 for examples; on the role of conceptual metaphors and embodiment in the metaphorical extension of spatial grams see further Luraghi 2016 for discussion and references).

More interesting are cases in which a relative marker that was originally confined to expressing location extends its range to all types of relativization, so that it comes to be the standard marker for the relativization of all roles on the AH in a given language. We label this type of development *total extension* of locative relatives. General relativizers with a locative origin have been documented in a number of genetically and areally unrelated languages, and include modern Greek *pou* (Nicholas 1998), Thai *thii* (Kullavanijaya 2008), Lao *thii* (Enfield 2007: 112-115), Biblical Hebrew *asher* (Watke & O'Connor 1990: 333-334; see Rubin 2005: 49-50 for discussion). Moreover, the generalization of locative relatives as general relativizers is a widespread phenomenon in several English-based pidgins and creoles, e.g. Tok Pisin *we*, Trinidad English Creole *weh*, ultimately based on English *where* (at least 13 languages that feature this pattern are reported in the *APiCS*, Michaelis *et al.* 2013). Clearly, metaphorical extension of the original spatial semantics to more abstract domains, which we have discussed for cases of partial extension, must play a role in the total extension to core syntactic roles as well. However, since in this case one also observes a widening in the syntactic scope of the relative marker, this cannot be the whole story.

Based on this evidence, it seems reasonable to posit the existence of a recurrent grammaticalization path LOCATIVE > RELATIVE. Such a grammaticalization path has already been cautiously proposed by Kuteva *et al.* (2019: 274), who however remark that “more crosslinguistic data are needed to establish this pathway of grammaticalization”. Kuteva *et al.* are right in pointing out that genuine examples of the LOCATIVE > RELATIVE shift remain few in number. However, we would like to stress that the few existing and often-quoted examples all come from cases in which the shift has been fully completed and the original locative marker has imposed itself as the standard general relativizer and is as such registered in reference grammars of the languages at hand. As we discuss in the next sections, data from grammars only represents the tip of the iceberg and a closer look at the actual use of locative relatives in written and spoken production reveals a much more complex and interesting picture.

A final note is in order on the dynamics of the extension of locative relatives. In the first place, it must be stressed that in several cases a LOCATIVE > RELATIVE shift is only inferred to have taken place in the case of general relativizers for which we can etymologically reconstruct a locative origin. In other words, in these cases we can only observe the initial and the final stage of the process and we can only offer a more or less well-informed speculation on how it actually took place. Cases in which the step-by-step diachronic development from the locative origin to the general relativizer stage is well documented are rather the exception. This is an especially important point to keep in mind in

evaluating the development of locative relatives against the AH. Taken at face value, the cases of extension from locative to subject relativization discussed above are all fully compatible with the AH, as the resulting synchronic distribution of the new general relativizer complies with the contiguity constraint. There is thus a more or less tacit assumption that the change is gradual and is a step-by-step shift from lower to higher positions in the hierarchy (Romaine 1984). In this shift, a key stage is arguably played by the extension of locative relatives to the relativization of non-spatial head nouns (thus Kullavanijaya 2008). However, it must be stressed that there is no historical real corpus data that demonstrates that the shift LOCATIVE > RELATIVE operates in a gradual fashion along the AH (see Nicholas 1998 and Kullavanijaya 2008 for discussion). In the next sections, we challenge the purported linearity of this shift, and discuss some interesting cases in which locative relatives skip to higher positions on the AH. To do so, we focus on the patterns of polyfunctionality of locative relatives in the languages of Europe, and we discuss a more in-depth case study from Italian.

4. From interlinguistic to intralinguistic variation

4.1. Standard and non-standard uses of locative relatives in the languages of Europe

The languages of Europe feature a number of different relativization strategies, including the use of relative pronouns, general relativizers, and gapping (see Cristofaro & Giacalone Ramat 2007 for an overview). Several European languages also have a dedicated locative relative marker that, according to normative grammars, can only convey a spatial value. This is the case, among others, of Italian *dove* (Serianni 2006 [1989] VII, § 247), English *where* (Huddleston *et al.* 2002: 1050), French *où* (Price 2008: 191 ff) and German *wo* (Götze & Hess-Lüttich 1999: 434). Nevertheless, in recent years studies based on non-standard varieties and on real spoken and/or written corpus data have evidenced that, at least in some varieties of these languages, the originally locative pronoun has broadened its functional domain to other types of relativization in actual usage.

An overview of this phenomenon can be found in Murelli (2011: 182-193). The author classifies locative relativizers of several European languages according to their range of functions, also taking non-standard varieties into account. According to Murelli, three are the main patterns of polyfunctionality of locative relativizers, and these can be organized on a cline of none/partial/total extension similar to the one that we have proposed in section 3.3. These are summarized in *Table 2* (the table offers a somewhat simplified version of Murelli's typology, see Murelli 2011: 184 for details).

	Locative	Oblique	Subj, DO, IO
EST <i>kus</i> POL <i>gdzie</i> RUM <i>unde</i> RUS <i>gde</i>	+	-	-
DUT <i>waar</i>	+	+	-
BUL <i>deto</i> FRE <i>dont</i> GER <i>wo</i> (dialects) GRE <i>pu</i> LIT/LTV <i>kur</i>	+	+	+

Table 2: *Use of locative relatives in some European languages (adapted from Murelli 2011: 184)*

First of all, there are elements that do not show any functional extension and can only encode a locative value (e.g. Russian *gde*). Then, there are languages in which locative relativizers show a partial extension to the oblique class (e.g. Dutch *waar*), and lastly there are languages/varieties where these elements are used to relativize all the higher position of the AH (e.g. Bulgarian *deto*, dialectal German *wo*).

Let us consider in more detail some of these cases, in order to evaluate these cases of functional extensions against the background of the AH.

A relatively well-known example of the functional extension of locative relatives in Europe is offered by German. As we have anticipated, according to reference grammars in Standard German *wo* can only relativize location. However, as discussed at length by Fleischer (2004: 224-227), German dialects offer a much richer picture. Specifically, in two Low Alemannic dialects (Basle, Obberrotweil) *wo* came to relativize also obliques (albeit combined with a preposition), direct and indirect objects, and even subjects. This behavior is illustrated in (8), in which *wo* relativizes the noun *e Fisch* ‘the fish’ which performs the role of subject within the RC.

(8) Basle (Low Alemannic, Germanic; Fleischer 2005: 181)

- a. *Das isch e Fisch wò fliegt*
 DEM.N be.PRS.3SG ART.N fish(N) REL fly.PRS.3SG
 “This is a fish that flies.”

Historically, it is clear that the use of *wo* for roles other than location (and obliques) must be an innovation, as also shown by the fact that this is a trait shared by fewer varieties (Fleischer 2004: 234). Commenting upon this development, as Fleischer explicitly states that “*wo* began its grammaticalization as a relative marker in the oblique position, and in some dialects was extended to higher position” (2004: 234). The author suggests that the functional extension of *wo* essentially followed the AH, partly because the synchronic uses of the form are fully compatible with it.

However, the precise diachronic process that led to this situation remains unclear. The case of Low Alemannic dialects is surely interesting in that it shows that the total extension LOCATIVE > RELATIVE also occurs in European varieties beyond the standard ones reported in grammars. However, it does not add much to our understanding of the development of locative relatives *per se*, as it fully complies with the diachronic predictions of the AH (cf. Romaine 1984).

More interesting are cases in which non-standard varieties show a distribution of locative relatives that, if taken at face value, goes against the AH, as these markers apply to non-contiguous regions of the hierarchy. Let us begin by discussing the case of Modern Greek. In the Standard variety of Greek, the pronoun *pou*, which originally had a locative function only (Nicholas 1998), has been established as a general relativizer. Consider example (9), in which *pou* is used to relativize a subject.

- (9) *Kséro ton ánthropo pou méni s tin Athina*
 know.PRS.1SG the man.ACC REL live.PRS.3SG in the Athens
 “I know the man who lives in Athens.” (Joseph 1983: 3)

Besides relativizing subjects, in Standard Greek *pou* may also apply to direct and indirect objects and obliques (Mackridge 1985: 249). Nevertheless, according to Joseph (1983), this situation does not hold for all the speakers. The acceptability test conducted by Joseph shows that the relativization strategy involving *pou* “is impossible at the indirect object position for many speakers while it is possible at the oblique position (and again impossible for all speakers at the genitive position)”. It thus seems that for the interviewed speakers the employment of bare *pou* as a general relativizer applies to a noncontinuous segment of the AH.

Finally, noteworthy are also those cases in which, markers that according to grammars are confined to expressing location, such as English *where* and Italian *dove*, do occur in a much richer range of contexts in actual use and in other sociolinguistic varieties of these languages.

In English, spoken data neatly shows that *where* in some sociolinguistically non-standard constructions may occur to relativize direct objects, as in (10), and even subjects, as in (11).

(10) *I meant, has he ever had a... girl **where** he didn't have to pay her* (SBC045)

(11) *And it's nothing like these movies that are out today like *Scream* and, what is it? *Nightmare on Elm Street* whatever it is, **where** they're actual horror movies* (American National Corpus)

These two examples show that, at least in some non-standard varieties of English, *where* has widened its functional domain in that, beside the oblique value, it can also encode subjects and direct objects. Interestingly, there is no evidence of *where* being used for intermediate positions on the AH (i.e. indirect object and other obliques; but as no thorough corpus investigation of *where* in spoken American has been carried out, this result must be taken as preliminary). In this respect, English spoken data seems to match Joseph's description of Modern Greek *pou*, with *where* skipping intermediate positions on the AH. Note that in cases in which *where* is extended to subject/object relativization, the role of subject/object in the RC is further marked by the occurrence of personal pronouns, e.g. *he* and *they* in (10) and (11), respectively. In this respect, examples like (10) and (11) actually represent new relativization strategies of the *pronoun-retention* type (see e.g. Keenan & Comrie 1977: 92, Comrie & Kuteva 2013). We return to this point in Section 5.1.2.

The distribution of dialectal German *wo*, and non-standard Greek *pou* and English *where* with respect to the AH is summarized in Table 3.

	Subj	DObj	IObj	Obl	Location	Gen	OComp
Low Alemannic <i>wo</i>	+	+	+	+	+	-	-
Gr. <i>pou</i>	+	+	-/+	+	+	-	-
Eng. <i>where</i>	+	+	-	-	+	-	-

Table 3: *Patterns of functional extension of locative relatives in some European languages*

To sum up, whereas data from dialectal varieties of German does not challenge the validity of the AH, evidence from spoken Greek and English shows a distribution of locative relative markers that goes against the AH, as these strategies apply to non-contiguous sections of the hierarchy. Clearly, since we know that these items were all locative to begin with, the challenge to the AH is essentially a diachronic one, *contra* Romaine's prediction (1984): in their functional extension, these locative relativizers skip from lower to higher positions on the AH, thus resulting in a synchronic violation of the continuity constraint. At this point, it still remains to explain how it is possible that such an unexpected functional extension took place, and which diachronic process is able to account for the synchronically puzzling distribution of these markers with respect to the AH. To address these questions, in the following section we zoom in on the case of Italian *dove* 'where'.

4.2. *Dove* in social varieties of Italian

In Standard Italian, the relativizer *dove* is typically described as confined to the relativization of locations, as in (12)a (Serianni 2006 [1989] VII, § 247, Cinque 1988: 468). However, previous studies have already noted that in some varieties the functional range of *dove* is wider than what reported for the Standard, and it can also be used to relativize subjects, as (12)b:

- (12) a. *Le case dove ho vissuto*
 ART.DEF.F.PL houses REL have.PRS.1SG live.PTCP
 “The houses where I have lived.”
- b. *Nel greco c’è un dativo dove può presentare una enne finale*
 in_ART.DEF.M greek there be.PRS.3SG ART.INDF.Mdative(M) REL
 can.PRS.3SG show.INF ART.INDF.F n(letter) final.F
 “In Greek there is a dative that can show a final *n*.” (Bernini 1989: 91)

Constructions like the one (12)b have been observed by several authors (e.g. Berruto 1983, Alfonzetti 2002, Murelli 2011, Cerruti 2017, Ballarè & Micheli 2018, Ballarè *et al.* 2019 among others), who often regard them as marginal.⁵ Taken at face value, examples such as (12)a-b suggests that *dove* extends to the relativization of higher roles on the AH while skipping intermediate positions, in a similar way as *pou* and *where* discussed in Section 4.1.

Nevertheless, a detailed account of how frequent constructions such as (12)b are in Italian is still missing. In this section we aim to fill this gap and present the results of a corpus-based analysis of the uses of *dove* in spoken Italian.

4.2.1. Data and methods

The study has been carried out on the ParlaTO corpus (Cerruti & Ballarè 2020) and aims at investigating the behavior of *dove* in different social varieties of spoken Italian (see also Ballarè *et al.* 2019). The corpus consists in almost 50 hours (552.461 tokens) of semi-structured interviews

⁵ Even if a diachronic analysis of *dove* falls out of the scope of this paper, it is interesting to note that evidence of the polyfunctionality of the investigated pronoun has been found also in Old Italian (see De Roberto 2008). This comes as no surprise: as discussed in detail by D’Achille (1990), many linguistic features which seem to characterize contemporary Italian have actually been attested in texts from the origins to the beginning of the 19th century.

collected in Turin between 2018 and 2019 and provides access to a wide range of metadata (such as educational status, occupation, age range, birth region, etc.) allowing for sociolinguistic analyses. We have extracted all occurrences of *dove* from the corpus, and manually selected only those in which *dove* behaves as a relative pronoun (thus excluding those in which it functions as interrogative pronoun etc.). The selected tokens have been classified according to linguistic and extralinguistic parameters. In doing that, we created a semasiological sociolinguistic variable, in that we were not investigating if “alternative ways of saying the same thing” (Labov 1972: 188) correlate with extralinguistic factors; instead, we were studying whether one form (i.e. *dove*) conveys different functions in different sociolinguistic varieties of language.

Concerning the extralinguistic parameter, we have considered as a sociolinguist variable the educational status of the speakers because, according to the literature, this is the most significant parameter to observe social variation in Italian (see Berruto & Cerruti 2015: 21 ff. for a discussion). Based on their education, speakers have been divided in two groups (see Table 3), and each occurrence of *dove* has been annotated for the social group of the speaker.

Social groups	Ed. 1	Ed. 2
	<ul style="list-style-type: none"> - Elementary school diploma; - Middle school diploma; - High school diploma <i>and</i> currently working and not studying. 	<ul style="list-style-type: none"> - PhD; - Master’s degree; - Bachelor’s degree; - University students.
Number of tokens	283.799	268.662

Table 3 Social groups - ParlaTO corpus

Turning to the linguistic side, we have selected three parameters that allow us to take into account the syntactic and semantic relations expressed by *dove* as well as the semantics of the relativized item. The linguistic parameters and their values are as follows:

- The syntactic relation expressed by the relative pronoun:
 - OBL: oblique
 - SUBJ: subject
 - OBJ: object
 - CONC: concern connective
- The semantic relation expressed by the relative pronoun:

- SPA: spatial
- NOSPA: non-spatial
- The semantics of the head:
 - LOC: location
 - ABSTR: abstract
 - TEMP: temporal
 - HUM: human
 - CONCR: concrete

The first parameter refers to the syntactic role of the item relativized by *dove*. In the first place, *dove* can be employed to relativize obliques, typically locations. As mentioned before, according to grammars, is its only standard use. However, in the corpus there are cases in which *dove* may also relativize subjects or even objects.

In addition, one also finds a wealth of occurrences in which *dove* cannot be characterized as relativizing a specific syntactic role, but rather acts as a more general linker between the main clause and the RC. This function has been labelled by Ballarè & Micheli (2018) *concern*. Specifically, in these cases, *dove* can be conceived as introducing a RC in which the head NP does not have a specific syntactic role, but rather serves as the sentence topic.⁶ This is for example the case in (13), in which the RC introduced by *dove* seems to provide additional information to the relativized head *madri* ‘mothers’. It is important to stress that this function of *dove* is not easily captured by any of the syntactic roles traditionally assumed for the AH. We return to the notion of *concern* in section 5.

- (13) *Chiaramente la FIAT è un po' come quelle*
 clearly ART.DEF.F Fiat(F) be.PRS.3SG a bit like DEM.PL.F
madri un po' oppressive dove non ti
 mother(F).PL a bit oppressive.PL.F REL NEG 2SG.OBL
manca mai nulla
 lack.PRS.3SG never nothing

“Clearly FIAT is a bit like those mothers a bit oppressive where you never need anything.”⁷

⁶ To be sure, we do not claim that the relativized NP has a topical status within the main clause or in the general discourse. We use *topic* in the narrow sense of Lambrecht as “the thing which the proposition expressed by the sentence is ABOUT” (1994: 118), to underscore the fact that the RC essentially provides additional information about the head noun. Note that in this respect, the *concern* use of *dove* partly blurs the line between restrictive and appositive RCs.

⁷ All the examples presented in sections 4.2.1 and 4.2.2 are extracted from the ParlaTO corpus.

The second parameter refers to the spatial/non spatial relation expressed by the relative pronoun and aims at distinguishing the different kind of obliques that *dove* can relativize. In fact, while in the majority of cases it has a spatial value, it can also acquire a non-spatial semantics, expressing temporal relationships, as in (14), or more abstract relationship in which the spatial dimension, if any, is purely metaphorical, as in (15).

Finally, we have annotated the semantics of the head to which *dove* refers. In addition to the expected locative heads, one also finds that temporal (e.g. (14)), concrete, abstract (e.g. (15)), and human NPs can function as heads of *dove*.

- (14) *Infatti c' è stato un periodo dove*
indeed there be.PRS.3SG be.PTCP A RT.INDF.M period(M) REL
hanno ripulito delle zone del Valentino
have.PRS.3PL clean_up.PTCP PART.PL.F area(F).PL of_ ART.DEF.M Valentino(M)
“As a matter of fact, there has been a time in which they cleaned up some areas of the Valentino park.”

- (15) *Ormai esiste la formula dove invece che pagar=ti*
by_now exist.PRS.3SG ART.DEF.F formula(F) REL instead that pay.INF=2SG.OBL
così ti assumo
thus 2sg.OBL hire.PRS.1SG
“By now there is a formula according to which instead of paying you in this way, I hire you.”

4.2.2. Results

The distribution of *dove* with respect to the different linguistic parameters and across the two social groups of speakers is provided in Table 4.

	Ed. 1		Ed. 2	
syntactic relation	OBL = 239		OBL = 191	
semantic relation	SPAT = 192	NON-SPAT = 47	SPAT = 180	NON-SPAT = 11
sem. of the head	LOC = 186	ABSTR = 21	LOC = 175	TEMPO = 4
	HUM = 4	TEMP = 15	CONCR = 4	ABSTR = 3
	CONCR = 2	LOC = 6	HUM = 1	CONCR = 3
		HUM = 5		LOC = 1

syntactic relation	CONC = 26	CONC = 9
sem. of the head	ABSTR = 18	LOC = 5
	LOC = 3	ABSTR = 4
	HUM = 3	
	CONC = 2	
syntactic relation	SUBJ = 6	SUBJ = 0
sem. of the head	LOC = 5	
	HUM = 1	
syntactic relation	OBJ = 1	OBJ = 0
sem. of the head	ABSTR = 1	
Tots.	272	200
	472	

Table 4: *Distribution of dove*

The two columns give account of the distribution of the occurrences of *dove* in the 2 social groups (i.e. Ed. 1 and Ed. 2), while the linguistic parameters are disposed over the rows of the table.⁸ As can be seen from Table 4, the cases in which *dove* behaves as a prototypical locative relative pronoun, expressing a locative relation and referring to spatial nouns are 361 on 472 (76.5%). As for the remaining 111 occurrences, in 69 cases (14.6%) it relativizes an oblique, even if it refers to non-locative nouns, as in (16), and/or it convey a non-spatial semantic relation, as in (17) and (18).⁹

(16) *Son tornato dal padrone dove lavoravo prima*
 be.PRS.1SG return.PTCP from_ART.DEF.M boss(M) REL work.PST.1SG before
 “I came back to the boss with whom I used to work before.”

(17) *C’ erano molti gruppi di amici dove si*
 there be.PST.3PL many.PL.M group(M).PL of friend.PL REL 3SG.REFL
stava volentieri al bar a giocare a calcio
 stay.PST.3SG gladly to_ART.DEF.M bar(M) to play.INF to football
 “There were many groups of friends with which we gladly stayed at the bar playing soccer.”

⁸ Note that, as said before, the parameter *semantic relation* is relevant only when associated with the oblique *syntactic function*.

⁹ In the oblique class, we have grouped together both argumental obliques as defined by Keenan & Comrie (1977) as well as circumstantial ones (see section 3.1).

(18) *Non mi viene in mente un aneddoto dove*
 NEG 1SG.OBL come.PRS.3SG in mind ART.INDF.M anecdote(M) REL
possa ricordar=mi qualcosa
 can.SBJV.3SG remember.INF=1SG.REFL something
 “I can’t think of an anecdote about which I can remember something.”

In the other 42 cases, *dove* does not express an oblique syntactic relation. In fact, in 35 occurrences (7.4 %) behaves as a concern connective, as defined in Section 4.2.1. At a closer look, in these cases, *dove* does not behave as a proper relative pronoun, in that it acquires a more discursive value allowing the speaker to manage the information flow, as in (19). This value can be expressed also in cases in which the head is a spatial noun, as in (20).

(19) *Sono sempre stata una dove i miei*
 be.PRS.1SG always be.PTCP.F ART.INDF.F REL ART.DEF.PL.M 1SG.POSS.PL.M
genitori non mi hanno mai vincolata
 parent(M).PL NEG 1SG.OBL have.PRS.3PL never bind. PTCP.F
 “I have always been one of those girls, such that my parents never gave me restrictions.”

(20) *Tu la sera devi trovar=ti in un*
 2SG.NOM ART.DEF.F evening(F) must.PRS.1SG find.INF= 2SG.OBL in ART.INDF.M
luogo dove c’ è un sacco di casino, dove
 place(M) REL there be.PRS.3SG a lot of mess REL
anche a te può anche non interessare lo
 also to 2SG.OBL can.PRS.3SG also NEG interest.INF ART.DEF.M
spaccio (...)
 dealing(M)
 “In the evenings you should find yourself a place where there is a real mess, where you might not even be interested in drug dealing.”

Lastly, in the corpus we have found 5 (1.3%) cases in which *dove* relativizes a subject, as in (21), (22), and (23), and 1 (0.2%) in which it relativizes an object, as (24).

(21) *C’ è uno sgabuzzino dove si usava da bagno*

there be.PRS.3SG ART.INDF.M closet(M) REL 3SG.REFL use.PST.3SG as toilet
 “There is a closet that was used as toilet.”

(22) *I ragazzini under 16 dove stan lì a far*
 ART.INDF.PL.M kid(M)PL under 16 REL stay.PRS.3PL there to do.INF
casino al giardinetto
mess to_ ART.DEF.M garden(M)
 “The kids under 16, who stay there making a mess in the park.”

(23) *Il paese dove siamo nati dove siamo*
 ART.DEF.M country.M REL be.PRS.1PL born.PTCP REL be.PRS.1PL
cresciuti, dove ci piace.
grew_up.PTCP REL 1PL.OBL like.PRS.3SG
 “The country where we were born, where we grew up, that we like”

(24) *Ho sempre fatto questo lavoro (...) dove l' ho*
 have.PRS.1SG always do.PTCP DEM.M job(M) REL 3SG.ACC have.PRS.1SG
fatto un po' per delle società da dipendente
do.PTCP a bit for PART.PL.F society(F).PL as employee
 “I have always done this job that I have done for a while for some societies as an employee.”

Even in our corpus, the use of *dove* as a subject/object relativizer is quite rare, thus confirming earlier descriptions of these constructions as marginal. It is important to stress that in 5 cases over 7 the head of the RC is a spatial noun (see ex. (21) and (22)) and this may have played a role in triggering the use of *dove*. Furthermore, to provide a deeper understanding of (23), it may be useful to refer to the notion of *retraction*, a phenomenon that characterizes oral productions and can be identified when “in formulating a next verbal component, the speaker re-uses a syntactic position which has already been filled by some other element before, and puts another element in this position” (Auer & Pfänder 2007: 59). In fact, in the example, the speaker uses *dove* three times in a list construction: while in the first two cases it relativizes an oblique, in the latter it relativizes a subject (see also Ballarè et al. 2019: 88-89). Nevertheless, this is not the whole picture, as *dove* is used to relativize core arguments even with non-spatial heads, as in (22) and (24).

Finally, data in Table 4 shows that, surprisingly, *dove* does not only refer to heads with a locative meaning. In fact, even if locative heads can be found in relation with every syntactic function, there

are always also cases of heads whose semantics is linked with other categories. In the majority of cases, *dove* refers to inanimate items (*locative, temporal, concrete* and *abstract* nouns) but what is striking is that there are also occurrences of *dove* referring to humans, as in (22)¹⁰.

At this point, we can discuss the data from a sociolinguistic perspective. Probably the most interesting aspect of the data reported in Table 4 is the distribution of *dove* across the two social groups. Even though the non-standard uses of the investigated form are numerically scarce and therefore the data should be interpreted with caution, the distribution of the syntactic functions of *dove* in the two social groups is statistically significant, according to the Fisher test with $p < 0.01$.¹¹ In Table 5, we report the percentages of the uses of *dove* in the considered social groups.

	<i>Synt. functions</i>	<i>Ed. 1</i>	<i>Ed. 2</i>
<i>Standard</i>	OBL. (Loc)	192 = 70.6%	180 = 90%
<i>Non-Standard</i>	(Other) OBL.	47 = 17.3%	11 = 5.5%
	CONC.	26 = 9.6%	9 = 4.5%
	SUBJ.	6 = 2.2%	0
	OBJ.	1 = 0.4%	0
<i>Tots.</i>		272 = 100%	200 = 100%

Table 5: *Syntactic functions of dove*

The data in Table 5 shows that in general in both groups standard functions are more robustly attested than the non-standard ones, even if with a noticeable difference (70.6% VS 90%). Furthermore, not only the formers employ the pronoun with *non-standard* values more frequently but, only in this social variety, *dove* can also cover more functions in that, even if sporadically, it can be employed to relativize subjects and objects.

If one arranges the different functions of *dove* attested in different social varieties of Italian on an implicational scale, as in Table 6, one could argue that in different varieties *dove* is situated at different stages of this hypothetical cline, with Ed. 1 speakers being more prone to use *dove* with its non-standard value. As already noted, Ed.1 speakers are the only one who employ *dove* to relativize subjects and objects: this means that their productions are the place where the last step of the extension took place (see *changes from below* in Labov 1994). What remains to be explained is why the form *dove* starts as a locative relativizer and extends to subjects and objects without spreading to

¹⁰ Furthermore, it is interesting to note the presence of *dove* is not restricted to inanimate and uncountable nominals; in fact, it can also occur with animate plural and singular nouns, as in (22) and (16), respectively.

¹¹ $p = 0.006295$, $N = 472$.

intermediate positions on the AH first, and what is the role and status of the function defined as CONCERN in this process. We discuss these issues in more detail in Section 5.

	<i>Standard</i>	<i>Non-standard</i>			
	Obl (loc)	(other) Obl	Conc	Subj	Obj
Grammars					
Ed. 2					
Ed. 1					

Table 6: *The functional extension of dove in social varieties of Italian*

5. From locative to subject (via concern): a possible pathway of change

In section 4.2, we have seen how corpus data shows that Italian *dove* is also used in contexts that do not easily match its original locative value and cannot be easily described by applying the criteria of the AH. We have described *dove* in these contexts as performing a connective function that we labelled *concern*. In the reminder of this section, we argue that it is precisely this concern function that provides the link between the original locative semantics of *dove* and its use as subject/object relativizer.

In the first place, it must be mentioned that this behavior is not isolated, and it is matched by other languages. A case in point is English (see Brook 2011). Compare example (25) with the Italian examples (13), (19) and (20).

(25) *We've got the statement **where** you immediately said I'm hundred percent sure this is the guy* (SBC008)

As discussed for (13), the passage in (25) shows that in real corpus data *where* is used in contexts that cannot possibly involve the encoding of a spatial relationship, as in these cases the head does not show any spatial dimension. A similar behavior has also been pointed out for French *où*. Indeed, even though the French locative relative *où* does not otherwise show signs of having undergone a strong functional extension similar to e.g. Italian *dove*, French spoken data points to the marginal use of the form to relativize abstract noun phrases in a fashion similar to (15) and (24), as exemplified in (26) (see also Larrivée & Skrovec 2016: 9).

(26) *Un domaine proper qui est précis qui est vraiment*

ART.INDF.M domain proper REL be.PRS.3SG precise REL be.PRS.3SG really
très objectif où on peut examiner le malade d'
 very objective REL IMPR can.PRS.3SG examine.INF ART.DEF.M patient(M) in
une façon rigoureuse
 ART.INDF.F way(F) rigorous.F

“A proper domain which is precise which is really very objective, where one can examine the patient in a rigorous way.” (ESLO1 corpus, ESLO1_ENT_001)

What examples (19), (20), (25), and (26) all have in common is that the originally locative relativizers no longer encode a relative relation of the spatial type. Instead, they start being used in contexts in which what is relevant is the topic – comment relationship between the main clause and the subordinate. In fact, as we have already remarked in Section 4.2.1, the RC elaborates upon a topic introduced in the main clause, to the point that *where*, *dove*, and *où* can be paraphrased as *as far as X is concerned*. The point is also made by Murelli (2011: 186-188) who observes that in these cases “the specialized relative element (...) cannot be claimed to explicitly relativize any syntactic position at all” (*ivi* 186). This is why this usage apparently escapes the strictly syntactic classification of the AH, as what matters in these contexts is not the syntactic role of the head within the RC but rather the discourse structure of the complex clause.

Such a connective function of Italian *dove* also explains apparently aberrant usages such as (27) and (28) (both from ParlaTO). In (27), *dove* essentially introduces a parenthetical sentence which serves to remind the hearer of the identity of the person that the speaker is talking about, and bears little syntactic connection with its semantic head, the possessive adjective *sua* ‘his’. In (28), *dove* does not even introduce a sentence, but rather serves as a linker between a topic *dizionari* ‘dictionaries’ and the comment *il piemontese e italiano* ‘Piedmontese and Italian’.

(27) *Oramai è da solo ero l'unica sua spalla*
 by_now be.PRS.3SG alone be.IMPF.1SG ART.DEF.F only.F 3SG.POSS.F shoulder(F)
destra dove lui è Mario Rossi
 right.F REL 3SG.NOM be.PRS.3SG M. R.

“By now he’s alone; I used to be his only right shoulder (he is Mario Rossi).”

(28) *Sono fornita di dizionari dove il piemontese e*
 be.PRS.1SG provided.PTCP.F of dictionary(PL) REL ART.DEF.M Piedmontese and
italiano

Italian

“I have dictionaries – those in Piedmontese and Italian. [I read books in Piedmontese.]”

5.1.1. Stage I: LOCATIVE > CONCERN

The development of the CONCERN use of locative relative pronouns can be explained as follows. In origin, such locative relatives are restricted to expressing location (or other spatial configurations).¹² Due to their inherent spatial meaning, these relative pronouns are easily prone to start relativizing other semantic roles. This is arguably the first step in the development of locative relatives. As we have discussed in Section 3.3, this is for example the case of the Lhasa Tibetan locative relative =*sa* that also expresses beneficiary, and of Yongning Na -*di33* that can also express purpose, as well as the case of Italian *dove* that can also be used to encode temporal ‘when’. As anticipated in Section 3.3, all these processes of semantic extension are relatively trivial, and they are ultimately based on the widespread tendency for the spatial domain and spatial grams in general to be metaphorically used to express other more abstract domains. The semantic shift from location to more abstract domains also leads to a higher frequency of occurrences of non-spatial nouns as heads of the RC introduced by locative relatives, thereby progressively weakening the specific connection of locative relatives with the spatial semantics (see e.g. Table 4).¹³ From the perspective of the AH, such semantic extensions have no consequence for the syntax of relativization, as the markers in question still essentially work within the oblique class, that is, semantic change does not yet trigger syntactic change.

Among possible semantic extensions to more abstract domains, it has been pointed out that spatial grams can also develop into markers of CONCERN. This shift is well attested in the languages of the world especially in relation to adverbs/adposition meaning ‘about’ shifting to ‘concerning’ and constitutes “an instance of a widespread process whereby spatial and temporal markers are grammaticalized in specific context to markers of “logical” grammatical relations” (Kuteva *et al.*

¹² More research is needed to assess to what extent the original semantics of the locative relative may influence its development. For instance, a common feature of English *where* and Italian *dove* is that they are underspecified, in the sense that they can express LOCATION, SOURCE, and GOAL (see example (5) and discussion in Section 3.1). One could speculate that it is their semantic underspecification that favors their development as general relativizers. However, this can hardly be the case, as a similar development has also been observed for Italian *in cui* ‘in which’, which given the occurrence of the preposition *in* was confined to location proper and nevertheless shows signs of a similar functional expansion as *dove* (see Ballarè *et al.* 2019).

¹³ Another process contributing to the extension of locative relatives to non-spatial heads is also metonymy. Consider example (16), in which the non-spatial head *padrone* ‘boss’ metonymically stands for ‘workplace’, which itself would be a typical spatial head of *dove*.

2019: 269, see already Heine & Kuteva 2002: 201). We propose that such a semantic shift may also apply to locative relative markers.

Now, as locative relatives are by nature clause linking devices, this leads to a situation in which locative relatives develop into more general markers of clause linkage with the discourse function of signaling that the content of the RC is about the NP head in the main clause (see fn. 6). This is arguably how the use of Italian *dove* and English *where* as markers of CONCERN came about. Clearly, a metaphor TOPIC IS LOCATION lies at the origin of this development, but one cannot simply dismiss such usages of e.g. It. *dove* as being metaphoric, because the metaphoric shift ultimately triggers a radical shift in the distribution of *dove*. Consider in particular examples (27) and (28), in which the new CONCERN value of *dove* allows its usage in contexts that are structurally incompatible with the original semantics and distribution of the marker. For these reasons, while we do wish to downplay the role of metaphoric extension, we believe that CONCERN can be rightfully considered a distinct function of *dove*.

5.1.2. Stage II: CONCERN > SUBJECT

The CONCERN function of locative relatives offers the natural bridge between the original locative semantics and subject relativization. Indeed, topic markers have been repeatedly pointed out to be a possible source of subject marking (see Lambrecht 1994: 131-149 and Shibatani 1991). It is thus conceivable that from cases in which locative relatives correlate with topicality they may be further extended to relativizing subjects, especially in languages in which the sentence topic usually coincides with the subject, as is the case of e.g. English and Italian. Once the syntactic shift from locative to subject has taken place, the same marker may also have been extended to the relativization of other core arguments, such as objects, as in the case of Eng. *where* in (10) and It. *dove* in (23).

More generally, the plausibility of the CONCERN > RELATIVE shift is supported by the existence of a general trend of discourse markers, including markers of topic, to develop into relative markers proper (Hendery 2012: 66). It is interesting to observe that in both the case of Italian and English the CONCERN usage mostly emerges in sub-standard varieties and in real spoken and written production, and this incipient change has not yet made it into the reference grammars of the respective languages.

A final note is in order on the syntax locative relative markers used for subject/object relativization. In these cases, we believe that the locative markers can be better viewed as general relativizers, as described in Section 3.1. This means that the locative relative does not carry any information as to the syntactic role of the relativized head and instantiate a [- case] strategy in the sense of Keenan & Comrie (1977). In other words, the use of a locative relative marker becomes the strategy to relativize subjects (and objects), but this does not necessarily entail that the subject role is now marked by the

relativizer. This analysis also explains why for example one observes differences in the syntax of Italian *dove* as compared to English *where*.

In Italian, in which null subjects are notoriously allowed, the occurrence of *dove* alone is sufficient as a subject relativization strategy, as in (22). However, as null objects are preferably avoided, it is not surprising that in (24) the syntactic role of the head within the RC is further marked by the enclitic pronoun *l(o)* ‘it’. This therefore brings about a split whereby *dove* alone can relativize the subject position, but in object position a heavier construction involving a resumptive pronoun is used, i.e. a *pronoun-retention* strategy (Comrie & Kuteva 2013). This complies with Keenan & Comrie’s (1977: 92) prediction that pronoun retention is more likely to occur on lower roles on the AH. Interestingly, a similar behavior has been observed by Fleischer (2004: 227) for some German dialects, in which *wo* may also co-occur with the pronoun *der, die, das*, and the pronoun is optional for higher roles on the AH but compulsory for lower roles (Fleischer 2004: 230).

In English, in which null arguments are virtually unallowed, the subject and object roles in RCs introduced by *where* must be expressed by a pronoun, thus giving rise to a pronoun-retention strategy for both positions, as in (10) and (11). It must be remarked that, while the existence of the pronoun-retention strategy may appear surprising in Italian and English, there is abundant evidence that this pattern already exists in non-standard and informal varieties of these languages with relativizers other than *dove* and *where* (see Cerruti 2019 for Italian and Radford 2019 Chap. 2 for English).

5.1.3. *Revising the locative > relative development*

To sum up, our proposal is that the fine-grained analysis of spoken data of substandard varieties Italian (as well as the parallel development in e.g. English and French) suggests the existence of another cline of development that involves locative relatives and that does not necessarily follow the diachronic predictions of the AH as discussed by Romaine (1984).

Crucially, the process that we have described cannot be merely explained in terms of syntactic expansion. Instead, we assume the existence of a complex shift that involves different components. What we observed is a first semantic shift LOCATION > CONCERN, which is based on metaphorical thinking and is exploited by speakers to serve specific discourse functions, that is, to manage discourse structure in terms of topic-comment relations. The second shift CONCERN > SUBJECT (> OBJECT) instead involves the syntacticization of a what is essentially a discourse-based relation. Overall, this complex change can be ultimately described in terms of grammaticalization, as among other things one observes that *dove* undergoes semantic bleaching and starts being used in a number of contexts previously incompatible with its original meaning, all of which are typical of grammaticalization processes (e.g. Lehmann 2016, Hopper & Traugott 2003, Himmelmann 2004).

Nevertheless, it must be remarked that the mechanisms that underlie the two subparts of the process are by no means identical, and the two sub-processes are in principle entirely independent from one another: the first part of the process is possible because of the original locative meaning of the form, while the second part takes as input the outcome of the first one, that is the CONCERN value, but there is no direct connection between LOCATIVE and SUBJECT as such. The process is schematized in Figure 1.

Figure 1: *The development of Italian dove*

The cline of development sketched for locative relatives, whereby they first acquire a general grounding function (cf. Givón 1987) and subsequently become grammaticalized as subject/object relativizers, is by no means exclusive to locative relatives. As a matter of fact, this type of development from discourse to syntax also more generally pertains to the rise of subordinators and subordinate clauses, which can be by and large understood as a process of grammaticalization of originally shallowly linked discourse units (see Matthiessen & Thompson 1988, Hopper & Traugott 2003: Chap. 7 among others for discussion).

6. Conclusions

As we have discussed in this work, the development of locative relatives proves a very interesting case to put the validity of the AH to a test. We started our investigation by observing that it is commonly believed that the development of locative relatives into general relativizers must take place in compliance with the AH, and that more specifically it is determined by the contiguity constraint. In other words, when locative relatives expand upwards on the AH, they should do so without skipping intermediate positions. This commonly accepted assumption is however only based on the observation of synchronic data from grammars: since the distribution of these general relativizers of locative origin ultimately conforms to the AH, then they are also believed to have extended in a step-by-step fashion from the bottom to the top of the AH (thus e.g. Romaine 1984, Fleischer 2004). Whereas this development is in principle perfectly imaginable, compelling historical evidence that confirms it is rather scanty, even in well-documented languages such as Greek.

We have challenged the universality of this diachronic process and have provided evidence that in some cases (Italian *dove* and possibly English *where*) locative relativizers may follow an alternative

path of development, whereby they start as locative relative pronouns and ultimately come to relativize subjects, without being used for any of the intermediate roles in the AH.

Crucially, in the case of English and Italian, this development emerges once we move away from typological generalizations based exclusively on data from descriptive grammars and we take into account also data of real language use, as the ones contained in spoken corpora. Furthermore, it must be noted that corpus data, in our case, are characterized by inherent variation: the paths of development discussed here are at an incipient phase and can be especially identified in sub-standard varieties and co-exist the original locative meaning of the scrutinized form.

These findings also have interesting bearings on the status of the AH from a diachronic perspective and makes the paper's result relevant to the ongoing debate on the status and explanation of typological hierarchies and universals more in general (see Cristofaro 2017, Cristofaro & Zúñiga 2018, Mithun 2018, Haspelmath 2019b). Again, the AH is commonly believed to ultimately reflect a principle of ease of processing, and it is also thought to serve as a constraint for possible diachronic changes. Both assumptions are challenged by the discussed data.

On the one hand, ease of processing, and more generally psycholinguistic reasons, have often also been invoked as an explanation to AH effects in RCs production, comprehension, and acquisition (both L1 and L2) (see e.g. Keenan & Hawkins 1987, Hawkins 1999, Song 2002, Diessel & Tomasello 2005, Shirai & Ozeki 2007, Luzi 2012). Specifically, more complex strategies (e.g. pronoun-retention, see discussion in section 4.2.2) are preferred on lower roles on the AH because relativization with these roles is easier to process, while less overt strategies are preferred for higher roles because of their easier processability. From this line of reasoning, it follows that roles that employ similar strategies have a comparable degree of processability. However, it is doubtful whether processing played a significant role in the shift from locative to subject relative markers. Specifically, our data does not unambiguously support the idea that locations and subjects receive the same marking because they have a similar processing status.¹⁴ On the contrary, as we have discussed the reason why the same strategy applies to both locations and subject is to be sought in a complex process that can be divided into two independent subparts, a semantic extension LOCATIVE > CONCERN and syntacticization CONCERN > SUBJECT.

On the other hand, our findings suggest that the AH should not be conceived as rigidly imposing constraints on all possible changes in the domain of the extension of locative markers. The case of Italian *dove* shows that changes that go against the AH can also take place under specific circumstances and following dynamics that are not predicted by the AH. Overall, our results call for

¹⁴ More psycholinguistic research may further contribute to clarify this point, but this clearly lies beyond the scope of our paper.

a more systematic rethinking of the synchronic and diachronic motivations that lie behind the configuration of the AH and its universal status. More research on the diachronic patterns that bring about the AH is needed to fully explore this domain.

Finally, our last remark is methodological in nature. As we hope to have shown, a careful combination of linguistic typology, corpus linguistics, and the study of (socio)linguistic variation in real use and in non-standard varieties can offer precious insights into language-specific typological oddities and further contribute to refining more general typological models.

Acknowledgement

An earlier version of this paper was presented at the *Association for Linguistic Typology (ALT) 2019 Meeting*, University of Pavia, 4-6 September 2019. We would like to thank the conference participants for their insightful remarks. We also would like to thank two anonymous reviewers whose critical comments and observations have helped us greatly improve our paper. The remaining shortcomings are, needlessly to say, our own. Guglielmo Inglese also acknowledges the support of the Research Foundation Flanders – FWO (grant n. 12T5320N). This paper results from close collaboration between the authors. For academic purposes, Silvia Ballarè is responsible of Sections 4, 5 and 5.1.3 whereas Guglielmo Inglese is responsible of Sections 2, 3, 5.1.1 and 5.1.2. Sections 1 and 6 have been written by both authors.

References

- Ahland, Coleen Anne 2012. *A grammar of Northern and Southern Gumuz*. Doctoral dissertation, Eugene, University of Oregon.
- Alfonzetti, Giovanna. 2002. *La relativa non-standard. Italiano popolare o italiano parlato?* Palermo: Centro di Studi Filologici e Linguistici Siciliani.
- Andrews, Avery. 2007. Relative clauses. In *Language typology and syntactic description. Vol. 2, Complex constructions* (2nd ed.), Timothy Shopen (ed.), 206–236. Cambridge: Cambridge University Press.
- Ballarè, Silvia & M. Silvia Micheli. 2018. Usi di *dove* nell’italiano contemporaneo: costruzioni relative e dinamiche di ristandardizzazione. *Linguistica e Filologia* 38, 29-56.
- Ballarè, Silvia, Massimo Cerruti & Eugenio Gorla. 2019. Variazione diastratica nel parlato di giovani: il caso delle costruzioni relative. In *Le tendenze dell’italiano contemporaneo rivisitate. Atti del 52esimo congresso internazionale di studi della Società di Linguistica Italiana (SLI), Bern, 6-8/09/2018*, 75-94. Milano: Officinaventuno.
- Bernini, Giuliano. 1989. Tipologia delle frasi relative italiane e Romanze. In *L’italiano tra le lingue romanze. Atti del XX Congresso Internazionale della Società di Linguistica Italiana (Bologna,*

- 25-27 settembre 1986), Fabio Foresti, Elena Rizzi & Paola Benedini (eds.), 85-98. Roma: Bulzoni.
- Berruto, Gaetano. 1983. L'italiano popolare e la semplificazione linguistica. *Vox Romanica* 42: 38-79.
- Berruto, Gaetano & Massimo Cerruti. 2015. *Manuale di sociolinguistica*. Torino: UTET.
- Brook, Marisa. 2011. One of those situations where a relative pronoun becomes a complementizer: A case of grammaticalization in progress... again. In *Proceedings of the 2011 Annual Meeting of the Canadian Linguistic Association*, Lisa Armstrong (ed.), 1-7.
- Cerruti, Massimo. 2017. Changes from below, changes from above. Relative constructions in contemporary Italian. In *Towards a new standard. Theoretical and empirical studies on the restandardization of Italian*, Massimo Cerruti, Claudia Crocco & Stefania Marzo (eds.), 32-61. Berlin/New York: Mouton de Gruyter.
- Cerruti, Massimo & Silvia Ballarè. 2020. Parlato: corpus del parlato di Torino. *Bollettino dell'Atlante Linguistico Italiano* 44: 13-38.
- Cinque, Guglielmo. 1988. La frase relativa. In *Grande grammatica italiana di consultazione, Vol. 1*, Lorenzo Renzi, Giampaolo Salvi & Anna Cardinaletti (eds.), 443-503. Bologna: Il Mulino.
- Comrie, Bernard & Tania Kuteva. 2013. Relativization Strategies. In *The World Atlas of Language Structures Online*, Matthew S. Dryer & Martin Haspelmath (eds.). Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Crahe, Maxime-Morvan. 2013. *Le breton de Languidic: étude phonétique, morphologique et syntaxique d'un sous-dialecte du breton vannetais*. Doctoral dissertation, Université Rennes 2.
- Cristofaro, Sonia. 2003. *Subordination*. Oxford: Oxford University Press.
- Cristofaro, Sonia. 2017. Implicational universals and dependencies. In *Dependencies in language: On the causal ontology of linguistic systems*, Nick J. Enfield (ed.), 9-22. Berlin: Language Science Press.
- Cristofaro, Sonia & Anna Giacalone Ramat. 2007. Relativization strategies in the languages of Europe. In *Europe and the Mediterranean as Linguistic Areas: Convergencies from a historical and typological perspective*, Paolo Ramat & Elisa Roma (eds.), 63-93. Amsterdam/Philadelphia: John Benjamins.
- Cristofaro, Sonia & Fernando Zúñiga. 2018. Synchronic vs. diachronic approaches to typological hierarchies. In *Typological hierarchies in synchrony and diachrony*, Sonia Cristofaro & Fernando Zúñiga (eds.), 4-27. Amsterdam/Philadelphia: John Benjamins.
- D'Achille, Paolo. 1990. Sintassi del parlato e tradizione scritta della lingua italiana. Analisi di testi dalle origini al secolo XVIII. Roma: Bonacci.

- de Vries, Mark. 2002. *The syntax of relativization*. Doctoral dissertation, University of Amsterdam.
- de Vries, Mark. 2005. The fall and rise of universals on relativization. *Journal of Universal Language* 6: 125-157.
- DeLancey, Scott. 1986. Relativization as nominalization in Tibetan and Newari. <<http://www.uoregon.edu/~delancey/papers/relnom.html>>
- De Roberto, Elisa. 2008. *Le proposizioni relative con antecedente in italiano antico*. Doctoral Dissertation, Università Roma Tre – Université Paris IV, Sorbonne.
- Diessel, Holger, & Michael Tomasello. 2005. A new look at the acquisition of relative clauses. *Language* 81 (4): 882-906.
- Enfield, Nick J. 2007. *A grammar of Lao*. Berlin/New York: Mouton de Gruyter.
- Fleischer, Jürg. 2004. A typology of relative clauses in German dialects. In *Dialectology meets typology: dialect grammar from a cross-linguistic perspective*, Bernd Kortmann (ed.), 211-244. Berlin/New York: Mouton de Gruyter.
- Fleischer, Jürg. 2005. Relativsätze in den Dialekten des Deutschen: Vergleich und Typologie. *Linguistik online* 24 (3): 171–186.
- Fox, Barbara A. 1987. The Noun Phrase Accessibility Hierarchy reinterpreted: Subject primacy or the absolutive hypothesis? *Language* 63 (4): 856–870.
- Givón, Talmy. 1987. Beyond background and foreground. In *Coherence and Grounding in Discourse*, Russel S. Tomlin (ed.), 175–188. Amsterdam/Philadelphia: John Benjamins
- Givón, Talmy. 2012. Toward a diachronic typology of relative clause. In *Relative Clauses in Languages of the Americas: A typological overview*, Bernard Comrie & Zarina Estrada-Fernández (eds.), 1-26. Amsterdam/Philadelphia: John Benjamins.
- Götze, Lutz & Ernest W. B. Hess-Lüttich (1999). *Grammatik der deutschen Sprache*. Munich: Bertelsmann.
- Haspelmath, Martin. 2019a. Differential place marking and differential object marking. *Language Typology and Universals (STUF)* 72 (3): 313-334.
- Haspelmath, Martin. 2019b. Can cross-linguistic regularities be explained by change constraints? In *Explanation in typology: Diachronic sources, functional motivations and the nature of the evidence*, Karsten Schmidtke-Bode, Natalia Levshina, Susanne Maria Michaelis & Ilja A. Seržant (eds.), 1-23. Berlin: Language Science Press.
- Hawkins, John A. 1999. Processing complexity and filler-gap dependencies across grammars. *Language* 75 (2): 244-285.
- Heine, Bernd & Tania Kuteva. 2000. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

- Hendery, Rachel. 2012. *Relative Clauses in Time and Space: A Case Study in the Methods of Diachronic Typology*. Amsterdam/Philadelphia: John Benjamins
- Himmelman, Nikolaus P. 2004. Lexicalization and grammaticization: opposite or orthogonal? In *What Makes Grammaticalization: A Look from its Fringes and its Components*, Walter Bisang, Nikolaus P. Himmelman & Björn Wiemer (eds.), 21-42. Berlin/New York: Mouton de Gruyter.
- Holmsted, Robert D. 2002. *The Relative Clause in Biblical Hebrew: a Linguistic Analysis*. Doctoral Dissertation, University of Wisconsin.
- Hopper, Paul J. & Elizabeth C. Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press.
- Huddleston, Rodney, Geoffrey K. Pullum & Peter Peterson. 2002. Relative constructions and unbounded dependencies. In *The Cambridge Grammar of the English Language*, Rodney Huddleston & Geoffrey K. Pullum (eds.), 1031-1096. Cambridge: Cambridge University Press.
- Huehnergard, John. 1997. *A Grammar of Akkadian*. Winona Lake (Indiana): Eisenbrauns.
- Joseph, Brian D. 1983. Relativization in Modern Greek: another look at the accessibility hierarchy constraints. *Lingua* 60 (1): 1-24.
- Keenan, Edward. 1985. Relative clauses. In *Language typology and syntactic description. Vol. 2, Complex constructions*, Timothy Shopen (ed.), 141–170. Cambridge: Cambridge University Press.
- Keenan, Edward & Bernard Comrie. 1977. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry* 8: 63-99.
- Keenan, Edward & Bernard Comrie. 1979. Noun phrase accessibility revisited. *Language* 55 (3): 649-664.
- Keenan, Edward & Sarah Hawkins. 1987. The psychological validity of the accessibility hierarchy. In *Universal grammar: 15 essays*, Edward Keenan (ed.), 60-88. London/Sydney: Croom Helm.
- Klamer, Marian. 1998. *A Grammar of Kambera*. Berlin/New York: Mouton de Gruyter.
- Kruspe, Nicole. 2004. *A Grammar of Semelai*. Cambridge: Cambridge University Press.
- Kullavanijaya, Pranee. 2008. A Historical Study of /thii/ in Thai. In *The Tai-Kadai Languages*, Anthony V. N. Diller, Jerold A. Edmonson & Yongxian Luo (eds.), 445-467. London: Routledge.
- Kuteva, Tania, Bernd Heine, Bo Hong, Haiping Long, Heiko Narrog, Seongha Rhee. 2019. *World lexicon of grammaticalization. Second, extensively revised and updated edition*. Cambridge: Cambridge University Press.
- Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

- Labov, William. 1994. The study of change in progress: observations in real time. In *Principles of linguistic change: internal factors*, William Labov (ed.), 73-112. Oxford: Blackwell.
- Lambrecht, Knud. 1994. *Informative structure and sentence form*. Cambridge: Cambridge University Press.
- Larivée, Pierre & Marie Skrovec. 2016. Les relatives en français vernaculaire. *SHS Web of Conferences, Volume 27, 2016, 5^e Congrès Mondial de Linguistique Française*.
- Lehmann, Christian. 1986. On the typology of relative clauses. *Linguistics* 24: 663-680.
- Lidz, Liberty A. 2010. *A Descriptive Grammar of Yongning Na (Mosuo)*. Doctoral dissertation, Austin, University of Texas at Austin.
- Luraghi, Silvia. 2016. The mapping of space onto the domain of benefaction and some unpredicted trends in semantic change. *Linguistics* 54 (2): 339-379.
- Luzi, Eleonora. 2012. Relativization strategies in SLA. *Second Language Research* 24 (4): 443-476.
- Mackridge, Peter. 1987. *The modern Greek language*. Oxford: Clarendon Press.
- Matthiessen, Chrisitan & Sandra A. Thompson. 1988. The structure of discourse and 'subordination'. In *Clause combining in grammar and discourse*, John Haiman & Sandra A. Thompson (eds.), 275-329. Amsterdam/Philadelphia: John Benjamins.
- Maxwell, Dan. 1982. Implications of NP accessibility for diachronic syntax. *Folia Linguistica Historica* 3: 135-152.
- Michaelis, Suzanne, Martin Haspelmath & the APiCS consortium. 2013. Subject relative clauses. In *The atlas of pidgin and creole language structures*, Susanne Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.). Oxford: Oxford University Press
- Mithun, Marianne. 2018. Deconstructing teleology: The place of synchronic usage patterns among processes of diachronic development. In *Typological hierarchies in synchrony and diachrony*, Sonia Cristofaro & Fernando Zúñiga (eds.), 111-128. Amsterdam/Philadelphia: John Benjamins.
- Murelli, Adriano. 2011. *Relative constructions in European non-standard varieties*. Berlin/New York: Mouton de Gruyter.
- Nicholas, Nick. 1998. *The story of pu: The grammaticalisation in space and time of a Modern Greek complementiser*. Melbourne: University of Melbourne.
- Poletto, Cecilia & Emanuela Sanfelici. 2017. Relative clauses. In *Manual of Romance Morphosyntax and Syntax*, Andreas Dufter & Elizabeth Stark (eds.), 804-846. Berlin/New York: Mouton de Gruyter.
- Price, Glanville. 2008. *A Comprehensive French Grammar*. Oxford: Blackwell.

- Radford, Andrew. 2019. *Relative Clauses: Structure and Variation in Everyday English*. Cambridge: Cambridge University Press.
- Romaine, Suzanne. 1984. Towards a typology of relative-clause formation strategies in Germanic. In *Historical Syntax*, Jacek Fisiak (ed.), 437-470. Berlin/New York: Mouton de Gruyter.
- Rubin, Aaron D. 2005. *Studies in Semitic grammaticalization*. Winona Lake (Indiana): Eisenbrauns.
- Serianni, Luca. 2006 [1989]. *Grammatica italiana: italiano comune e lingua letteraria. Suoni, forme, costrutti*. Torino: UTET.
- Shibatani, Masayoshi. 1991. Grammaticization of topic into subject. In *Approaches to Grammaticalization, Volume II. Types of grammatical markers*, Elizabeth C. Traugott & Bernd Heine (eds.), 93-134. Amsterdam/Philadelphia: John Benjamins.
- Shirai, Yasuhiro & Hiromi Ozeki. 2007. The acquisition of relative clauses and the noun phrase accessibility hierarchy: A universal in SLA? Introduction. *Studies in Second Language Acquisition* 29 (2):155-167.
- Smits, Helena J. 2017. *A grammar of Lumun*. Doctoral dissertation, Rijksuniversiteit te Leiden.
- Song, Jae Jung. 2002. Linguistic typology and language acquisition: The accessibility hierarchy and relative clauses. *Language Research* 38 (2): 729-756.
- Stolz, Thomas, Sander Lestrade & Stolz Christel. 2014. *The Crosslinguistics of Zero-Marking of Spatial Relations*. Berlin/New York: Mouton de Gruyter.
- Topping, Donald M. 1973. *Chamorro Reference Grammar*. Honolulu: University of Hawaii.
- Wang, P., R. Hunt, J. McGriff & R. E. Elkins. 2006. *The Grammar of Matigsalug Manobo*. The Summer Institute of Linguistics and the University of Texas at Arlington.
- Watke, Bruce K. & Michael P. O'Connor. 1990. *An Introduction to Biblical Hebrew Syntax*. Winona Lake (Indiana): Eisenbrauns.
- Wheeler, Max, Alan Yates & Nicolau Dols. 1999. *Catalan: A Comprehensive Grammar*. London: Routledge.
- Yap, Foong Ha, Karen Grunow-Hårsta & Janick Wrona. 2011. Introduction. In *Nominalization strategies in Asian languages*, Foong Ha Yap, Karen Grunow-Hårsta & Janick Wrona (eds.), 1-60. Amsterdam/Philadelphia: John Benjamins.

Corpora

American National Corpus: americannationalcorpus.org

Enquêtes Sociolinguistiques à Orléans (ESLO Corpus): eslo.huma-num.fr

ParlaTO Corpus: corpusparlato.com

Santa Barbara Corpus of Spoken American English (SBC): linguistics.ucsb.edu/research/santa-barbara-corpus

Authors' information

Silvia Ballarè, silvia.ballare@unibo.it

Università di Bologna, Dipartimento di Filologia Classica e Italianistica

Via Zamboni, 32

40126 Bologna

Italy

Guglielmo Inglese, guglielmo.inglese@kuleuven.be (corresponding author)

KU Leuven, Department of Linguistics

Blijde-Inkomststraat 21 - box 3308

3000 Leuven

Belgium