CP and COMP in Diachrony

Nigel Vincent
The University of Manchester

Proceedings of the LFG'19 Conference

Australian National University

Miriam Butt, Tracy Holloway King, Ida Toivonen (Editors)

2019

CSLI Publications

pages 314–333

http://csli-publications.stanford.edu/LFG/2019

Keywords: complementizers, accusative and infinitive, causatives, control, recomplementation, Latin, Romance

Vincent, Nigel. 2019. CP and COMP in Diachrony. In Butt, Miriam, King, Tracy Holloway, & Toivonen, Ida (Eds.), *Proceedings of the LFG'19 Conference, Australian National University*, 314–333. Stanford, CA: CSLI Publications.

Abstract: This paper explores the ways in which a single Latin construction, the accusative and infinitive (AcI), has been replaced in different Romance languages. The parallel correspondence architecture of LFG provides an account which is more illuminating and theoretically more economical than that available to approaches which mediate all aspects of grammatical structure through a single set of syntactic categories and projections. Both categories and functions are seen to have their own diachronic profiles and the changes they exhibit over time do not necessarily proceed in parallel. More generally, the paper aims to show how both synchronic and diachronic data are relevant to the construction of theories about the structure and organization of human languages.

1. Introduction[1]

The domain of complement clauses is one where modern grammatical theories differ markedly both from traditional grammar and from each other. Consider the sentence in (1):

(1) Sarah believes that the train will be late.

Traditional grammar would label the string *that the train will be late* as noun clause object on the grounds that a) it could be replaced by a noun phrase such as *the rumour*, b) a noun phrase here would constitute the direct object of the verb *believe*, and c) the string *the train will be late* could stand as an independent finite clause with its embedded role here being signalled by the 'subordinating conjunction' *that*. Within a framework such as LFG with its distinction between f-structure and c-structure, debates have mainly centred around the object part of this traditional definition. Thus, in contrast to Bresnan & Kaplan's original proposal for a separate closed function COMP to be assigned to a constituent like *that the train will be late* Dalrymple & Lødrup (2005) argue that where a verb can take a direct nominal object, as *believe* does in English, this implies that the clausal complement should also have that function. COMP would then be reserved for verbs like *hope* which do not admit a nominal object. Others have gone a stage further and argued that COMP is redundant and all the functions of the embedded clause can be subsumed within OBJ or OBL (Alsina *et al* 2005). This debate is ongoing, with for example Patejuk & Przepiórkowski (2016) and Szűcs (2018) refining and providing further empirical evidence for the COMP-free approach while Belyaev *et al* (2017) argue for the continued recognition of a distinction between OBJ and COMP and the relevance of both in the domain of verbal complementation.

[1] This paper started life as a presentation at the teach-in on LFG and diachrony whch preceded LFG '19. I am grateful to the organizers, Wayan Arka and Jane Simpson, for inviting me to participate, to my co-presenters, Kersti Börjars and Louisa Sadler, to Mary Dalrymple, to those who attended and to the referees for their comments and suggestions. Responsibility for errors of fact or interpretation remains of course my own.

By contrast, there has been relatively little discussion within LFG of the categorial side of things. Although LFG is a framework which allows non-binary branching and exocentric configurations, both anathema to cartographic and nanosyntactic approaches, it is common to find the concept of CP carried over without comment from the Chomskyan tradition as the standard way to represent c-structures of constituents which begin with items like English *that,* French *que* and Hungarian *hogy*, and with it of course the implication that such items are heads. Yet such an assumption is by no means necessary. In the words of Pollard & Sag (1994: 44-5):

> "We are not claiming that the analysis of complementizers as heads is untenable, only that the fundamental intuition underlying such proposals raises as many questions as it answers … But if complementizers are not heads, then what are they? We will take the position that they are a subspecies of *marker*. On our account, a marker is a word that is 'functional' or 'grammatical' as opposed to substantive, in the sense that its semantic content is purely logical in nature (perhaps even vacuous). A marker, so-called because it formally *marks* the constituent in which it occurs, combines with another element that heads that constituent."

To this we may add the diachronic observation that the items that fall under the label of complementizer are always the product of processes of grammaticalization, and in that sense are different from lexical categories like noun and verb where core members may remain stable over centuries and even millennia. In other words, whatever is a C now will have been something else in the past.[2] And yet to date the debate has been exclusively based on synchronic evidence. In the present paper, therefore, we aim to introduce a diachronic dimension by means of a case study: the history of complementizers and complement clauses from Latin through to modern Romance. In section 2 we set out the Latin background before considering in sections 3 through 7 a variety of Romance developments and then in section 8 drawing some general conclusions.

2. Latin and the accusative and infinitive construction (AcI)

Latin had a variety of clausal complementation strategies but the central one for verbs whose semantics imply a propositional complement (thinking, saying, promising, hoping, knowing, believing, etc), and the one we will focus on here, was the one that goes by the traditional name accusative and infinitive construction (AcI) as exemplified in (2) - (4):

(2) sese confestim supsequi dicit
REFL.3SG immediately follow.INF say.PRS.3SG
'He$_i$ says that he$_i$ will follow you immediately' (Caesar *Gall* 6.29.5)

[2] Here and throughout, in order to avoid confusion, we will use C to refer to the category of complementizer and COMP to refer to the function, although the latter is also commonly used as a categorial label in the general literature.

(3) in aqua numquam credidi voluptatem
in water.ABL never believe.PERF.1SG pleasure.ACC.FSG
inesse tantam
in-be.INF such.ACC.FSG (Plautus *Rud.* 458)
'I have never believed that there was such pleasure in water'

(4) populus me vere iurasse iuravit
people.NOM.SG me.ACC truly swear.PERF.INF swear.PERF.3SG
'The people swore that I had sworn truly' (Cic *Fam* 5.2.7)

In these examples the governing verb takes a complement expressing the propositional content of the statement, belief or oath with the subject argument of the embedded verb in the accusative (*sese, tantam voluptatem, me*) and the verb in the infinitive, either present (*supsequi*, *inesse*) or perfective (*iurasse*). The accusative of the embedded subjects here must be generated clause internally since in many instances the governing verb either does not take a direct object, as with *dicere* 'say', or governs a different case, as with *credere* 'believe' which takes the dative (*crede.*IMP *mihi.*DAT 'believe me!'). Nor is the item in the accusative in semantic terms an argument of the governing verb. Note too that the embedded subject and the main clause subject can be coreferential as in (2) where the accusative of the AcI is the reflexive pronoun *sese*.

If we follow the account of this construction in Jøhndal (2012: 79-82), we can therefore assign these verbs the PRED values in (5):

(5) a. 'dicere <SUBJ, COMP>'
b. 'credere <SUBJ, COMP>'
c. 'iurare <SUBJ, COMP>'

In fact, however, nothing crucial hangs on assigning the second argument here the function COMP; the analysis would go through if we chose to follow Alsina *et al* (2005) and Patejuk & Przepiórkowski (2016) and assign it the function OBJ instead. Moreover, there are (admittedly rare) instances such as (6) in which an AcI (*iuraturas in feminae verba praetorias cohortis* 'that the praetorian cohorts would swear allegiance to a woman') can be co-ordinated with a simple NP (*consortium imperii* 'share of the power'):[3]

(6) quod consortium imperii iuraturas-que
that share.ACC power.GEN swear.FUTPRT.ACC.FPL-and
in feminae verba praetorias cohortis
in woman.GEN word.ACC.PL praetorian. ACC.FPL cohort. ACC.FPL

[3] To be precise, in example (6) the element *iuraturas*, marked with the co-ordinating affix *-que*, is the future participle of *iurare*, which taken together with the verb *esse* 'be' forms a future periphrasis. However in the AcI the auxiliary in its infinitival form is, as here, often omitted.

… speravisset
hope.PLUPERF.SUBJ.3SG
'that she had hoped for a share in the empire and that the praetorian cohorts would swear allegiance to a woman' (Tacitus *Ann* 14.11)

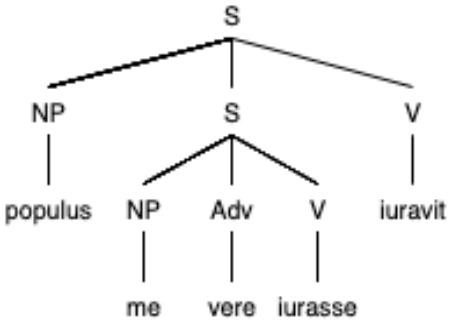
The pattern here is parallel to the English and Polish examples in (7) and (8) cited by Patejuk & Przepiórkowsk (2016) as a motivation for preferring OBJ to (X)COMP as the function to be assigned to clausal and infinitival complements (and see already Sag *et al* 1985):

(7) Pat remembered the appointment and that it was important to be on time.

(8) Nie chciał pić ani kanapki
NEG want.PST drink.INF nor sandwich.GEN
'He didn't want to drink nor (did he want) a sandwich'

What is key, however, is that an analysis of these patterns in terms of f-structure eliminates the need to postulate an empty complementizer head, so that the c-structure assigned for example to (4) would be as in (9):

(9)



Since the AcI, unlike a small clause, has the full range of tense, aspect and argument structure associated with a main clause, Minimalist or cartographic frameworks have little choice but to represent it as a CP. This is the analysis proposed, for example, in Oniga (2014: Ch 23) where the AcI is a CP with a zero complementizer which assigns accusative case, even though in general complementizers do not assign case and in this construction the C can never be overtly realized.

3. Complementizers and growing syntax in Romance

Central as the AcI is to the syntax of complementation in Latin, it comes over time to be replaced by a finite pattern introduced by items such as French *que* and Italian *che,* which derive from the Latin neuter relative pronoun *quid*. However, more frequently attested in Latin texts is the form *quod* which in origin had a causal value:

(10) cum tibi agam gratias
while you.DAT make.PRS.SUBJ.1SG thanks.ACC
quod me vivere coegisti
because me.ACC live.INF compel.PERF.2SG (Cic *Att* 3.3.1)
'while I may give thanks to you because you forced me to stay alive'

And in (11) we can see it being used to mark the complement of *credere* 'believe' in a text where the speaker is being identified as uneducated and uncultured and hence suggesting that this usage was part of popular speech at the time (2nd cent CE):

(11) credo nunc quod Pudentilla me
believe.PRS.1SG now C Pudentilla.NOM me.ACC
in eo tempore non amabat
in that.ABL time.ABL NEG love.IMPERF.3SG (Apuleius *Apol* 79)
'I now believe that at that time Pudentilla did not love me'

It is *quod* which is the etymological source of the complementizer *co/cu* which survives in southern Italian dialects as in the Salentino example (12) and is already found in the earliest Italian text from 960 CE in (13):

(12) oyyu krai ku bbene lu Maryu
want.PRS.1SG tomorrow C come.PRS.3SG DEF Mario
'I want Mario to come tomorrow'

(13) sao ko kelle terre … trenta anni
know.PRS.1SG C that.FPL land.FPL 30 year.PL
le possette parte Sancti Benedicti
it.FPL possess.PST.3SG party Saint Benedict
'I know that those lands have belonged to the party of St Benedict for 30 years'

Interestingly, this complementizer is especially found, as in (12), with clauses that would take the subjunctive in those dialects which preserve that form, whereas causal clauses in Latin always take the indicative, thus suggesting a significant restructuring over time consistent with the loss of the inherent causal meaning.

By contrast, many southern dialects also have an indicative complementizer *ca/ka* which derives from another Latin causal marker *quia*, as in the Old Siclian example (14) from Rinaldi (2005: 473):

(14) dicu ka dichi beni
say.PRS.1SG C say.PRS.2SG well
'I say that you speak well'

This change must have started early since Bennett (1910: I,130) in his grammar of early Latin based on texts from the period before 100 BCE observes: "In apposition with neuter pronouns … the causal notion is usually very slight, *quia* having the force rather of 'that'."

Developments such as these — which we have only been able to sketch: see Ledgeway (2005) for a fuller treatment and further references — raise two questions of more general relevance in the present context. The first concerns the categorial status to assign to these items before they develop the functions exemplified here. One answer would be that even when they have semantic content of their own such as the causal meanings of Latin *quod* and *quia* they are nonetheless complementizers, so that the change is one involving semantic bleaching but not change of category. In this respect then the class of complementizers would be similar to prepositions, where it is common to recognise a distinction between items that have grammatical functions such as English *of* and French *de* and those with semantic content evidence by pairwise contrasts such as *before* vs *after* and *off* vs *on*. Taking this route would also provide a response to the observation by Pollard & Sag quoted above: only some complementizers would have a purely marking function but this would not stop them being treated as heads of CPs any more than it stops a constituent like *of my cousin* being defined as a PP in a construction like *proud of my cousin*. However, an argument against this view is provided by a 15th century Salentino example like (15) (cited in Ledgeway 2005: note 30):

(15) adivene perché ca Adamo lassao
happen.PRS.3SG because C A leave.PST.3SG
lo sua signo
DEF POSS.3SG sign
'it happens because Adam left his mark'

Here the complementizer *ca* co-occurs with the the word *perché* 'because'. Ledgeway's solution is to exploit the split CP hypothesis first put forward by Rizzi (1997) and developed extensively within the cartographic approach to clause stucture since that time. The item *ca* can then be assigned to the lowest functional head Fin while *perché* inhabits the specifier slot associated with the Interrogative head.

The second question follows on from the first, namely how are we to represent the mechanisms of change that are at work in these examples? Börjars *et al* (2016) argued that the development of grammaticalized definiteness markers in North Germanic was a case of 'growing syntax'. That is to say, rather than postulate a universal category DP with only some languages having an overt realization of D, it is proposed that the sole universal category is NP but that in some languages a D slot, and with it a DP projection, comes into existence over time via the well attested process of grammaticalization. By the same token, it might be suggested that there is only evidence for a CP in these structures once the C has emerged, once again via the mechanism of grammaticalization. The difference here, however, is that Latin did have CPs in other context as witness the items *ut* and *ne* in (16) and (17), which introduce subjunctive complements, respectively positive and negative, of the verbs *velle* 'want' and *timere* 'fear':

(16) si vis ut loquar
if want.PRS.2SG C speak.PRS.SUBJ.1SG
'if you want that I should speak' (Martial 5.52.6)

(17) haec … ne impediantur timeo
these things C-NEG hinder.PRES.PASS.SUBJ.3PL fear.PRS.1SG
'I fear these things may be hindered' (D Brutus 6 Cic *Fam*)

It seems then that in this case the CP has not so much 'grown' as 'spread'. However, we defer further discussion of these issues to sections 6 and 7 below and turn instead to two other developments in Romance occasioned by the loss of the AcI.

4. Complex predicate formation: causative and perception verbs

Among the classes of verbs that in Latin could govern an AcI were causatives (18) and perception verbs (19):

(18) ventus … fecit … spissescere nubem
wind.NOM.SG make.PERF.3SG thicken.INF cloud.ACC.SG
'the wind caused the cloud to thicken' (Lucretius 6.176)

(19) cum illaec autumare illum audio
when that.ACC.NEUT.PL say.INF that.ACC.M.SG hear.PRS.1SG
'when I hear that man say those things' (Plautus *Am* 416)

In this instance, however, the diachronic development was not the replacement of the AcI by a finite clause but the fusing of the original two clauses into one through the formation of a complex predicate construction as in (20) and (21):

(20) he fet veure el problema al director
have.PRS.ISG do.PSTPRT see.INF the problem to.the director
'I made the director see the problem' (Catalan)

(21) ho udito uscire Paolo
have.PRS.ISG hear.PSTPRT go out.INF Paolo
'I heard Paolo go out' (Italian)

It is natural to assume that this reanalysis took place before the wholesale decline of the AcI pattern, with the consequence that these structures were not affected by the shift to clauses with overt complementizers described in the previous section.

A change like this, discussed in more detail in Börjars & Vincent 2017: 651-655), fits naturally within a framework like LFG since it is in essence a reorganization at the level of argument and f-structure, and can be handled directly in these terms rather than mediated through c-structure. The starting

point is the representation for these verbs as in (22) and parallel to what we have already seen in (5) for *credere, dicere* and *iurare*:[4]

(22) a. 'facere <SUBJ, COMP/OBJ>'
b. 'audire <SUBJ, COMP/OBJ>'

The change then consists in the arguments of the embedded infinitival verb becoming dissociated from it and attaching instead to the light verb which heads the new complex predicate construction:[5]

(23) a. 'fare-V <SUBJ, OBJ, OBJ_Θ>'
b. 'udire-V <SUBJ, OBJ, OBJ_Θ>'

The new pattern is monoclausal whereas its historical antecedent was biclausal. Diagnostics for this changed state of affairs include first the fact that if the object is cliticised it must attach to the light verb and not to lexical verb of which it is a semantic argument. Thus, the clitic object version of (21) is *l'ho udito uscire* and not **ho uditolo uscire*. Second, if the lexical verb is unergative or unaccusative, the OBJ function of the complex predicate expresses that verb's semantic subject but if the lexical verb is transitive then its subject is forced to assume the OBJ_Θ role, hence *al director* in (20). In addition, monoclausal structures cannot be iterated. Contrast the grammaticality of iterated biclausal causatives in English examples such as *Bill made the director make his assistant answer the letter*.

5. Control verbs

So far, with complementizers we have seen developments that affect c-structure largely in isolation from f-structure while with complex predicate formation the essential shifts affect argument and f-structure, with any changes in syntactic constituency being consequential thereon. In this section, we will examine instead the changes which affect the control verb *velle* 'want', changes which concern both f- and c-structure. Once again the starting point is the AcI as in (24):

(24) volo te uxorem domum ducere
want.PRS.1SG you.ACC wife.ACC home.ACC lead.INF
'I want you to take a wife' (Plautus *Aul* 149)

[4] We use the notation COMP/OBJ to indicate that nothing hangs on the choice between either the COMP-based account of the OBJ one, although the fact that both *facere* 'do' and *audire* 'hear' can also occur with simple nominal objects suggests that the OBJ-based analysis might be preferable.

[5] We use Italian for exemplificatory purposes here but the same would hold for other Romance reflexes of these verbs such as French *faire,* Spanish *hacer, oír* and indeed for cases in which the lexical realization of the light verb component of the construction has changed as with French *entendre* 'hear' (< Latin *intendere* 'stretch, direct attention to') or Portuguese *mandar* 'make' (< Latin *mandare* 'send'). For further discussion of the argument assigning mechanisms involved here see Alsina (1996) and Butt (2010).

At the same time we have also seen that this verb may take a finite CP complement, as in (16). Both these examples involve different subjects in the main and embedded clauses, while in the same subject construction the most common pattern is a simple infinitive as in (25):

(25) potare ego hodie, Euclio, tecum volo
drink.INF I today Euclius you-with want.PRS.1SG
'I want to drink with you today, Euclius' (Plautus *Aul* 569)

In addition, in the words of Jøhndal (2012: 92), 'surprisingly, we also find the AcI under coreference' as in (26) (= his 110), though as he goes on to note examples of this type are less frequent that the more usual plain infinitive in the same subject construction:

(26) volo me placere Philolachi
want.PRS.1SG me.ACC please.INF Philolaches.DAT
'I want to please Philolaches' (Plautus *Mos* 167)

Given what we have seen so far, it is less surprising that the AcI disappears in both its same and different subject variants, leaving a pattern of alternation between a bare infinitive and a finite CP as in the Italian examples in (27):

(27) a. voglio partire domani
want.PRS.1SG leave.INF tomorrow
'I want to leave tomorrow'

b. voglio che tu parta domani
want.PRS.1SG C you.NOM leave.PRS.SUBJ.2SG tomorrow
'I want you to leave tomorrow'

What is less expected is that the same subject variant in (27a) has both a monoclausal and a biclausal version, as can be seen from the alternative positions of the clitic *ci* 'there' in (28):

(28) a. voglio andarci
want.PRS.1SG go.INF-there
'I want to go there'

b. ci voglio andare
there want.PRS.1SG go.INF

(28a) and (28b) are synonymous but (28b) has undergone so-called 'restructuring' to become a single clause as further evidenced by the fact that in the periphastic perfect *volere* requires the auxiliary *essere* 'be' appropriate to *andare* 'go' rather than *avere* 'have' which it requires in isolation: hence *ci sono voluto andare* 'I wanted to go there' and not **ci ho voluto andare* but *ho voluto una birra* 'I wanted a beer'.

How then are we to model the diachronic trajectory here? As we have seen in (5), Jøhndal (2012) proposes to assign the function COMP to AcI across the board, which has the effect of treating both (24) and (26) as instances of anaphoric control. By contrast, he proposes to treat the bare infinitive type (25) as an instance of functional control with the infinitival predicate being assigned the function XCOMP. The loss of the AcI has the effect, then, of creating a clear alternation between different subject/anaphoric control and same subject/functional control. This closely parallels the same diachronic sequence postulated in the account in Börjars & Vincent (2019) of the mechanisms underlying the development of **wil-*, the Germanic cognate of Latin *velle*, into the modern English auxiliary *will* via the 'want' meanings seen in Old English *willan* and modern Swedish *vilja*, with functional control seen as the intermediate stage between anaphoric control (and more particularly 'quasi-anaphoric control' in the sense of Haug 2013) and the PRED-free tense/aspect value of English *will*. More generally, such accounts provide a natural way of modelling the kind of semantic 'bleaching' standardly associated with the process of grammaticalization.

In summary, then, the history of the Latin 'want' verb *velle* and its Romance descendants French *vouloir* and Italian *volere* provides evidence of two distinct diachronic trajectories, one having to do with f-structure and one with c-structure.[6] Such developments can be easily accommodated within a framework like LFG but create an analytical problem for cartographic approaches, where a decision has to be made as to whether to accord an item like *volere* the status of an independent main verb or to treat it as a functional head. Grano (2015: 89) opts for the latter solution: 'Following Cinque (2004), I take the cross-linguistically robust restructuring of *want* as decisive in classifying *want* as a functional head in the inflectional layer of the clause.' The problem then is that as such it cannot govern a CP. To accommodate examples like (27b) Grano is obliged to postulate an intervening silent HAVE as the lexical head of 'want' clauses plus the further assumption that the complement of HAVE is not a CP but a *v*P. This in turn requires him to deny complementizer status to the Italian item *che* in (27b) despite the fact that *che* has a standard complementizing function in clauses dependent on verbs of saying, thinking and the like (Grano 2015: 83, note1). If, on the other hand, he had chosen to classify *want* as a lexical V head, his framework has no obvious way to handle the alternation seen in (28) without a further set of arbitrary assumptions. There is not space here to go into a detailed analysis of these proposals, but it suffices to note the problems that arise within an approach in which one variant has to be given derivational priority over the other, problems that disappear in a model such as LFG where differing c-structures can be mapped onto the same f-structure.

Further evidence of the way c-structure and f-structure may develop independently is to be seen in southern Italian dialects such as Salentino (Calabrese 1993), where we find the patterns in (29):

[6] There are similar patterns to be seen in other Romance languages such as Spanish, Portuguese and Sardinian, but with the additional complication that the relevant lexical item is not a reflex of *velle* but the result of a lexical semantic shift of Latin *quaerere* 'seek' to yield Spanish/Portuguese/Galician *querer* and Sardinian *kerrere.*

(29) a. voggyu lu kattu
want.PRS.1SG it buy. PRS.1SG
b. lu voggyu kattu
it want.PRS.1SG buy. PRS.1SG
c. voggyu ku lu kattu
want.PRS.1SG C it buy. PRS.1SG
I want to buy it'
d. *lu voggyu ku kattu

The difference here is that the complement of the 'want' verb in these dialects is expressed by a finite form rather than the infinitive even with the same subject construction. Restructuring is still possible, however, as evidenced by the equivalence of (29a) and (29b), but if the complementizer *cu* (<Lat *quod*) is present as in (29c) then restructuring is blocked and hence the ungrammaticality of (29d).

Conversely, elsewhere in southern Italy and in Sardinia it is the infinitive which generalises leading to patterns like old Sicilian (30) and modern Sardinian (31):

(30) a. eu nun vi voglu veniri
I NEG there want.PRS.1SG come.INF
'I do not want to come there' (Rinaldi 2005: 152)

b. vulissi homu tu non chi essiri
want.PST.SUBJ.3SG one you.NOM NEG there be.INF
'one would like you not to be there' (Bentley 2014: 99)

(31) a. non kèlio vénnere
NEG want.PRS.1SG come.INF
'I do not want to come'

b. non kèlio a vénnere tue
NEG want.PRS.1SG C come.INF you.NOM
'I do not want you to come' (Jones 1992)

It is to be noted here that in different ways the monoclausal/biclausal distinction is still evident: in (30a) the clitic precedes the 'want' verb while in (31a) there is no complementizer in contrast to the presence of *a* in (31b). Note too that in both (30b) and (31b) the subject of the infinitive is in the nominative, thus marking this out as a Romance development rather than a continuation of the Latin AcI.[7] Thus, once again parallel argument structures map onto different grammatical categories.

[7] Note that the nominative plus infinitive construction here is different from the one that goes by that name in Latin. The latter is simply a passivized variant of the AcI: *Marcus*.NOM *abire*.INF *dicitur*.PRES.PASS.3SG 'Marcus is said to be leaving' (Jøhndal 2012: 61).

6. Prepositional complementizers and split CP

What we have seen in the case of the Latin and Romance 'want' verbs is on the one hand the replacement of the AcI by a finite complement clause when the subjects differ and the continuity of the bare infinitive construction as the only option when the subjects of the main and complement clauses coincide. This, however, is a combination of properties virtually unique to 'want'. The more common situation is the kind of alternation seen in the French examples (32) and (33), where the dependent infinitive is introduced by a marker of its own such as *de* or *à*:

(32) a. J'ai décidé de partir demain
I-have.PRS.1SG decide.PSTPRT DE leave.INF tomorrow
'I have decided to leave tomorrow'

b. J'ai décidé qu'on partira demain
I-have.PRS.1SG decide.PSTPRT C-one leave.FUT.3SG tomorrow
'I have decided that we will leave tomorrow'

(33) Pierre m'a invité à venir demain
Pierre me-have.PRS.3SG invite.PSTPRT A come.INF tomorrow
'Pierre has invited me to come tomorrow'

The issue then is how to model these items. Etymologically there is no doubt that we are dealing with reflexes of the Latin prepositions *ad* 'to' and *de* 'from', which can also be seen in expressions like *à Paris* 'to/in Paris' and *de Londres* 'from/of London'. The complication is that in Latin prepositions do not co-occur with infinitives so the functions exemplified in (32) and (33) are Romance innovations, where they serve as non-finite alternants of items like *que* and hence the label 'prepositional complementizer' which they have acquired in the literature. Within LFG the choice lies between treating them as prepositions that take infinitival complements, thus yielding structures like (34a), or as complementizers as in (34b):

(34) a. [PP [P de] [VP partir demain]], [PP [P à] [VP venir demain]]

b. [CP [C de] [VP partir demain]], [CP [C à] [VP venir demain]]

There are arguments in favour of both. Formal identity might lead one to prefer the prepositional solution, whereas the pattern of finite/non-finute alternation seen in (35) argues in favour of the complementizer account:

(35) a. *avant de partir demain* 'before leaving tomorrow'

b. *avant que tu partes demain* 'before you leave tomorrow'

Alternatively one can seek to import solutions developed within other frameworks. Thus, Abeillé *et al* (2006) introduce the concept of a 'weak head' for precisely these cases, where a weak head is characterised as having the status of a 'prep-word', that is to say the same as a regular preposition,

but it is weak in the sense that it yields its head value to the item with which it co-occurs so that overall the structure is for selection purposes headed by the V and not by the P. In LFG terms, this is very similar to the role played by non-projecting items (see Vincent & Börjars 2019 for further discussion).

A different approach is that adopted within the cartographic model of the left periphery proposed by Rizzi (1997) and alluded to above. Rizzi notes that whereas *que* and the following material in an example like (35b) can be separated by fronted of parenthetical elements, the sequence *de partir* in (35a) can only be separated by verbal clitics as in *d'en partir* 'leave from there' or *d'y aller* 'go to there'. He therefore proposes to break C down into a series of hierarchically organized functional heads, with a finite complementizer like *que* occupying the highest head, labelled Force, while items like *à* and *de* occupy the lowest head, labelled Fin. In other words, in a string like *avant de partir*, there would be a full lexical preposition *avant* 'before' linked to an infinitive by a complementizing particle *de*, which has here lost its etymological status as a preposition, so that the structure is similar to that discussed for the string *perché ka* 'because that' in example (15). In general, LFG has avoided the proliferation of functional heads that is characteristic of the cartographic approach but, as the data from this section and the following one suggest, this may be one instance where the price is worth paying.

7. Recomplementation

The phenomena we have considered so far play to LFG's strengths insofar as they involve patterns of interaction between different levels of structure with no ontological or derivational priority being given to one type of structure above all others. In particular, there is no central role for categorial syntactic representation. We turn our attention now to something which comes with the historical development of complementizers and which at first sight looks to argue strongly for a configurational account, namely complementizer doubling or what in the recent literature has come to be so-called recomplementation. This is the phenomenon whereby complementizers are repeated around a fronted element as in the English examples in (36) - (38):

(36) The party opposite said [**that** if we cut 6 billion from the budget, **that** it would end in a catastrophe] [David Cameron, Prime Minister's Questions in the UK House of Commons]

(37) I'm glad [**that,** whoever talked Strauss into it, **that** they did [Geoff Boycott, BBC Radio 5]

(38) 'Forster once wrote **that** if he had to choose between betraying his friends or his country, **that** he hoped he would have the courage to betray his country.'
[Christopher Catherwood *The Cuckoos' Nest. Five Hundred Years of Cambridge Spies*, Cambridge, Oleander Press, 2013, p.59]

One immediate response to cases such as these might be to dismiss them as errors. Examples like (36) and (37) are drawn from off-the-cuff spoken language — a response to a question in Parliament in the case of (36) and sports commentary in (37) — a genre which inevitably involves hesitations and repetitions that go beyond the bounds of grammar. And occasional written examples such as (38) might be challenged as oversights that more careful proofreading could have eliminated.

However, in Romance at least, the phenomenon is too richly attested in a range of modern and medieval varieties for this to be a convincing escape route. Thus, when discussing the evidence from early Italian dialects, Ledgeway (2005: 3008) observes: 'The examples to be considered, though not so numerous in each single text as to be legitimately considered a core grammatical phenomenon, do however occur in sufficient number and across a wide range of texts from different regions to be interpreted as the reflex of a regular structural phenomenon'. The fact too that our example (40) is from a canonical literary figure like Boccaccio makes it hard to dismiss such cases out of hand. Representative instances then are given here in (39) from modern Portuguese and (40) from old Italian:[8]

(39) Duvido **que** a Ana **que** goste de ópera
doubt.PRS.1SG C to Ana C please.PRS.SUBJ.3SG of opera
'I doubt whether Ana likes opera' (Mascarenhas 2014)

(40) ti priego **che**, se egli avviene
you beg.PRS.1SG C if it happen.PRS.3SG
ch'io muoja **che** le mie cose ed ella
C I die.PRS.SUBJ.1SG C the my things and her
ti sieno raccomandate
you be.PRS.SUBJ.3SG entrusted.PSTPRT
'I beg you that, if I die, my things and her should be entrusted to you'
(Boccaccio *Decameron* 2,7)

In (39) the complementizer *que* appears at the beginning of the complement clause and then again after the fronted topicalized item *a Ana*. In (40) we see a similar pattern, with the difference that the two complementizers appear before respectively the protasis and the apodosis of the embedded conditional sentence. On the usual assumption that the protasis of a conditional sentence is a kind of topic these examples can be made to fit very neatly into Rizzi's extended left periphery with the first occurrence of the complementizer in the Force head and the second in the Fin head and with the topicalized element as specifier of the Topic head:

(41) [$_{\text{Force P}}$ [$_{\text{Force}}$ que] [$_{\text{TopP}}$ [$_{\text{PP}}$ a Ana] [$_{\text{Top}}$ Ø]] [$_{\text{FinP}}$ [$_{\text{Fin}}$ que]]]

[8] Further examples and references to the literature can be found in Wanner (1995), Paoli (2003, 2007) and Munaro (2016). Salvesen & Walkden (2017) also cite examples from Old English.

Indeed, just such an analysis is proposed by Ledgeway (2005) and Villa-Garcia (2015), while Radford (2018) develops a similar account for the spoken English examples (37) and (38).

The challenge for LFG is then to see how this kind of data can be accommodated. One possibility might be to treat this as CP recursion with the fronted element located in the specifier position of the higher C:

(42) [$_{CP}$ [$_{C}$ que] [$_{TopP}$ [$_{PP}$ a Ana] [$_{Top}$ Ø]] [$_{CP}$ [$_{C}$ que]]] …

However, this either implies implausibly that a complementizer can take a CP as its own complement or it is simply a notational variant of the split CP analysis. In this connection, it is instructive that Zipf & Quaglia (2017) propose an LFG-analysis of a different Italian phenomenon which includes what they call a C-structure template (their Figure 7) akin to (42) and then comment in a footnote (p.399, note 6) that this 'is not meant to represent the case of CP recursion but rather two different C-related projections' adding that nonetheless they do not adopt Rizzi's labels but 'prefer remaining neutral to the specific implications of these projections'. It is hard, however, to see what 'remaining neutral' in these circumstances can mean; it would appear that *de facto* if not *de nomine* they have incorporated the concept of split CP into the range of phrase structures permitted within LFG. Nor is there anything inherently implausible about such a conclusion. There is no universally fixed limit to the range of c-structure categories that natural language data require to be recognised if they are to be analysed in proper detail. The question is rather whether a categorial analysis is the right solution for any subset of such data. In the present context it is hard to avoid the conclusion that this is indeed the right solution, and that therefore a c-structure with a split CP will need to be deployed even within a framework such as LFG.

At the same time it is easy to understand the reluctance to go down this route since it can lead to the explosion of functional heads that is characteristic of recent nanosyntactic work (see for example Baunaz 2018 and references there). An alternative therefore would be to follow the idea of Sag & Pollard and abandon the idea of a complementizer as a head and treat it as simply a structural marker that can be inserted as pragmatic circumstances dictate. Such an account would be consistent with the occasional attested instances of complementizer tripling as in the old Neapolitan (43), taken from a letter dated 1353, where the function of the complementizers appears to be to break the text down into rhythmical or rhetorical chunks:[9]

[9] In order to facilitate legibility I have deliberately not glossed this example but hopefully the literal translation plus the complementizers in bold will make the intended structure clear.

(43) Pregove, madama, per l'amor di Dio, **che** de chilli dinare che eo agio vostri **che** si non vi fusse troppo sonço **che** mi 'ndi impristiti una unça.
'I beg you, lady, for the love of God, **that** of that money that I have of yours **that** if it wasn't too much trouble **that** you should lend me some.

For the present we leave open the question as to which these two analytical routes it is preferable to follow. Either way the possibility within LFG of consigning the pragmatic interpretation of the fronted elements to an independent dimension of i-structure (Dalrymple & Nikolaeva 2011) means that the number of slots in the c-structure can be kept to a minimum.

8. Conclusion

In summary, then, what the present paper has sought to do is explore the various ways in which a single Latin construction, the AcI, has been replaced with different structures in different Romance languages and in a range of different syntactic contexts. In the course of the analysis we have seen how the parallel correspondence architecture of LFG, with its separation of a-, f-, c- and i-structure, provides an account which is both more illuminating and theoretically more economical than that available to approaches which mediate all aspects of grammatical structure — and therefore all aspects of change — through a single set of syntactic categories and projections. We have demonstrated that both categories and functions have their own diachronic profiles and that the changes they exhibit over time do not necessarily proceed in synch with each other. At the same time we have raised some questions about the precise nature of such categories and whether for example Rizzi's split CP model needs to be incorporated into LFG, in particular as a way of dealing with the phenomenon of recomplementation. More generally, our work has been inspired by the conviction that any theory or framework needs to be able to accommodate both synchronic and diachronic data and that there is no reason to privilege one over the other if the aim is to understand the mechanisms and processes at work in the organization of human language.

References

Abeillé, Anne, Olivier Bonami, Danièle Godard & Jesse Tseng. 2006. The syntax of French *à* and *de*: an HPSG analysis. In Patrick Saint-Dizier (ed.) *Dimensions of the Syntax and Semantics of Prepositions*. 147-162. Dordrecht: Springer.

Alsina, Alex. 1996. *The Role of Argument Structure in Grammar. Evidence from Romance*. Stanford, Ca: CSLI.

Alsina, Alex, Tara Mohanan & K. P. Mohanan. 2005. How to get rid of the COMP. In Miriam Butt & Tracy Holloway King (eds) *Proceedings of the LFG 2005 Conference,* 83-103. Stanford: CSLI, http://csli-publications.stanford.edu/

Baunaz, Lena. 2017. Decomposing complementizers: the functional sequence of French, Modern Greek, Serbo-Croatian, and Bulgarian

complementizers. In Lena Baunaz, Karen De Clercq, Liliane Haegeman & Eric Lander (eds) *Exploring Naonsyntax.* 149-179. Oxford: Oxford University Press.
Belyaev, Olga, Anastasia Kozhemyakina & Natalia Serdobolskaya. 2017. In defense of COMP: complementation in Moksha Mordvin. In Miriam Butt & Tracy Holloway King (eds) *Proceedings of the LFG 2017 Conference,* 83-103. Stanford: CSLI, http://csli-publications.stanford.edu/
Bennett, Charles E. 1910. *Syntax of Early Latin.* Boston: Allyn and Bacon.
Bentley, Delia. 2014. On the personal infinitive in Sicilian. In Paola Benincà, Adam Ledgeway, Nigel Vincent (eds) *Diachrony and dialects. Grammatical change in the dialects of Italy,* 96-115. Oxford: Oxford University Press.
Börjars, Kersti, Pauline Harries & Nigel Vincent. 2016. Growing syntax: the development of a DP in North Germanic. *Language* 92: e1-e37.
Börjars, Kersti & Nigel Vincent. 2017. Lexical-Functional Grammar. In Adam Ledgeway & Ian Roberts (eds) *The Cambridge Handbook of Historical Syntax*, 642-663. Cambridge: Cambridge University Press.
Börjars, Kersti & Nigel Vincent. 2019. Modelling step change: the history of WILL-verbs in Germanic. In Nuria Yañez-Bouza, Willem Hollmann, Emma Moore & Linda van Bergen (eds) *Categories, Constructions and Change in English Syntax*. 283-314. Cambridge: Cambridge University Press.
Butt, Miriam. 2010. The light verb jungle: still hacking away. In Mengistu Amberber, Brett Baker & Mark Harvey (eds) *Complex predicates. Cross-linguistic perspectives on event structure*. 48-78. Cambridge: Cambridge University Press.
Calabrese, Andrea. 1993. The sentential complementation of Salentino: a study of a language without infinitival clauses. In Adriana Belletti (ed.) *Syntactic Theory and the Dialects of Italy.* 28-98. Turin: Rosenberg & Sellier.
Cinque, Guglielmo. 2004. 'Restructuring' and functional structure. In Adriana Belletti (ed.) *Structures and Beyond: The Cartography of Syntactic Structures.* 132-191. Oxford: Oxford University Press,
Dalrymple, Mary & Helge Lødrup. 2000. The grammatical functions of complement clauses. In Miriam Butt & Tracy Holloway King (eds) *Proceedings of the LFG 2000 Conference.* Stanford: CSLI, http://csli-publications.stanford.edu/
Dalrymple, Mary & Irina Nikolaeva. 2011. *Objects and Information Structure.* Cambridge: Cambridge University Press.
Grano, Thomas. 2015. *Control and Restructuring.* Oxford: Oxford University Press.
Haug, Dag. 2013. Partial control and anaphoric control in LFG. In Miriam Butt & Tracy Holloway King (eds) *Proceedings of the LFG 2013 Conference*, 274-294. Stanford: CSLI, http://csli-publications.stanford.edu/
Jøhndal, Marius L. 2012. *Non-finiteness in Latin.* PhD dissertation, University of Cambridge.
Jones, Michael. 1992. *Sardinian Syntax*. London: Routledge.

Ledgeway, Adam. 2005. Moving through the left periphery: the dual complementiser system in the dialects of southern Italy. *Transactions of the Philological Society* 103: 339-396.
Mascarenhas, Salvador. 2014. Complementizer doubling in European Portuguese. *Rivista di Grammatica Generativa* 36. 105-116.
Munaro, Nicola. 2016. A diachronic approach to complementizer doubling in Italo-Romance and the notion of downward reanalysis. *Rivista di Grammatica Generativa* 38:215-228.
Oniga, Renato. 2014. *Latin: A Linguistic Introduction*. Oxford: Oxford University Press.
Paoli, Sandra. 2003. *COMP and the left periphery: comparative evidence from Romance*. PhD dissertation, Department of Linguistics, University of Manchester.
Paoli, Sandra. 2007. The fine structure of the left periphery: COMPs and subjects. Evidence from Romance. *Lingua* 117:1057-79.
Patejuk, Agnieszka & Przepiórkowski, Adam. 2016. Reducing grammatical functions in LFG. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King & Stefan Müller (eds) *Proceedings of the Joint 2016 Conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*, 541-559. Stanford: CSLI, http://csli-publications.stanford.edu/
Pollard, Carl & Ivan Sag. 1994. *Head-driven phrase structure grammar.* Chicago: University of Chicago Press.
Radford, Andrew. 2018. *Colloquial English. Structure and variation.* Cambridge: Cambridge University Press.
Rinaldi, Gaetana Maria. 2005. *Testi d'archivio del Trecento*. Palermo: Centro di studi filologici e linguistici siciliani, (Collezione di testi siciliani dei secoli XIV e XV, 24-25).
Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.) *Elements of Grammar.* 281-337. Dordrecht: Kluwer.
Sag, Ivan, Gerald Gazdar, Thomas Wasow & Steven Weisler. 1985. Co-ordination and how to distinguish categories. *Natural Language and Linguistic Theory* 3: 117-171.
Salvesen, Chrisine & George Walkden. 2017. Diagnosing embedded V2 in Old English and Old French. In Eric Mathieu & Robert Truswell (eds) *Micro-change and Macro-change in Diachronic Syntax*. 168-181. Oxford: Oxford University Press.
Szűcs, Péter. 2018. A COMP-less approach to Hungarian complement clauses. In Miriam Butt & Tracy Holloway King (eds) *Proceedings of the LFG 2018 Conference,* 325-342. Stanford: CSLI, http://csli-publications.stanford.edu/
Villa-Garcia, Julio. 2015. *The Syntax of Multiple* que *Sentences in Spanish*. Amsterdam: John Benjamins.
Vincent, Nigel & Kersti Börjars. 2010. Grammaticalization and models of language. In Elizabeth Traugott & Graeme Trousdale (eds) *Gradience, gradualness and grammaticalization,* 279-297. Amsterdam: Benjamins.
Vincent, Nigel & Kersti Börjars. 2019. Heads and history. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds) *Syntactic Architecture and its Consequences: Synchronic and Diachronic*

Perspectives. Vol 1: *Syntax inside the Grammar*. 135-159. Berlin: Language Science Press.
Wanner, Dieter. 1995. Les subordonnées à double complémentateur en roman mediéval. In Giovanni Ruffino *et al* (eds) *Atti del XXI Congresso Internazionale di Linguistica e Filologia Romanza, Palermo, 18-24 settembre 1995*. Tübingen: Niemeyer, Vol 1, pp. 421-433.
Zipf, Jessica & Stefano Quaglia. 2017. Asymmetries in Italian matrix *wh*-questions: word order and information structure. In Miriam Butt & Tracy Holloway King (eds) *Proceedings of the LFG 2017 Conference,* 387-405. Stanford: CSLI, http://csli-publications.stanford.edu/