

Not All Complementisers Are Late: a first look at the acquisition of illocutionary complementisers in Catalan and Spanish*

NÚRIA BOSCH

UNIVERSITY OF CAMBRIDGE

ABSTRACT This paper analyses the emergence of illocutionary complementisers (in the sense of [Corr 2016](#)) through a corpus study with Catalan and Spanish children. The production of illocutionary complementisers by ten Catalan- and Spanish-speaking children in the CHILDES database is quantified and compared to the production of finite embedding complementisers. The results indicate that illocutionary complementisers emerge early in the child production data, often well before embedding complementisers first appear. These preliminary findings, which illustrate important developmental differences between kinds of complementisers, are hard to account for in approaches that take functional categories to mature bottom-up, with left-peripheral knowledge developing last. I argue, instead, that the early emergence of illocutionary complementisers favours a view which takes the C-domain to be present early on in child grammars.

1 TWO COMPLEMENTISERS, TWO ACQUISITION TIMINGS

The acquisition of complementisers and subordination is typically taken to be a relatively late phenomenon, in comparison to other very early phenomena, such as the acquisition of word-order and head-directionality ([Tsimpli 2014](#)). The earliest forms of subordination include so-called preconjunctionals, namely subordinate clauses which lack the target-language complementiser. The emergence of complementisers, such as Catalan and Spanish subordinator and relativiser *que*, is a later development during the period of early word-combinations (e.g. see [Armon-Lotem 2005](#), [Clahsen & Penke 1992](#) for discussion on Hebrew and German). Converging with these observations, theoretical approaches have frequently advocated for initial unavailability of the CP domain; consider [Radford's \(1988\)](#) Small Clause Hypothesis, [Rizzi's \(1993/1994\)](#) Truncation Hypothesis and [Friedmann, Belletti & Rizzi's \(2021\)](#) Growing Trees Hypothesis. The latter of these, similarly to Radford, proposes a

* Immense thanks go to Theresa Biberauer for supervising my MPhil thesis and the small side-project from which this paper stems. All her encouragement, patience and support (both academic and pastoral) has been much appreciated throughout my time at Cambridge. Any remaining errors or inaccuracies are my own. This research was generously supported by a Benefactors' Scholarship (Hayes Scholarship) from St John's College and a Cambridge Opportunity Master's Studentship from the Cambridge Trust.

more general maturational mechanism whereby functional spines develop bottom-up. These theoretical approaches are argued to help account for, i.a. the timings regarding the development of complementisers.

However, another kind of main-clause complementiser attested across Ibero-Romance and used prevalently in spoken contexts has gone unstudied in the context of the emergence of complementisers. These are so-called illocutionary complementisers; that is, complementisers that do not function as a subordinator that heads a complement or relative clause in Ibero-Romance and instead introduce non-embedded matrix clauses, with several illocutionary functions. A canonical example of a complementiser which can introduce complement clauses or relative clauses is illustrated in (1), while examples of illocutionary complementisers heading a main clause are provided in (2):¹

(1) Subordinating complementisers

- a. *Li he dit **que** aquesta tarda vaig a* [Catalan]
 CL.IO= AUX.1SG told that this afternoon go.1SG to
Barcelona a veure un concert al Liceu
 Barcelona to see a concert in.the Liceu
 ‘I have told him/her that I’m going to Barcelona this afternoon to see a concert in El Liceu.’
- b. *No podia creer **que** hubiesen ganado* [Spanish]
 not can.IMPF.3SG believe that AUX.SUBJ.IMPF.3PL won
la lotería
 the lottery
 ‘He/she couldn’t believe that they’d won the lottery.’

(2) Illocutionary complementisers

- a. *Ai, **que** t’atrapo!* [Catalan]
 hey that.EXCL CL.DO=catch.1SG
 ‘I’m coming to get you!’ (Corr 2016: 88)
- b. *No hagas esto, **que** luego mamá se* [Spanish]
 not do.SUBJ.2SG this that.CONJ then mum CL.REFL=
enfada
 get.angry.3SG
 ‘Don’t do this, because then mum gets angry.’

Illocutionary complementisers therefore represent a new and potentially productive testbed for the claim that CP-material and, particularly, complementisers emerge

¹ Unless otherwise noted, the Catalan and Spanish examples provided throughout this paper are the author’s own and reflect the dialects of Central Catalan and Peninsular Spanish.

late in the learning path. The present paper shows that this use of *illocutionary* complementisers emerges substantially earlier than subordinating complementisers, instantiating novel evidence for early access to a CP domain in child grammars.

The relative emergence of this kind of complementiser compared to prototypical embedding complementisers could have rather far-reaching consequences for approaches to the acquisition of functional categories. Early uses of illocutionary complementisers will raise several questions. For example, how can any timing differences between these two complementiser types be explained? And, looking at the bigger picture, how does early emergence of illocutionary complementisers square with contemporary approaches to the development of functional categories, particularly those that assume a ‘delayed’ maturation of the CP domain?

The development of so-called illocutionary complementisers in children acquiring Ibero-Romance languages is yet to be systematically studied and contrasted to the acquisition of subordinating complementisers. In this paper, I present a preliminary corpus study on child Catalan and Spanish complementisers, showing that at least some illocutionary complementisers emerge on average earlier than subordinating complementisers and often from the earliest files. I argue, therefore, that illocutionary complementisers represent a potentially strong case study in favour of early availability of discourse and speaker/hearer-oriented material. Such a finding is at odds with many maturational approaches that assume functional categories mature bottom-up, with CP material developing at the very end of the learning path. On the other hand, I suggest that these discrepancies between complementiser types are to be expected due to (at least) two factors: firstly, given the main clause vs. embedded clause nature of the use of these two kinds of complementisers, illocutionary complementisers require less syntactically elaborate grammars (no knowledge of embedding) and thus should, in this respect, be accessible earlier. I contend, as well, that these patterns are coherent with approaches to acquisition and language variation that take discourse and interactional content to be acquisitionally privileged and salient and that take functional spines to initially develop ‘inwardly’, such as in Heim & Wiltschko’s (2021) Inward Growing Spine Hypothesis (cf. also Wiltschko’s, 2021, and Hinzen & Wiltschko’s, 2023, Bridge Model), Biberauer’s 2018 Hypothesis and Biberauer & Roberts’s (2015) emergent categorial hierarchy (see also Bosch *in progress* for empirical support and further theoretical expansion).

This paper is organised as follows: [section 2](#) introduces illocutionary complementisers in Ibero-Romance and some of their syntactic properties, focusing on Catalan and Spanish data. In [section 3](#), I present the theoretical background and hypothesis and also provide an overview of approaches arguing for early acquisition of CP-based and speaker-hearer material. [Section 4](#) outlines the corpus study, its methodology and, subsequently, its results and [section 5](#) offers a discussion of the data’s preliminary theoretical implications. [Section 6](#) concludes.

2 ILLOCUTIONARY COMPLEMENTISERS IN IBERO-ROMANCE: TYPOLOGY AND SYNTACTIC PROPERTIES

Complementisers are typically assumed to involve subordinating conjunctions that transform clauses into complements of a matrix clause (Rosenbaum 1967, Lakoff 1968, Bresnan 1972). In Spanish and Catalan, finite complement clauses are often introduced by the complementiser *que*, as exemplified below:

- (3) a. *M'adono* ***que*** *encara no ha arribat* [Catalan]
 CL=realise.1SG that yet not AUX.3SG arrive.PART
 'I realise that he/she hasn't arrived yet.'
- b. *Me dijiste* ***que*** *se había arrepentido* [Spanish]
 CL.IO= tell.PST.2SG that CL.REFL= AUX.IMPF.3SG regret.PART
 'You told me that he/she had regretted (it).'

A conspicuous property of the Ibero-Romance language family, particularly varieties in the Iberian Peninsula, however, is their pervasive use of what seems to be the complementiser *que* introducing matrix clauses; namely, the complementiser in Ibero-Romance does not always function as a prototypical subordinator that introduces an embedded clause. These uses of *que* are attested across Ibero-Romance (though with varying degrees of frequency and robustness) and they come with a range of illocutionary functions. Diachronically, these illocutionary complementisers arise from main-clause constructions where the core function of *que* as a subordinator (e.g. as in 1 and 3) has been bleached in favour of utterance- and speaker-hearer-oriented functions, which are often emphatic in nature. Illocutionary *que* provides evidence of a grammatical device – a semantically-bleached subordinating complementiser – whose morpholexical material has been recycled for the purpose of encoding illocutionary information in the main clause (Corr 2016: 1-3).

Corr (2016) distinguishes at least three separate kinds of illocutionary complementisers based on both their distributional and interpretive properties. We will adopt Corr's classification and, thus, distinguish between EXCLAMATIVE, QUOTATIVE and CONJUNCTIVE *que*. These conversation- and speech-act-oriented uses of *que* will be collectively referred to as ILLOCUTIONARY complementisers. Given the child data to be considered, attention will focus on Catalan and Spanish.

Exclamative *que* introduces an indicative clause (typically, though not always, declarative) and has the illocutionary force of an exclamation:

- (4) a. *Ala, que ho has llençat tot* [Catalan]
 hey that.EXCL CL.DO= AUX.2SG throw.PART everything
al terra!
 on.the floor
 'Hey! You've thrown everything on the floor!'

- b. *¡Que la abuela Rosa me ha* [Spanish]
 that.EXCL the grandmother Rosa CL.IO= AUX.3SG
comprado una Nintendo nueva!
 buy.PART a Nintendo new

‘Grandma Rosa has bought me a new Nintendo!’

Quotative *que* constructions involve reported speech clauses introduced by *que*. These, importantly, do not require a retrievable *verbum dicendi* for the utterance to be felicitous:

- (5) Context: the speaker is asked who had just phoned.

- a. *Era la Carme. Que em trucava* [Catalan]
 was the Carme that.QUOT CL.DO= phone.IMPF.3SG
per felicitar-me
 to congratulate=CL.DO

‘It was Carme. She phoned me to wish me a happy birthday.’

- (6) a. *A: No entiendo qué acabas de decir* [Spanish]
 not understand.1SG what finish.2SG to say.INF
 ‘I don’t understand what you’ve just said.’

- b. *B: ¿Qué?*
 what

‘What?’

- c. *A: Que no entiendo qué dices*
 that.QUOT not understand.1SG what say.2SG

‘(I’ve said that) I don’t understand what you’re saying.’

Quotative *que* can also introduce a report of a recent utterance made by the addressee, often requesting for confirmation:

- (7) a. *A: M’acompanyes a Vic?* [Catalan]
 CL.DO=accompany.2SG to Vic

‘Are you coming to Vic with me?’

- b. *B: Que si t’acompanyo a Vic?*
 that.QUOT if CL.DO=accompany.1SG to Vic

‘You’re asking if I’m coming to Vic with you?’

- c. *A: Sí!*
 yes

‘Yes!’

- d. B: *D'acord, ja vinc!*
 okay already come.1SG
 'Okay, I'll come!'

A final third kind of illocutionary *que* involves conjunctive uses of *que*, where its function involves contextualising utterance information for the addressee. It is typically associated with causality, usually linking two root clauses in a speech act causal relation. Often, though not necessarily, the first root clause is an imperative clause as in (8).²

- (8) a. *No li diguis això a la Paula que és un* [Catalan]
 not CL.IO= tell.SUBJ.2SG this to the Paula that.CONJ is a
secret
 secret
 'Don't tell this to Paula because it's a secret.'
- b. *Dame el diario, por favor, que siempre* [Spanish]
 give=CL.IO the newspaper for favour that.CONJ always
me gusta repasarlo antes de ir a dormir
 CL.IO= like go.OVER=CL.DO before of go.INF to sleep.INF
 'Give me the newspaper, please. I always like going over it before going to bed.'

The present study will also consider instances of polar interrogatives optionally headed by the complementiser *que* in Catalan, which are available in Standard Eastern Ibero-Romance. They are, thus, unavailable in Spanish:³

- (9) *Que vindràs al final a veure la pel·lícula?* [Catalan]
 that.INT come.FUT.2SG in.the end to watch.INF the film
 'Are you coming to watch the film in the end?'

The interpretative value of this INTERROGATIVE *que* varies dialectally. In Balearic, Central and north-western Catalan, interrogative *que* introduces a neutral polar question, whereas it can only introduce counter-expectation questions in north-central Catalan, Rossellonese and Valencian (Prieto & Rigau 2007). All Catalan data to be discussed here, however, is taken from five children growing up in Catalonia, primarily in Barcelona; they therefore speak a Central Catalan variety.

Main-clause instances of illocutionary *que* such as those illustrated in (2), whilst discourse-oriented, are not easily catalogued under the umbrella of information

² The distribution of conjunctive complementisers is not limited to Ibero-Romance, but they are also observed across Romance (see Corr 2016: 226-227 and references therein for data from Romanian, Southern Italian Dialects and Rhaeto-Romance and Prins 2014 for Southern Italian Dialects specifically).

³ Balearic Catalan also allows the conjunction *o* 'or' to head this structure (e.g. *O vindran a Ciutadella?*, 'are they coming to Ciutadella?'; Prieto & Rigau 2007: 1). These instances are not relevant here.

structure that traditionally characterises the CP domain. Rather, illocutionary uses are primarily attested in spontaneous speech, charting the conversational dynamics between interlocutors, and operative in here-and-now contexts (Corr 2016). Although linguistic elements involved in the encoding of these kinds of conversational pragmatics were until relatively recently thought to lack formal syntactic properties, I will adopt neo-performative proposals (building on Ross 1970) that take such utterance- and conversation-related information to be formally represented in the syntax. These works embed illocutionary complementisers as part of the C-domain or of a higher Speech-Act-oriented domain (following, i.a. Speas & Tenny 2003, Giorgi 2010, Haegeman & Hill 2013, Wiltschko 2014, 2021, Corr 2016, 2022, Miyagawa 2022).

Therefore, we will broadly (and simplifying grossly) assume that some structural organisation like the one in Figure 1 is largely on the right track for natural language, though I will remain agnostic as to exactly what functional projections (cartographic or otherwise) populate each of these domains. Although Figure 1 contains a speaker-hearer domain at the outer vP edge (cf. also Belletti 2004 on the ‘low IP area’), this paper will only discuss the higher SAP domain directly above the traditional CP domain.

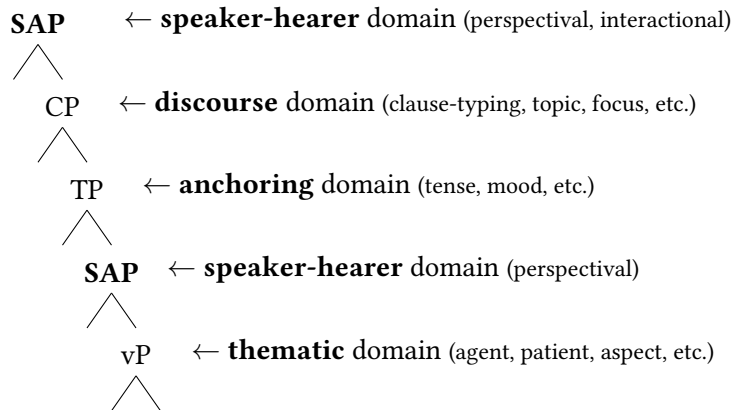
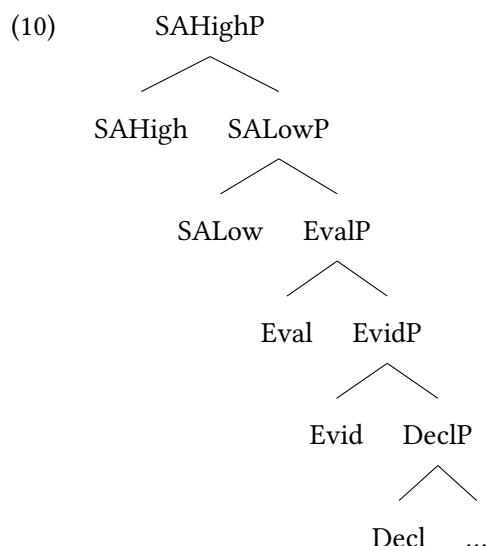


Figure 1 Clausal structure with speech-act layers (Biberauer 2018: 4).

Specifically, Corr (2016) proposes that the three types of ‘illocutionary’ *que* outlined above each have distinct distributional and interpretational properties (e.g. with respect to interpolation, VP deletion, variable binding, among other diagnostics). On the basis of these discrepancies, they are argued to correspond to separate functional projections at the height of the cartographic left-periphery and the so-called *Utterance Phrase* (or UP), a dedicated utterance domain above the CP (analogous to the higher SAP in Figure 1). Illocutionary complementisers are hypothesised to be ‘scattered’ across both the speech-act (UP) and discourse domain (CP). It is proposed that exclamative and conjunctive complementisers pattern differently to quotative complementisers, with the former being located in a higher, speech-act-oriented layer (in SALowP for exclamative *que* and SAHighP for conjunctive *que*), while quotative complementisers behave like a C-head (interpreted in Corr 2016 as a C-head

that heads a dedicated Evidential projection). These complementisers' properties and their associated projections are described in greater depth below. Traditional embedding complementisers, in contrast, are assumed to be located in Rizzi's (1997) ForceP or in Corr's (2016) DeclarativeP below (as the head encoding declarative clause-typing; see also Ledgeway 2012: 179).⁴ Interrogative complementisers have also been suggested to be C-based (e.g. Prieto & Rigau 2007 place interrogative *que* in Rizzi's, 1997, Fin head).

Corr's (2016) cartographic hierarchy is illustrated in (10). SAHighP and SALowP correspond to functional layers within the UP domain, while EvaluativeP, EvidentialP and DeclarativeP involve functional projections at the UP/CP border, encoding speaker belief in and evidence for the proposition, and declarative clause-typing, respectively (Corr 2016: 232). These are commonly collapsible into the Rizzian ForceP (Rizzi 1997).



A range of distributional diagnostics indicate that both conjunctive and exclamative *que* operate at the speech-act domain (Corr's SAHighP and SALowP) rather than the cartographic C-space (from EvalP and projections below). For instance, exclamative *que* constructions can be formed from and precede wh-exclamatives and wh-interrogatives (e.g. Sp. *¡Que qué diablos te pasa, mocosa inmadura!* 'What the hell's wrong with you, you snivelly kid!'; Corr 2016: 130). As a result, exclamative *que* must surface higher vis-à-vis the dedicated projections for wh-exclamatives and wh-interrogatives. Furthermore, both conjunctive and exclamative *que* cannot be readily embedded, yet embedding quotative complements is possible and in fact highly productive (cf. 11b and 12a). This 'embeddability' property is to be expected from CP-based material, but not from speech-act-located constituents — embedding speech acts, although a logical possibility, is in reality very rare (see

⁴ Other instances of non-illocutionary *que* in Ibero-Romance, such as recomplementation *que* and jussive *que* are, in some approaches, hosted by Topic and Fin, respectively (Villa-García 2012, *et seq.*).

Krifka 2014). Such a mismatch motivates the inclusion of quotative *que* as part of the CP-domain (specifically, in EvidP and above DeclP, the position for embedding complementisers) and quotative/exclamative *que* as part of a higher UP domain.

(11) Incompatibility of embedding with exclamative and conjunctive *que*

- a. **Vaig dir que que en Mario va* [Catalan]
 AUX.PST.1SG say.INF that that.EXCL the Mario AUX.PST
callar a la fi!
 shut.up.INF in the end
 ‘I said that Mario shut up at last!’ (Corr 2016: 231)
- b. *Ets estúpid? *Vaig dir (.) que* [Catalan]
 be.2SG stupid AUX.PST.1SG say.INF that.CONJ
m’estàs trepitjant!
 CL.DO=be.3SG stepping
 ‘Are you an idiot? (I said that you’re stepping on me!).’ (Corr 2016: 231)

(12) Quotative complements can be embedded

- a. *Te he preguntado que quién es aquel* [Spanish]
 CL.IO= AUX.1SG ask.PART that.QUOT who is that
nuevo estudiante allá
 new student there
 ‘I’ve asked you who that new student there is.’

Other diagnostics further support this conclusion. According to Krifka (2001), the logical operation of disjunction is only interpretable at the propositional level. This is because a sentence with two disjoint clauses cannot be paraphrased as co-ordinating two acts of assertion, thus cancelling the illocutionary force of the sentence. This predicts disjunction to be licit with propositional (thus CP-internal) elements, but illicit with SAP-based material. In this context, both conjunctive *que* and exclamative *que* cannot be disjoined, unlike quotative *que*, which does permit it (cf. 13 and 14). The former thus pattern alike, suggesting they operate at the level of the speech act; again, to the exclusion of quotative *que*.

(13) Exclamative and conjunctive *que* cannot be disjoined

- a. **Oh, que he menjat massa o que* [Catalan]
 oh that.EXCL AUX.1SG eat.PART too-much or that.EXCL
em fa mal el cap!
 CL.IO= do.3SG pain the head
 ‘Oh, I’ve eaten too much, or my head hurts!’ (Corr 2016: 230)

- b. *No me pises, que llevo chanclas #o* [Spanish]
 not CL.DO= step.SUBJ.2SG that.CONJ wear.1SG flip-flops or
que soy sensible
 that.CONJ am sensitive
 ‘Don’t step on me, I’m wearing flip-flops (or I’m sensitive).’ (Corr 2016: 230)

(14) Quotative *que* allows for disjunction

- a. *Y él que llegábamos tarde, que no* [Spanish]
 and he that.QUOT arrive.IMPF.1PL late that.QUOT not
se podía salir con nosotros o que
 CL.IMP= can.IMPF.3SG go.out.INF with us or that.QUOT
teníamos que protestar por el retraso
 have.IMPF.3SG that complain.INF for the delay
 ‘And he kept saying that we were late, that he couldn’t go out with us or that we should complain about the delay’. (Corr 2016: 167)

These observations, among other diagnostics (for which see Corr 2016), single out conjunctive and exclamative *que* in Ibero-Romance as a class displaying syntactic behaviour distinct from complement clauses headed by the finite complementiser *que* in C and also from quotative *que*. Among these UP-based complementisers (and simplifying slightly), exclamative *que* is argued to lexicalise SALow, as it precedes topics in the CP but follows vocatives higher in the UP. Conjunctive *que*, on the other hand, is suggested to be localised in SAHigh, on the grounds that utterance-oriented items to the left of conjunctive *que* are banned and, thus, conjunctive *que* must occupy a higher head in the UP, above SALowP (see Corr 2016 for further discussion on their properties). In contrast, quotative *que* displays the formal behaviour of a prototypical C-head, introducing an embeddable sentence (and, in Corr’s system, lexicalising Evid, a functional projection at the UP/CP boundary). Although the precise location of these complementisers within the left periphery will not be at the forefront of our discussion, it is worth emphasising their placement as part of a high discourse (CP) and interactional domain (SAP or, in Corr’s terms, UP): quotative *que* resides at the edge of CP (which broadly coincides with that of embedding complementisers, also assumed to be located high in the C-domain) and exclamative and conjunctive *que* in the higher SAP domain. This syntactic property will become important in light of hypotheses that suggest comparatively late development of the highest layers of the clausal spine.

There are other structures in Catalan and Spanish that are somewhat reminiscent of the ones just introduced. Cruschina & Remberger (2018) discuss cases of complementisers being preceded by a functional element which morphologically coincides with an adjective or adverb (see 15). They suggest that the functional elements are directly merged in a speech-act-related phrase above ForceP (labelled

SentenceP, from [Speas & Tenny 2003](#)), while the main-clause complementisers are located in Force. These constructions, attested across Romance, are also relevant for our purposes and will be analysed in the corpus study introduced in the next section.

- (15) a. *Evidentment **que** estudiaré per a aquest examen!* [Catalan]
 obviously that study.FUT.1SG for to this exam
 ‘I’ll obviously study for this exam!’
- b. *¡Claro **que** entendió!* [Spanish]
 clear that understand.PST.3SG
 ‘Of course he/she understood!’

Illocutionary complementisers will be contrasted in the corpus study in [section 4](#) to traditional EMBEDDING or SUBORDINATING complementisers in Catalan and Spanish. As described earlier, these introduce, for instance, relative clauses ([16a](#)) and complement clauses ([16b](#)):

- (16) a. *La Maria, **que** estava al davant de l’atri,* [Catalan]
 the Mary who was in.the front of the-atrium
va anunciar els guardonats al premi literari
 AUX.PST.3SG announce.INF the winners to.the prize literary
 ‘Mary, who was standing in front of the atrium, announced the winners of the literary prize.’
- b. *Ayer nos enteramos de **que** van a* [Spanish]
 yesterday CL.REFL= find.out.PST.1PL of that go.3SG to
ser padres
 be.INF parents
 ‘Yesterday we found out that they are expecting a child.’

The key assumption which we will adopt from the works just introduced, then, is that illocutionary complementisers are hosted in the discourse and interactional domains (e.g. the SAP domain and CP domain in [Figure 1](#)) – a theoretical analysis we will subsequently harness to test the predictive differences among current approaches to the acquisition of functional categories (see [section 3](#)). However, as briefly noted above, the exact hypothesised location of these complementisers in the left periphery and its empirical adequacy is going to be largely orthogonal to the aims of this paper.

Having introduced these two broad kinds of complementisers and their properties, I now turn to discussing the theoretical background of the corpus study and its hypothesis in the next section.

3 THEORETICAL BACKGROUND AND HYPOTHESIS

An analysis of the acquisition of illocutionary complementisers is significant for a range of current proposals regarding the acquisition of functional categories.

A prevalent hypothesis, most recently defended in [Friedmann et al. \(2021\)](#) and [Friedmann & Reznick \(2021\)](#) under the name of the Growing Trees Hypothesis, suggests that functional categories are acquired following an innately-determined ordering of development, which specifies that the cartographic hierarchy in Universal Grammar develops bottom-up ([Figure 2](#)). The proposal finds its roots in earlier hypotheses such as [Radford's \(1988\)](#) Small Clause Hypothesis and [Rizzi's \(1993/1994\)](#) Truncation Hypothesis.

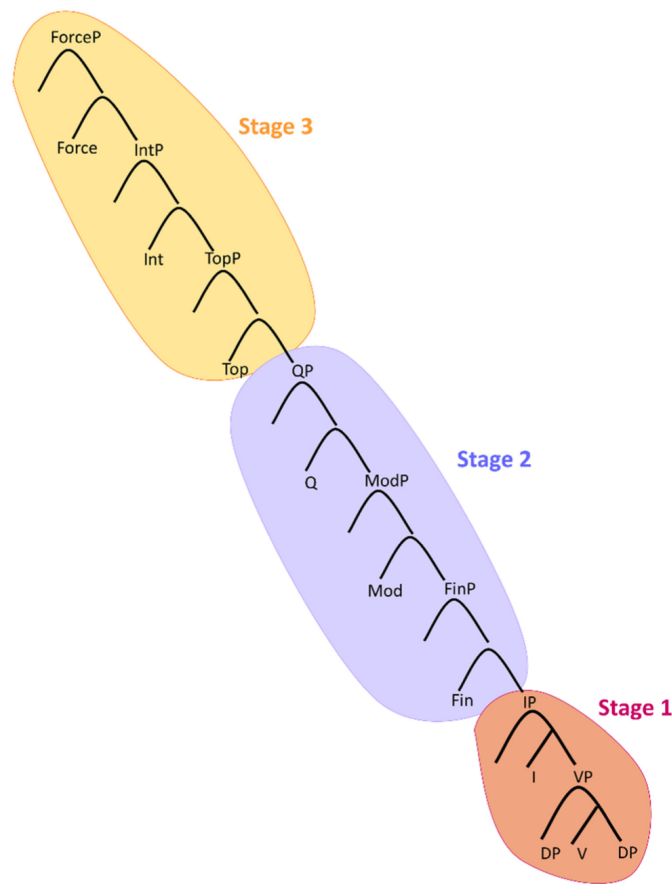


Figure 2 Stages of acquisition of the clausal domain in the Growing Trees Hypothesis ([Friedmann et al. 2021: 12](#)).

These approaches would predict delayed development of the left periphery compared to lower clausal domains such as VP and TP. The predicted relative development of embedding complementisers vs. illocutionary ones is unclear as [Friedmann et al. \(2021\)](#) do not discuss interactional language *per se*, but, in any event, the predictions would either expect both complementisers to become available at similar times

(that is, after TP and VP knowledge have emerged) or, possibly, that some illocutionary complementisers would emerge later than embedding complementisers, by virtue of being located in a higher left-peripheral domain (SAP/UP for exclamative and conjunctive *que*). As illocutionary complementisers are located at the very top of the left periphery, neither prediction would therefore expect (some) illocutionary complementisers to start emerging well before embedding complementisers. In other words, the development of illocutionary and embedding complementisers is predicted to be closely linked, by virtue of bottom-up maturation and their similar hierarchical placement (for comparable approaches, see [Vainikka & Young-Scholten 2011](#), [Diercks & Bossi 2021](#), [Diercks, Johnson, Bar-Sever & Bossi 2023](#)).

However, whilst bottom-up maturation has arguably been the dominant perspective in the acquisition literature, it is not obvious from other work in language acquisition and variation that expressive, illocutionary and discourse content should necessarily be developmentally ‘delayed’ in comparison to material located at other layers of the clausal domain such as vP or TP. In other words, it is not self-evident that making developmental predictions primarily on the basis of structural height (in the adult or UG-given grammar) should be the preferred mode of reasoning. There may be independent reasons for certain structural elements (whether located high in the clausal domain or not) to emerge early, particularly in emergentist approaches that do not assume a UG-given functional spine ([Biberauer & Roberts 2015](#), [Ramchand & Svenonius 2014](#), [Biberauer 2019](#), [Ramchand 2023](#)). For one, [Roeper \(2007, et seq.\)](#) already noted the importance of a ‘parallel’ language possibly being in play in early word combinations alongside simple referential forms — that is, a language that reflects primarily attitudes and feelings, besides propositional and referential language. A host of other terms, denoting seemingly ‘vague’ but nonetheless highly expressive meaning (such as *uh-oh*, *well*, *huh*, *oops*) abound in early utterances. As [Roeper](#) notes (p. 40), these expressive words tend to occur before and after sentences, sometimes even in co-occurring constructions (e.g. *Well, gee, yes, maybe I can*, in response to *Can you sled down that hill?*). Albeit remarkably hard to characterise semantically, they are among the first words a child uses and, he suggests, ‘may give the biggest clue to the quality of children’s thoughts’ ([Roeper 2007: 39](#)). Joint work with Christopher Potts also underscored the possibility that early child combinations might primarily denote purely expressive meaning, besides any conceivable propositional content ([Potts & Roeper 2006](#)).

Proceeding with a convergent line of reasoning, there is a small, but nonetheless growing, range of theoretical approaches to acquisition that argue for an acquisitionally privileged role of some speaker-hearer and discourse content, not just lexically, but also syntactically — the approaches introduced below, then, do anticipate that at least some illocutionary complementisers might emerge early on. This, notably, will underscore a predictive difference of these approaches compared to bottom-up maturation: such a prediction is *ipso facto* ruled out in a bottom-up approach. Since these discourse/perspectival elements are located at the CP or SAP domain, they should emerge late and closely approximate the emergence of subordinating complementisers. It is worth emphasising head-on, however, that neither [Roeper \(2007\)](#) nor any of the approaches outlined in what follows necessarily predict that

all relevant discourse and interactional material in a language will emerge early and/or simultaneously. Importantly, in this context, they expect early production and acquisition of *some* illocutionary complementisers, but critically without any presuppositions that the *complete* inventory of illocutionary complementisers (and other discourse elements) should emerge in one fell swoop.

Roeper & Rohrbacher (1994) and Galasso (2003), for instance, argue that the production data in English shows evidence of early knowledge of CP, in particular wh-questions; viz. Galasso's 'Empty Middle' approach, whereby the earliest stages project CP>VP before developing into CP>IP>VP or richer elaborations thereof.⁵ Tsimpli (2005) similarly provides data from child Greek that supports early acquisition of the left-periphery based on the difference between LF-interpretable and LF-uninterpretable features: LF-interpretable features are taken to be responsible for a distinction between (i) discourse-related features (e.g. focusing and topicalisation/dislocation), and (ii) between peripheral LF-interpretable features (focusing or wh-questioning) on the one hand, and non-peripheral, uninterpretable features such as inflectional features, on the other (Tsimpli 2005: 185-186). In the two children studied, phenomena such as focalisation and wh-questioning (involving LF-interpretable features) emerge early on. This is followed by Clitic-Left Dislocation and clitic doubling (involving discourse-related and uninterpretable features). The inflectional domain (which requires non-peripheral uninterpretable features) only clearly emerges subsequently. These results are taken to suggest that 'peripheral' positions are grammatically represented from the earliest stages.

Other works have also underscored the early emergence of some CP-based knowledge. van Kampen (2010), notably, has argued for a 'typological guidance' approach to acquisition, whereby the major typological properties of a language are invariably the first to be acquired and are the 'system's bootstrap for learnability' (van Kampen 2010: 264). Based on Dutch data, she makes the case for early acquisition of the V2 rule, with early topic-comment structures and operator-comment structures with [+FINITE] verbs displaying surface V2/V1 order (e.g. *moet_{operator} papa_{topic} doen_{comment}* 'daddy has to do (it)'; Sarah week 120) setting the stage for the full acquisition of V-to-C shortly after. Additionally, the acquisition data suggests that finiteness, not tense, is likely the trigger for verb movement to the C-domain, as tense and agreement markings are established after verb movement is already apparent (see also van Kampen 2009a, 2009b, on early topic-comment structures and the acquisition of V2 and wh-questions). In this approach, the early acquisition of some CP-related knowledge is a consequence of major typological characteristics of the child's L1 guiding the learner, and, in Dutch, some of these are encoded precisely in the C-domain (as is the case in other V2 and V1 languages). As these approaches indicate, then, the possibility that some discourse knowledge emerges early representationally, potentially even before TP-based material, is not unlikely.

⁵ See also Perkins & Lidz (2021) and Perkins, Ying, Williams & Lidz (2021) for behavioural evidence that 18- and 20-month-olds can represent abstract non-local dependencies like wh-questions and, according to preliminary results, that they can subsequently harness them to constrain hypotheses during verb learning. Thanks to Laurel Perkins for providing me with a copy of the slides of the latter reference.

More generally, however, they argue, in different forms, against a strictly bottom-up approach to development. To the extent that illocutionary complementisers are located at either the CP or SAP domain, these results could anticipate a potential early emergence of illocutionary complementisers.

In a more recent proposal, Heim & Wilschko (2021) and Wilschko (2021) propose an inward maturational pathway of the universal and interactional spine in Wilschko (2014, 2021), arguing that the interactional and classification-based layers (the latter corresponding, roughly, to vP) mature first. This hypothesis rests on two interrelated models of clausal organisation, namely the Universal Spine Hypothesis and the Interactional Spine Hypothesis, which we introduce in turn below. These are jointly displayed in Figure 3.

Firstly, Wilschko (2014) argues for a universal spine of functional categories, which are variably realised by language-specific content. Each layer is linked to a function, essential for the configuration of reference and propositional meaning and with each higher layer presupposing the existence of lower-level projections. These unfold as follows: the first functional layer, *Classification*, ‘catalogues’ events and individuals into subcategories (e.g. telic vs. atelic events; mass vs. count nouns). *Point of View* (PoV) relates the previously classified event or individual to a particular perspective or viewpoint (e.g. in familiar Indo-European languages, PoV is substantiated by temporality as outer aspect⁶). *Anchoring* maps the event or individual to the utterance context (e.g. commonly via tense or, in other languages, person or location; see Ritter & Wilschko 2014). Finally, *Linking* serves to map the anchored event or individual to the discourse (hosting, among others, wh-words, topics and other elements typically associated with the traditional CP domain). Broadly, then, the Universal Spine Hypothesis argues for a *substantivist* view in which grammatical categories are constructed on a language-specific basis (i.e. the substantive content of categories is crosslinguistically non-universal), but, all the same, the spine restricts the types of categories that languages construct and the hierarchical ordering they display.

Most recently, and incorporating the insights from the neo-performative insights from Ross (1970) and subsequent work, this universal spine has been expanded in Wilschko (2021) to incorporate a syntactic encoding of speech-act information. The interactional spine framework explores the grammatical underpinnings of interactional language and proposes to extend the universal spine by including two more syntactic domains, which syntactically dominate propositional structure (i.e. the Linking, Anchoring and Classification domains; approximately, from CP to vP in Figure 1). Within this conversational domain, a lower *Grounding* or *Grouping* layer, divisible into Ground-Speaker and Ground-Addressee, and a *Responding* layer are postulated. The core function of the Grounding layer is to enable ‘the speaker to configure the propositional content of the utterance so that the addressee can update their knowledge state to include it’ (Wilschko 2021: 72). Lastly, the Response layer serves the function of managing ‘the moves that serve to synchronize the interlocutors’ knowledge states’ (Wilschko 2021: 72), mapping the interactive

⁶ cf. so-called inner aspect, which denotes lexical aspect or *Aktionsart* (see Travis 2010).

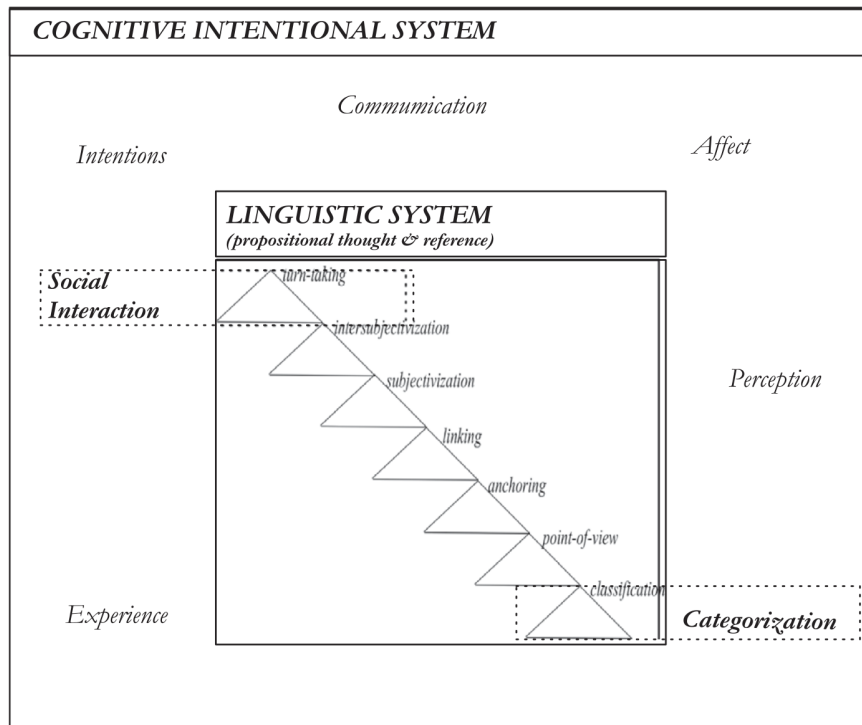


Figure 3 The Bridge Model (Hinzen & Wiltschko 2023: 75).

component between participants in conversation. This model is referred to as the Interactional Spine Hypothesis.

In this context, then, and as far as acquisition is concerned, the interactional and universal spine proposed in Wiltschko’s work is hypothesised to develop inwardly, starting from the edges and then developing towards the centre of the spine — this proposal receives the name of the Inward Growing Spine Hypothesis in Heim & Wiltschko (2021) (cf. also Wiltschko 2021). In Hinzen & Wiltschko (2023), a model of general language and cognition, the Bridge Model, is proposed, which views linguistic cognition as resting on two partially pre-linguistic pillars (perceptual categorisation and social-communicative interaction); grammar acts as a bridge that mediates both (as displayed in Figure 3). Being argued to be phylogenetically prior, the approach takes the thematic and categorisation-based domain (Wiltschko’s, 2014, Classification head) and the interactional and speech-act-based domain (Wiltschko’s, 2021, Responding head) to develop first. Development and maturation begin in these two phylogenetically-prior domains. Subsequently, Linking (hosting, e.g. wh-questions or embedding complementisers) matures. A third step sees the division of RespP into RespP and GroundP, and Anchoring matures to accommodate tense and other deictic elements (e.g. demonstratives for the nominal domain). The final predicted stage is one which subdivides GroundP into Ground-Addressee and Ground-Speaker as well as Anchoring into Anchoring and PoV. Data on the acquisition of English *huh* is presented as preliminary evidence for part of Heim

& Wiltschko's (2021) Inward Growing Spine Hypothesis. As a result, their Inward Growing Spine would accordingly predict child production data to present evidence of at least some illocutionary complementisers early on, given the proposed inward maturational timeline.

In a proposal predating Heim & Wiltschko (2021), Biberauer (2018) also converged on the hypothesis that some interactional and speaker-hearer-oriented material may be early-acquired, albeit for somewhat different reasons. Specifically, Biberauer (2018, *et seq.*) argues that structural edges (in particular, *phasal* edges) will be of particular significance in acquisition, as these constitute the locus of here-and-now and speaker-hearer-oriented material. This is captured in her Peripheral Speaker-Hearer Hypothesis, or PSHH, defined below and whose logic is explained in the following paragraphs (Biberauer 2018: 4):

(17) **The Peripheral Speaker-Hearer Hypothesis (PSHH)**

Speaker-hearer perspective is formally encoded at the edges of phasal domains, where phasal domains are independently signalled, realizationally (PF) and interpretively (LF) privileged structural domains, the precise identity of which differs from language to language, and the 'size' of which may also differ from derivation to derivation language-internally (i.e. the 'dynamic'⁷ perspective on phases).

What is of importance in this approach is that the peripheral nature of these speaker-hearer materials – and, therefore, their potential significance in acquisition – fall out from independent motivations, stemming from the formal properties of perspectival elements and the dynamics of phase-based derivations. Firstly, speaker-hearer and perspectival elements such as discourse particles are often acategorical elements, meaning they do not alter the category of the element they combine with (for instance, a focused DP remains a DP and can still be selected as an internal or external argument). Further, optional and ungrammaticalised discourse particles are analysed as being formal-feature-less ($[F]$ -less), since, for instance, the distribution of these elements is not constrained to specific functional positions⁸ (see Biberauer 2017 on several modals and functional particles in many East Asian and African languages, among others). Assuming that selection relies on the presence of $[F]$ s, these discourse particle roots, and similarly $[F]$ -less interactionally-oriented elements more generally, also cannot be selected, and they cannot intervene for selection either. Note that this does not imply all discourse particles and interactional elements are fully $[F]$ -less (see Biberauer 2017, especially section 4.4, on grammaticalised Q-particles and other grammaticalised particles with acategorical but nonetheless

⁷ See so-called dynamic or contextual approaches to phases (i.a. den Dikken 2007, Gallego 2006, 2010, Harwood 2013, Wurmbrand 2013, Bošković 2014).

⁸ One could, of course, ascribe speaker-hearer-related features to these elements. The perspective taken in Biberauer (2018), however, is one which hinges on Jakobson's (1941) and Jakobson & Halle's (1956) contrastivity requirement for formal features (initially suggested for phonology), whereby an $[F]$ is postulated by the acquirer if, and only if, it is required to capture a contrast in the system (see also Biberauer 2019 on Maximise Minimal Means, feature postulation and so-called 'departures from Saussurean arbitrariness').

[*F*]-sensitive properties). The point is that these acategorical elements will lack the kind of [*F*]-specifications required to be integrated into a [$\pm V$] spine (e.g. a [+*V*] clausal Extended Projection, EP), by virtue of either lacking [*F*]s entirely (which is the case for ungrammaticalised particles, as discussed above) or possessing [*F*]s which are not selected for (e.g. [FOCUS] or [TOPIC]; again, see Biberauer 2017).

Their formal nature therefore raises the question of how and when these elements are merged into the derivation, particularly if Merge is taken to be a [*F*]-sensitive operation (Pesetsky & Torrego 2006, 2007, Wurmbrand 2014). Biberauer (2017, 2018) argues that the device of lexical arrays (LAs) is one possible solution and, in turn, a means with which to understand both the properties of speaker-hearer elements and their role in the system (where an LA denotes the set of items selected from the lexicon to be used in the derivation within a given phase; see also Chomsky 2000). If we assume that Merge first operates on elements that can be merged based on their [*F*]s, such perspectival and [*F*]-less elements will only be able to leave their LA in a given phase once all the elements specified with [*F*]s (which *can* be selected) have been merged. In other words, unable to integrate with the clausal EP via standard selection relationships, syntactically inert perspectival elements will always be ‘last out’ of their LA and their respective phasal domains, and must therefore merge as phase-peripheral adjuncts. These ‘underspecification’ properties (namely, lack of [*F*]s, making them formally inert, but also lack of syntactic relations, such as c-selection, from which they could gain an interpretation) also imply that the interpretation at LF of these particles is instead compositionally modulated based on the projecting head and hierarchical position where they are merged. Often, this will be the phase-edge-located speaker-oriented structure that has been argued to dominate the CP domain (this is the case for illocutionary complementisers; see also discussion below). It follows, then, that information from the discourse and ‘here-and-now’ context has to be resorted to in order to interpret these types of underspecified items, a result which aligns with the very nature of perspectival elements.

Altogether, therefore, this predicts that these kinds of perspectival items would necessarily be *phase peripheral*. This is the logic of the *Last-out mechanism* specified below, which helps rationalise why perspectival elements would be peripheral (Biberauer 2018: 9):

(18) ***The Last-out mechanism***

[*F*]-less elements must be last out of the Lexical Array/LA defining their phasal domain: being unselectable by other elements, and also not able to select themselves, such elements can only leave their LA when all the [*F*]-bearing elements – which can select/be selected – have been merged.

In other words, phasal edges provide an opportunity to formally integrate unselectable elements as part of the system, since perspectival elements will always be last out of their LA, i.e. last-merged in the phasal domain, as a consequence of being [*F*]-less.

If perspectival, speaker-hearer elements – such as modal particles, emphatic polarity particles, illocutionary complementisers, among others – are necessarily

phase peripheral, these edge-based elements would furthermore serve a key bootstrapping heuristic in acquisition (namely, signalling domain edges) and so, at least some of these perspectival elements, may be expected to be acquired early⁹ (see the discussion in [Biberauer 2019](#) for some synchronic and diachronic consequences, also in the context of spoken, vernacular varieties).

Therefore, the PSHH makes important predictions in our context, insofar as illocutionary complementisers are located at phase edges, encoding, i.a. perspectival and speaker-hearer information (see [section 2](#)). In particular, PSHH predicts an acquisitionally significant role of illocutionary information by virtue of the role of structural *edges* in the system and their associated perspectival nature, as outlined above. Crucially, and as already discussed in [section 2](#), illocutionary complementisers cannot be readily analysed in the same way as subordinating C-elements. Besides being mostly located higher up in the clause than embedding complementisers and often in the SAP domain, they also cannot be selected (being main-clause elements and always peripheral), thereby displaying the same formal profile as the elements that the PSHH is concerned with. As a consequence, PSHH likewise foresees a possible developmental mismatch between the emergence of illocutionary vs. embedding complementisers – the peripheral and speaker-hearer facet of illocutionary complementisers is something acquirers are expected to be particularly drawn to. From a PSHH perspective, a key task in acquisition involves integrating $[F]$ -less elements or items not fully specified for $[F]$ s into the system (by virtue of the simple fact that children do not begin the learning task with a prior set of fully specified lexical and functional elements). Therefore, a mechanism and heuristic that allows them to do so is, learnability-wise, highly helpful and, since the pivotal role of phasal edges facilitates interpretation via anchoring to the ‘here-and-now’ context, it ratifies why these elements would have interactional and perspectival content.

Unlike the Inward Growing Spine Hypothesis, however, the prediction that illocutionary complementisers should be early-acquired importantly does not find its aetiology in some biological mechanism that trails through an innate clausal hierarchy and dictates acquisition orderings, as would be the case in maturational approaches. Instead, it is phasal edges and their distinctive role in the system that facilitate crucial syntactic domain-size learning in a language acquisition context and comprise productive ‘ways in’ for elements that have not yet been fully formally integrated into the structure (see also [Biberauer 2019](#)).

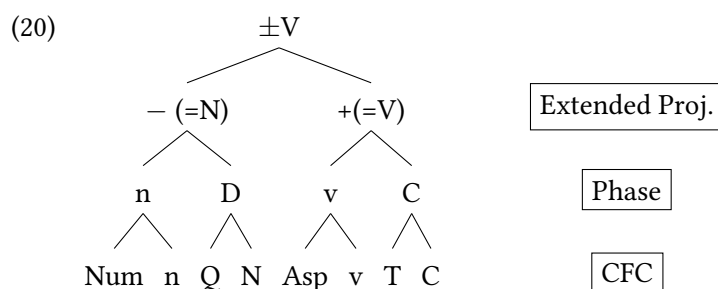
Finally, a similar theoretical proposal is made in [Biberauer & Roberts \(2015\)](#). This approach represents an attempt to unify independently suggested formal hierarchies and encodes these as a single formal hierarchy that is not UG-given, but instead an emergent property of the interaction of the three factors of language design (namely, UG, the Primary Linguistic Data and third-factor, general-cognitive principles, see [Chomsky 2005](#); cf. also [Biberauer 2019](#)). Crosslinguistically, according to [Biberauer](#)

⁹ Consider also various suggestions in the acquisition literature whereby structural elaboration in child grammars may progress via an initial Adjunction (thus $[F]$ -free) stage, as also proposed in [Biberauer \(2018\)](#) (see, i.a. [Hoekstra & Jordens 1996](#), and [Roy, Copley & McCune 2016](#) on Merge and Juxtaposition in development).

& Roberts (2015), clauses can be analysed at different levels of ‘magnification’ or ‘granularity’ (adapted from Biberauer & Roberts 2015: 6):

- (19) **Extended Projection (V)** (Grimshaw 1991) > **phase** (C, v) (Chomsky 2001, 2008) > **Core Functional Category** or **CFC** (C, T, v) (Chomsky 2000, 2001) > ‘**cartographic fields**’ (e.g. Tense, Mood, Aspect, Topic, Focus) > **semantically distinct heads** (as in Cinque 1999, Speas & Tenny 2003, Frascarelli & Hinterhölzl 2007, Haegeman & Hill 2013, among others).

(19) can be reformulated as a hierarchy, which unifies and identifies overall similarities between independently proposed hierarchies and functional sequences in the literature (see the references in 19).¹⁰ The hierarchy thus begins at an initially ‘undivided’ categorial space and gradually increments the level of grain with which the system is analysed:



Importantly for present purposes, the hierarchy in (20) represents a learning path, in which acquirers successively divide the categorial space into ever more fine-grained units (cf. also the Successive Division Algorithm in Drescher 2009). In Bosch (in progress) I provide corpus data from five languages that supports such a developmental progression, at least partially. After initially establishing a basic distinction between predicates/arguments (possibly aided by prosodic cues; see the discussion in Biberauer 2019) and subsequently one between nominal/verbal Extended Projections (the Extended Projection layer in 20), I suggest that the earliest developmental stages display evidence of a further subdivision between an underspecified discourse/interactional domain (hosting, e.g. early wh-words and interactional language) and a thematic domain (the Phasal stage in 20). The TP domain emerges either simultaneously with the former or, less clearly, a little after basic discourse and thematic material (the Core Functional Category or CFC stage). Like the Inward Growing Spine Hypothesis and the PSHH, therefore, Biberauer & Roberts’s (2015) approach argues too for early availability of illocutionary and discourse information in developing grammars.

Therefore, several separate and recent proposals in the literature of acquisition and variation converge on the proposal that some discourse, interactional and ‘here-and-now’ linguistic material is expected to be acquisitionally most accessible.

¹⁰ There may still be a level of nanosyntactic organisation, as Biberauer & Roberts (2015: 4, fn. 1) note, but we leave this aside as it is orthogonal to the present discussion.

I will adopt here the neo-emergentist perspective in [Biberauer & Roberts \(2015\)](#), [Biberauer \(2018, 2019\)](#) and the associated findings in [Bosch \(in progress\)](#), according to which development proceeds in the successive-division and edge-centred manner just outlined. According to [Biberauer & Roberts \(2015\)](#), the child initially makes a basic predicate/argument (or ‘archi’ N/V) distinction (see [Douglas 2018](#), [Song 2019](#) and [Biberauer 2019](#) on ‘archi-V’ and ‘archi-N’ features and categories), which then aids them with making a first representational division into the verbal and nominal Extended Projections. Subsequently, the learning path proceeds by making a coarse-grained subdivision between discourse/interactional (CP- and SpeechActP-internal) material and thematic (vP-internal) material.¹¹ In contrast to bottom-up perspectives, this approach *does* expect at least some illocutionary complementisers to be harnessed early on because of their interactional and discourse nature. These are also predicted to surface before embedding complementisers are first detectable, owing to the fact that illocutionary complementisers are attested in main clause contexts and thus do not require knowledge of subordination. The predictions made by this approach, as far as illocutionary complementisers are concerned, are identical to those made by [Heim & Wiltschko’s \(2021\)](#) Inward Growing proposal; namely, that at least some illocutionary complementisers will emerge early and before embedding complementisers. The present work will not discuss their predictive differences beyond the domain of complementiser types (though see [Bosch in progress](#) for some discussion on this latter point).

To reiterate, however, none of these approaches are committed to the absolute claim that *all* illocutionary complementisers (in their complete inventory of subtypes) will necessarily emerge before any embedding complementisers; the vital prediction that matters in the present context is that they do expect early production of *some* illocutionary complementisers, a scenario which would be disallowed in bottom-up maturational approaches. This work will not establish the relative development of each subtype of illocutionary complementiser (conjunctive, quotative, exclamative, interrogative, etc.), compared to embedding ones, and so will leave open the possibility that some kinds of illocutionary complementisers may develop later than others. This is left for future work.

With this in mind, we have paved the way for the presentation of the corpus study in the next section, which will discuss its methodology and results.

4 CORPUS STUDY

4.1 Summary of structures analysed

To summarise the discussion in [section 2](#), the following constructions were analysed and quantified in every corpus. As discussed earlier, these structures are hypoth-

¹¹ Therefore, I will not aim to tease apart here whether CP emerges before/after the Speech-Act domain. This would entail a larger project and, necessarily, this paper is more modest in scope. For now, I take these two layers to be potentially encoded as a single underspecified domain at the earliest stage, in line with [Biberauer & Roberts’s \(2015\)](#) hierarchy. This working and speculative assumption inevitably requires further work, so it is perfectly possible that, as more developmental data becomes available, it will have to be dropped or reshaped.

esised to be located in a variety of positions spanning both the utterance-related layer and the traditional CP (proposition-oriented) domain.

i. Illocutionary complementisers

- a. Exclamative
- b. Conjunctive
- c. Quotative
- d. Interrogative (in Catalan only)
- e. Adverb/Adjective + *que*

ii. Subordinating complementisers

- a. Complement clauses
- b. Relative clauses introduced by *que* (including free relatives¹²)

4.2 Methodology

Using the CLAN programme and its *kwal* command (MacWhinney 2000), we automatically extracted all occurrences of *que* and their conversational contexts for 10 Catalan and Spanish children in the CHILDES corpora. After manually excluding utterances where *que* did not correspond to a complementiser, this yielded $N = 1319$ utterances from children aged 0;11 to 4;08 that contained a complementiser. This is summarised in Table 1. The utterances were annotated for type of complementiser, either illocutionary or subordinating. The number of instances in which each kind of complementiser was produced per file was also counted. As the data is very often ambiguous and it is hard to discern the kind of illocutionary complementiser produced (e.g. discriminating between quotative and exclamative complementisers), I did not quantify each subtype of illocutionary complementiser and I treated them as a heterogeneous group, aiming simply to compare the (un)availability of illocutionary vs. embedded-clause complementisers. Detailed coding of these subtypes would be a welcome avenue of future work, however.

Cases of complementisers followed by an adjective only (such as in Cat. *Que bonic!* ‘how beautiful!’) were disregarded, on the grounds that these do not feature a finite verb (and so may not be reliable diagnostics for CP and Speech-Act-related projections). Importantly, they can also be conflated with other structurally and interpretively similar exclamative constructions, such as some kinds of wh-exclamatives. Therefore, only instances of sentences with a complementiser and a verb were counted. Any ambiguous, unclear or unintelligible utterances were excluded. Consecutive repetitions of the same construction were also excluded, as were identical imitations of an utterance by an adult.

¹² Free relatives are included in the analysis on the grounds that these often feature the complementiser *que* in Spanish and Catalan. For example, Sp. *Prefiero el que comí ayer* ‘I prefer the one I ate yesterday’. Just like with relative clauses more generally, any free relatives that do not include *que* are excluded from the analysis (e.g. Cat. *M’agrada on vius* ‘I like where you live’).

Language	Corpus	Children	<i>N</i>
Catalan	Serra/Solé	Laura	203
		Gisela	201
		Àlvar	15
		Guillem	105
	Júlia	Júlia	4
Spanish	Llinàs/Ojea	Irene	90
		Yasmin	42
	Aguado-Orea/Pine	Juan	244
	Aguirre	Magín	295
	Vila	Emilio	120
Total		10	1319

Table 1 CHILDES corpora studied (Serra & Solé 1989, Bel 2001, Llinàs-Grau & Ojea 2000, Aguado Orea & Pine 2015, Aguirre 2000, Vila 1990) and utterances with complementisers.

Additionally, I adopt Mean Length of Utterance (MLU) as the guiding measure for syntactic development, to be used to compare the stages across all children. As several authors have noted (Clahsen, Penke & Parodi 1993, Paradis & Genesee 1997, Caprin & Guasti 2009, Friedmann et al. 2021), age is to be avoided in both intralinguistic and crosslinguistic comparisons, to evade the high variability in linguistic development that is observed across 2-year-olds and to provide a developmental metric that, in later studies, will make it possible to match the children studied with other children acquiring the same or other languages. We calculated the word-based MLU for each file and each child. MLU was calculated via the `mlu` program in CLAN, by running the command ‘`mlu +t*CHI -t%MOR *.cha`’. The lowest and highest MLU values in each child’s production were used to provide the MLU range for their production.

The details for each child studied, their MLU range and the number of files analysed are reported in Table 2.

4.3 Results

Table 3 reports the breakdown of productions by type of complementiser for each child. Overall, among the 1319 total utterances with complementisers, 1010 of them (77%) corresponded to examples with illocutionary and interrogative complementisers and 309 corresponded to subordinating complementisers (23%). As expected, given the main clause nature of illocutionary complementisers, these are much more frequent in the child production data.

Language	Corpus	Children	Age range	Files analysed	MLU range
Catalan	Serra/Solé	Laura	1;07-4;00	19	1.03-3.47
		Gisela	1;07-4;02	20	1.02-3.51
		Àlvar	1;02-3;01	21	1.0-3.37
		Guillem	1;01-4;00	34	1.01-3.88
	Júlia	Júlia	1;07-2;06	17	1.15-2.74
Spanish	Llinàs/Ojea	Irene	0;11-3;02	40	1.0-4.94
		Yasmin	1;10-2;09	47	1.29-3.21
	Aguado-Orea/Pine	Juan	1;10-2;05	65	1.34-3.39
	Aguirre	Magín	1;07-2;10	29	1.24-3.07
	Vila	Emilio	0;11-4;08	35	1.0-3.23

Table 2 Children studied in the CHILDES database and summary information.

Language	Children	Illocutionary	Embedding
Catalan	Laura	155 (76.4%)	48 (23.6%)
	Gisela	148 (73.6%)	53 (26.4%)
	Àlvar	9 (60%)	6 (40%)
	Guillem	85 (81%)	20 (19%)
	Júlia	3 (75%)	1 (25%)
Spanish	Irene	58 (64.4%)	32 (35.6%)
	Yasmin	36 (85.7%)	6 (14.3%)
	Juan	164 (67.2%)	80 (32.8%)
	Magín	248 (84.1%)	47 (15.9%)
	Emilio	104 (86.7%)	16 (13.3%)
Total		1010 (76.6%)	309 (23.4%)

Table 3 Proportion of use by type of complementiser.

We provide examples of illocutionary complementiser productions in (21), for Catalan children, and (22), for Spanish. At the earliest stages, utterances with illocutionary complementisers are syntactically simple, typically involving an illocutionary complementiser alongside a verb and a (null/overt) subject.

- (21) a. *Que ja no fa mal?* [Guillem; MLU 1.99]
 that.INT already not make.3SG pain
 ‘Does it not hurt anymore?’

- b. *Ai, **que** crema!* [Laura; MLU 1.35]
 ouch that.EXCL burn.3SG
 ‘Ouch, it’s burning!’
- c. *Espera’t, **que** estic fent això* [Gisela; MLU 2.61]
 wait=CL.REFL that.CONJ am doing this
 ‘Wait, I’m doing this.’
- d. ***Que** no hi és* [Àlvar; MLU 1.84]
 that.QUOT not CL.LOC= is
 ‘(I’ve already said) It’s not there.’
- (22) a. ***Que** no qu(i)ero* [Juan; MLU 1.58]
 that.QUOT not want.1SG
 ‘(I said) I don’t want to.’
- b. *¡**Que** viene!* [Magín; MLU 1.78]
 that.EXCL come.3SG
 ‘He/she is coming!’
- c. *Ay, no, **que** me harán daño a* [Emilio; MLU 2.2]
 ouch no that.CONJ CL.IO= do.FUT.3PL harm to
la barriga
 the tummy
 ‘Ouch, no, they’ll hurt my tummy.’

In contrast, the first instantiations of subordinating complementisers are already syntactically more sophisticated, by virtue of being attested at a later developmental stage (see discussion below). Some examples illustrating early relative clauses and complement clauses are given in (23). Note that several of the English translations below feature a null *that*-complementiser. Whilst we would not be able to diagnose availability of complementation in child English exclusively on the basis of production of *overt* complementisers, this problem does not arise with the Catalan and Spanish data. This is because null complementisers are generally ungrammatical in both Catalan and Spanish (except in an infrequent set of typically formal contexts with subjunctive mood in Spanish and, to a lesser extent, in Catalan; see, i.a. [Etxepare 1996](#), [Antonelli 2013](#), [Llinàs-Grau & Fernández-Sánchez 2013¹³](#)).

¹³ [Llinàs-Grau & Fernández-Sánchez \(2013\)](#) additionally argue that, although these constructions are superficially similar in that a complementiser is missing, the constructions that allow a null *que* in Catalan and Spanish do not involve the same process as the ones that may underlie *that*-deletion in English.

Not All Complementisers Are Late

- (23) a. *Saps que no vindrà ningú?* [Gisela; MLU 2.61]
 know.2SG that not come.FUT.3SG no.one
 ‘Do you know no one is going to come?’
- b. *Una vegada hi havia un nen que* [Júlia; MLU 2.74]
 one time CL.LOC= AUX.IMPF.3SG a boy that
es diu Andreu
 CL.REFL= say.3SG Andreu
 ‘Once upon a time, there was a boy named Andreu.’
- c. *En una capsa que hi ha aquí* [Àlvar; MLU 2.82]
 in a box that CL.LOC= AUX.3SG here
 ‘In a box that’s here.’
- (24) a. *Quiero que sea un zapato* [Yasmin; MLU 2.47]
 want.1SG that be.SUBJ.3SG a shoe
 ‘I want it to be a shoe.’
- b. *¿No ves que estaba con la raqueta?* [Irene; MLU 3.23]
 not see.2SG that was with the racquet
 ‘Don’t you see it was next to the racquet?’
- c. *Había un señor muy malo que* [Juan; MLU 2.73]
 AUX.IMPF.3SG a man very bad who
tiraba flechas
 throw.IMPF.3SG arrows
 ‘There was a very bad man who was throwing arrows.’

We turn now to the predictions made by the different accounts outlined in [section 3](#). As per a bottom-up maturational account, both kinds of complementisers are predicted to be late acquisitions and to surface at roughly similar times, on the grounds that these are hosted at the very top of the functional spine and that a rigid, biologically-timed developmental pathway specifies the acquisition ordering of functional categories. In contrast, this is not borne out in the children studied. What we find instead is that the first illocutionary complementisers typically appear well before embedding complementisers and never later. In a couple of children only (Júlia and Emilio), both types of complementisers emerge simultaneously in the production data. Although the ten children differ in the number of complementisers they produce, these patterns recur in all of them.

The MLU stage at which illocutionary complementisers emerge for every child is reported in [Table 4](#).¹⁴

¹⁴ The calculation of the average MLU excludes outlier values, which were Júlia’s (cf. also discussion later).

Language	Children	Illocutionary	Embedding
Catalan	Laura	1;10.22	3;00.02
		1.15 MLU	2.42 MLU
	Gisela	1;08.24	2;08.00
		1.13 MLU	2.61 MLU
	Àlvar	2;02.06	2;06.25
		1.84 MLU	1.91 MLU
	Guillem	2;02.28	2;11.25
		1.54 MLU	2.44 MLU
Spanish	Júlia	2;06.25	2;06.25
		2.74 MLU	2.74 MLU
	Irene	1;08.09	1;08.26
		1.88 MLU	2.28 MLU
	Yasmin	1;10.08	2;05.18
		1.93 MLU	2.47 MLU
	Juan	1;11.11	2;01.21
		1.58 MLU	1.77 MLU
	Magín	1;09.01	1;10.00
		1.78 MLU	2.73 MLU
	Emilio	2;04.17	2;04.17
Average		1.66 MLU	2.31 MLU

Table 4 Emergence of illocutionary and embedding complementisers.

We ran a paired samples t-test in R (R Core Team 2022) to confirm that the MLU at the point of emergence is different for illocutionary and subordination complementisers. Any outliers in the dataset were excluded. This only involved Júlia’s MLU value for the emergence of illocutionary complementisers, which was radically different than for the other children. As a result, I excluded her data for both illocutionary and embedding complementisers from the t-test. The results confirmed that illocutionary complementisers were significantly more likely to emerge earlier than embedding complementisers ($t(17) = 5.5808, p < .001$).

On average, the two kinds of complementisers emerged at the MLU values given in Table 5.

	Illocutionary	Embedding
Catalan	MLU 1.42 (range 1.15-1.84)	MLU 2.35 (range 1.91-2.61)
Spanish	MLU 1.87 (range 1.58-2.18)	MLU 2.29 (range 1.77-2.73)
Combined	MLU 1.66 (range 1.15-2.18)	MLU 2.31 (range 1.77-2.73)

Table 5 Average and range of MLU values across language groups for the emergence of illocutionary and embedding complementisers.

Finally, the children’s acquisition of both kinds of complementisers over time is shown in [Figure 4](#), which plots their development according to the children’s MLU (with the y-axis indicating number of occurrences attested for each complementiser type). Each child’s individual development and their graphs are collated in the [Appendix](#).

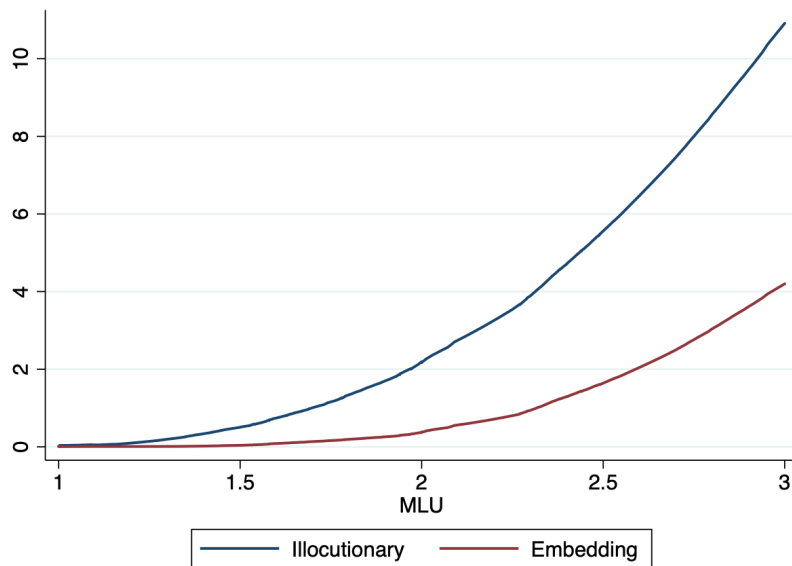


Figure 4 The development of complementisers in the Catalan and Spanish children.

As can be seen from [Figure 4](#) and the slopes of the curves, illocutionary complementisers both emerge earlier and develop faster in frequency than their subordinating counterparts. A one-sample Kolmogorov-Smirnov test was carried out to check that the two distributions are statistically significantly different from each other, which rejected the null hypothesis that the two curves are equal ($D = 2.0000$, $p < .001$).

In summary, illocutionary complementisers are accessible early on in most of the children studied and are almost always produced several files before embedding complementisers emerge. In four out of the ten children studied (Laura, Gisela,

Guillem and Juan), these emerge from the earliest multiword utterances, by around 1.5 MLU or even earlier. We conclude, therefore, that the acquisition of illocutionary complementisers and embedding complementisers is developmentally dissociable and unrelated, unlike what the predictions of a bottom-up maturational account would lead us to expect.

5 DISCUSSION

The results of the corpus study show that illocutionary complementisers emerge significantly earlier than embedding complementisers, supporting the common predictions made by approaches to the acquisition of functional categories that anticipate early emergence of the CP and the SAP domains (e.g. [Roeper & Rohrbacher 1994](#), [Galasso 2003](#), [Tsimpli 2005](#), [Roeper 2007](#), [van Kampen 2010](#), [Biberauer & Roberts 2015](#), [Biberauer 2018](#), [Bosch in progress](#); recall [section 3](#)). As expected by these approaches, illocutionary complementisers often emerge well before embedding complementisers are evident and never emerge later than the latter in the ten children studied. This is hypothesised to be due to both their speaker-hearer and main-clause nature. We have also established that, for some children at least, some illocutionary complementisers appear available from the earliest MLU stages (around 1.5 MLU or sometimes even earlier), indicating that a CP domain is possibly accessible at this stage for these children. This finding supports multiple recent approaches that propose children make efficient use of interactional/discourse, edge-based and ‘here-and-now’ cues in building their incipient grammars (as discussed in [section 3](#)). As noted earlier, the current study does not establish whether *all* types of illocutionary complementisers emerge before embedding complementiser (a separate empirical question). The important point, nonetheless, is that the first instances of illocutionary complementisers are never attested after the emergence of embedding complementisers.

This observation, although necessarily preliminary, could have significant consequences for developmental theories that pursue a bottom-up maturational track, and which would predict CP knowledge to emerge substantially late, with illocutionary and embedding complementisers developing at comparatively ‘delayed’ stages. The early availability of illocutionary complementisers contradicts such proposals and suggests instead that, whilst embedding complementisers may indeed be later phenomena compared to, for instance, argument structure and basic tense/agreement marking (as bottom-up proposals predict), not all complementisers are equally stagnant at early stages. I preliminarily suggest, therefore, that bottom-up approaches do not seem well-suited to account for the emergence of this grammatical knowledge, at least in Catalan and Spanish, where illocutionary complementisers abound in both child production, spoken language and, thus, carer input.¹⁵ This paper supplements the findings in [Bosch \(in progress\)](#), where it is suggested that the children studied abide by the predictions made by [Biberauer & Roberts’s \(2015\)](#) emergent categorial

¹⁵ Although this study did not quantify the adult data, a cursory examination of the adult utterances in CHILDES indicates that illocutionary complementisers are very frequent in the input, at least impressionistically.

hierarchy and that their data does not support bottom-up structural development as far as emergence (first use of a given structure) is concerned.

These findings must be treated as preliminary, due to the limited and exploratory nature of the corpus study presented here and the relatively small sample ($N = 10$). Nonetheless, the patterns reported in this study had thus far gone unnoticed and represent novel evidence that makes the case for early availability of a discourse/interactional domain in child grammars (cf. [Roeper 2007](#) on the importance of early expressive language). More attention should also be devoted to the role of maturation in language development: while it is clear that embedding complementisers emerge later than the first illocutionary complementisers, it remains to be determined whether this is simply because embedding complementisers require more sophisticated syntactic knowledge accommodating subordination or whether this is due to some maturational timeline (the Inward Growing Spine Hypothesis would predict Linking to be accessible later due to maturational constraints). If it is correct, however, that some illocutionary complementisers such as interrogative and quotative complementisers are located in the CP or Wiltschko's Linking domain (and the child encodes them as such), their early emergence may be at odds with the Inward Growing Spine Hypothesis. A separate corpus study analysing the relative development of the different subtypes of illocutionary complementisers would be needed to test for this. Generally, the data presented here does not have direct implications regarding the cause behind the relative delay of embedding complementisers. A theoretically more parsimonious approach would be one that foregoes appeal to innate developmental mechanisms such as maturation and resorts, for instance, to the differences between main and embedded contexts and their syntactic complexity, as well as to the salience of interactional/discourse content in acquisition. Notwithstanding the advantage of these 'early-CP' and 'inward-growing' approaches in accounting for the presented data, it remains to be seen whether they make the correct predictions for acquisition more generally and, importantly, whether development of syntactic categories necessarily needs to be modelled via a maturational mechanism, as suggested by Heim and Wiltschko (see [Bosch in progress](#) for arguments that a neo-emergentist approach may fare better).

Additionally, further research is required to determine whether these findings recur in other Spanish and Catalan children and, particularly, in other Ibero-Romance languages, notably Portuguese. Examination of crosslinguistic differences in the acquisition of illocutionary complementisers could also be helpful. As seen in [Table 4](#) and [Table 5](#), Catalan children appear faster than the Spanish children studied in producing illocutionary complementisers. Conceivably, this could be due to the relatively higher prevalence of illocutionary complementisers in Catalan, given the frequent use of interrogative complementisers in spoken language, which are unavailable in Spanish. The sample per language (5 children) is likely too small to make any conclusive statements, but future work should aim to probe for these timing differences further. One should also pay due attention to dialectal differences; for instance, whether developmental timings differ between children speaking different dialects, such as different varieties of Latin American vs. Peninsular Spanish, or even Catalanian Spanish vs. Spanish spoken in other areas of the

Iberian Peninsula. Finally, this paper mainly employs naturalistic corpus data, but experimental tests, such as comprehension studies, would also be revealing as regards the state of children's knowledge.

6 CONCLUSION

In this paper, I investigated the emergence of illocutionary complementisers in child Catalan and Spanish and compared it to the use of embedding complementisers. I established that illocutionary complementisers consistently emerge earlier in child speech than embedding complementisers. Since at least some illocutionary complementisers are often attested at early MLU values, I argued that this finding is hard to accommodate in bottom-up maturational approaches to the acquisition of functional categories, which propose the CP domain is acquired last. This developmental pattern can instead be understood with approaches which suggest an interactional/discourse domain alongside vP are first to develop in the learning path. The developmental differences between the two kinds of complementisers feasibly also come down to the distinction between main/embedded clauses, with embedding complementisers appearing later owing to the increased syntactic sophistication of the subordination structures of which they form part. More broadly, the early emergence of illocutionary complementisers lends additional support to the salience of speech-act and discourse material in acquisition and suggests further study of their acquisition can likely enhance our understanding of the make-up of early child grammars. Echoing the sentiment in, i.a. [Roeper \(2007\)](#), [van Kampen \(2010\)](#), [Biberauer \(2018\)](#) and [Heim & Wiltchko \(2021\)](#), their development helps underscore the possibility that *peripheral* and thus far underdiscussed elements may nonetheless be *core* stepping stones in the process of grammar construction.

REFERENCES

- Aguado Orea, J. & J. Pine. 2015. Comparing different models of the development of verb inflection in early child Spanish. *PloS ONE* 10(3). e0119613.
- Aguirre, C. 2000. *La adquisición de las categorías gramaticales en español*. Madrid: Ediciones de la Universidad Autónoma de Madrid.
- Antonelli, A. 2013. On the Left Periphery of Spanish Complementizerless Clauses. In J. Cabrelli Amaro & et al. (eds.), *Selected Proceedings of the 16th Hispanic Linguistics Symposium*, 15–26. Somerville, MA: Cascadilla Proceedings Project.
- Armon-Lotem, S. 2005. The Acquisition of Subordination: From Preconjunctivals to Later Use. In D. Ravid & H. B.-Z. Shyldkrot (eds.), *Perspectives on Language and Language Development*, 192–204. Kluwer.
- Bel, A. 2001. *Teoria lingüística i adquisició del llenguatge: Anàlisi comparada dels trets morfològics en català i en castellà*. Barcelona: Institut d'Estudis Catalans.
- Belletti, A. 2004. Aspects of the low IP area. In L. Rizzi (ed.), *The structure of CP and IP: The Cartography of syntactic structures*, vol. 2, 16–51. Oxford: Oxford University Press.

- Biberauer, T. 2017. The Final-over-Final Condition and Particles. In M. Sheehan, T. Biberauer, I. Roberts & A. Holmberg (eds.), *The Final-Over-Final Condition: A Syntactic Universal*, 187–296. Cambridge, MA: MIT Press.
- Biberauer, T. 2018. Peripheral significance: a phasal perspective on the grammaticalisation of speaker perspective. Talk presented at *DiGS 20* (York), June. https://www.york.ac.uk/media/languageandlinguistics/documents/conferences/digs20/Biberauer_handout_DiGS20.pdf.
- Biberauer, T. 2019. Factors 2 and 3: Towards a principled approach. *Catalan Journal of Linguistics* 45–88.
- Biberauer, T. & I. Roberts. 2015. Rethinking formal hierarchies: a proposed unification. *Cambridge Occasional Papers in Linguistics* 7(1). 1–31.
- Bosch, N. in progress. *Emergent Syntax and Maturation: a neo-emergentist approach to syntactic development*: University of Cambridge MPhil thesis.
- Bošković, Ž. 2014. Now I’m a Phase, Now I’m Not a Phase: On the Variability of Phases with Extraction and Ellipsis. *Linguistic Inquiry* 45(1). 27–89.
- Bresnan, J. W. 1972. *Theory of complementation in English syntax*: Massachusetts Institute of Technology dissertation.
- Caprin, C. & M. T. Guasti. 2009. The acquisition of morphosyntax in Italian: A cross-sectional study. *Applied Psycholinguistics* 30(1). 23–52.
- Chomsky, N. 2000. Minimalist inquiries: the framework. In R. Martin, D. Michaels & J. Uriagereka (eds.), *Step by step: essays on minimalist syntax in honor of Howard Lasnik*, 85–156. Cambridge, MA: MIT Press.
- Chomsky, N. 2001. Derivation by Phase. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*, 1–50. Cambridge, MA: MIT Press.
- Chomsky, N. 2005. Three factors in language design. *Linguistic Inquiry* 36(1). 1–22.
- Chomsky, N. 2008. On Phases. In R. Freidin, C. P. Otero & M. L. Zubizarreta (eds.), *Foundational issues in linguistic theory*, 133–166. Cambridge, MA: MIT Press.
- Cinque, G. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford: Oxford University Press.
- Clahsen, H. & M. Penke. 1992. The acquisition of agreement morphology and its syntactic consequences: New evidence on German child language from the Simone-corpus. In J. Meisel (ed.), *The Acquisition of Verb Placement: Functional Categories and V2 Phenomena in Language Acquisition*, 181–224. Dordrecht: Kluwer.
- Clahsen, H., M. Penke & T. Parodi. 1993. Functional Categories in Early Child German. *Language Acquisition* 3(4). 395–429.
- Corr, A. 2016. *Illocutionary complementisers and utterance syntax*: University of Cambridge dissertation.
- Corr, A. 2022. *The Grammar of the Utterance: How to Do Things with Ibero-Romance*. Oxford: Oxford University Press.
- Cruschina, S. & E.-M. Remberger. 2018. Speaker-oriented syntax and root clause complementizers. *Linguistic Variation* 18(2). 336–358.
- Diercks, M. & M. Bossi. 2021. Minimalist syntax is psychologically real: Lessons from (counter)cyclicity. *lingbuzz/005671*.

- Diercks, M., K. Johnson, G. Bar-Sever & M. Bossi. 2023. Developmental Minimalist Syntax. *lingbuzz/007134*.
- den Dikken, M. 2007. Phase Extension Contours of a theory of the role of head movement in phrasal extraction. *Theoretical Linguistics* 33(1). 1–41.
- Douglas, J. 2018. Māori subject extraction. *Glossa: a journal of general linguistics* 3(1). 110.
- Dresher, B. E. 2009. *The Contrastive Hierarchy in Phonology*. Cambridge: Cambridge University Press.
- Etxepare, R. 1996. Null Complementizers in Spanish. *Anuario del Seminario de Filología Vasca “Julio de Urquijo”* 32(2). 469–496.
- Frascarelli, M. & R. Hinterhölzl. 2007. Types of topics in German and Italian. In K. Schwabe & S. Winkler (eds.), *On information structure, meaning and form: Generalizations across languages*, 87–116. Amsterdam: John Benjamins.
- Friedmann, N., A. Belletti & L. Rizzi. 2021. Growing Trees: The acquisition of the left periphery. *Glossa: a journal of general linguistics* 6(1). 31.
- Friedmann, N. & J. Reznick. 2021. Stages rather than ages in the acquisition of movement structures: Data from sentence repetition and 27696 spontaneous clauses. *Glossa: a journal of general linguistics* 39(1). 143.
- Galasso, J. 2003. *The Acquisition of Functional Categories: A Case Study*. Indiana University, IUCL Publications.
- Gallego, Á. 2006. Phase sliding. Unpublished Ms., Universitat Autònoma de Barcelona.
- Gallego, Á. 2010. *Phase Theory*. Amsterdam: John Benjamins.
- Giorgi, A. 2010. *About the Speaker: Towards a Syntax of Indexicality*. Oxford: Oxford University Press.
- Grimshaw, J. 1991. Extended projection. Unpublished Ms., Brandeis University.
- Haegeman, L. & V. Hill. 2013. The syntacticization of discourse. In R. Folli, C. Sevdali & R. Truswell (eds.), *Syntax and its Limits*, 370–390. Oxford: Oxford University Press.
- Harwood, W. 2013. *Being progressive is just a phase: dividing the functional hierarchy*: Ghent University dissertation.
- Heim, J. & M. Wiltschko. 2021. Acquiring the form and function of interaction: a comparison of the acquisition of sentence-final particles and tag questions in the Brown corpus. Talk presented at the *LAGB Annual Meeting 2021*, 8 September.
- Hinzen, W. & M. Wiltschko. 2023. Modelling non-specific linguistic variation in cognitive disorders. *Journal of Linguistics* 59. 61–87.
- Hoekstra, T. & P. Jordens. 1996. From adjunct to head. In T. Hoekstra & B. Schwartz (eds.), *Language Acquisition Studies in Generative Grammar*, 119–149. John Benjamins.
- Jakobson, R. 1941. *Kindersprache, Aphasie, und allgemeine Lautgesetze*. Uppsala: Uppsala Universitets Årsskrift.
- Jakobson, R. & M. Halle. 1956. *Fundamentals of language*. The Hague: Mouton de Gruyter.
- Krifka, M. 2001. Quantifying into question acts. *Natural Language Semantics* 9(1). 1–40.

- Krifka, M. 2014. Embedding illocutionary acts. In T. Roeper & M. Speas (eds.), *Recursion: Complexity in cognition*, 59–87. Springer.
- Lakoff, R. 1968. *Abstract syntax and Latin complementation*. Cambridge, MA: MIT Press.
- Ledgeway, A. 2012. *From Latin to Romance: Morphosyntactic typology and change*. Oxford: Oxford University Press.
- Llinàs-Grau, M. & J. Fernández-Sánchez. 2013. Reflexiones en torno a la supresión del complementante en inglés, español y catalán. *Revista de la Sociedad Española de Lingüística* 43. 55–88.
- Llinàs-Grau, M. & A. I. Ojea. 2000. PhonBank Spanish Llinàs/Ojea Corpus. doi:[10.21415/T5K01Q](https://doi.org/10.21415/T5K01Q).
- MacWhinney, B. 2000. *The CHILDES project: The database*. Mahwah, New Jersey: Lawren Erlbaum Associates, Inc.
- Miyagawa, S. 2022. *Syntax in the Treetops*. Cambridge, MA: MIT Press.
- Paradis, J. & F. Genesee. 1997. On Continuity and the Emergence of Functional Categories in Bilingual First-Language Acquisition. *Language Acquisition* 6(2). 91–124.
- Perkins, L. & J. Lidz. 2021. Eighteen-month-old infants represent nonlocal syntactic dependencies. *Proc Natl Acad Sci U S A* 118(41).
- Perkins, L., Y. Ying, A. Williams & J. Lidz. 2021. Object wh-questions with unknown verbs are transitive for 20-month-olds. Talk presented at *Boston University Conference on Language Development* 46 (Boston).
- Pesetsky, D. & E. Torrego. 2006. Probes, Goals and Syntactic Categories. In Y. Otsu (ed.), *Proceedings of the 7th Annual Tokyo Conference on Psycholinguistics*, 25–60. Keio: Hituzi Publishing.
- Pesetsky, D. & E. Torrego. 2007. The syntax of valuation and the interpretability of features. In S. Karimi, V. Samiian & W. Wilkins (eds.), *Phrasal and Clausal Architecture*, 262–294. Amsterdam: John Benjamins.
- Potts, C. & T. Roeper. 2006. The narrowing acquisition path: From expressive small clauses to declaratives. In L. Progovac, K. Paesani, E. Casielles-Suárez & E. Barton (eds.), *The Syntax of Nonsententials: Multidisciplinary perspectives*, 183–2001. Amsterdam: John Benjamins.
- Prieto, P. & G. Rigau. 2007. The Syntax-Prosody Interface: Catalan interrogative sentences headed by *que*. *Journal of Portuguese Linguistics* 6(2). 29–59.
- Prins, K. 2014. *Overt complementizers in root clauses in southern Italian dialects*: Universiteit Leiden Master’s thesis.
- R Core Team. 2022. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria. <https://www.R-project.org/>.
- Radford, A. 1988. Small children’s small clauses. *Transactions of the Philological Society* 86(1). 1–43.
- Ramchand, G. 2023. The Symbolic Domain. Talk presented at *V-NYI 6 Winter School*, 23 January.
- Ramchand, G. & P. Svenonius. 2014. Deriving the functional hierarchy. *Language Sciences* 46. 152–174.

- Ritter, E. & M. Wiltschko. 2014. The composition of INFL. *Natural Language & Linguistic Theory* 32(4). 1331–1386.
- Rizzi, L. 1993/1994. Some notes on linguistic theory and language development: The case of root infinitives. *Language Acquisition* 3(4). 371–393.
- Rizzi, L. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Kluwer.
- Roeper, T. 2007. *The Prism of Grammar: How Child Language Illuminates Humanism*. Cambridge, MA: MIT Press.
- Roeper, T. & B. Rohrbacher. 1994. Null Subjects in Early Child English and the Theory of Economy of Projection. *IRCS Technical Reports Series* 162.
- Rosenbaum, P. 1967. Phrase structure principles of English complex sentence formation. *Journal of Linguistics* 3(1). 103–118.
- Ross, J. R. 1970. On declarative sentences. In *Readings in English Transformational Grammar*, 166–183. Waltham, MA: Ginn.
- Roy, I., B. Copley & L. McCune. 2016. The development of Merge and recursion in the transition from single words to sentences: Tracking the development of syntactic complexity in early child language. Unpublished Ms., CRNS Paris 8 and Rutgers University.
- Serra, M. & R. Solé. 1989. CHILDES Catalan Serra/Solé Corpus. doi:[10.21415/T58W3N](https://doi.org/10.21415/T58W3N).
- Song, C. 2019. *On the Formal Flexibility of Syntactic Categories*: University of Cambridge dissertation.
- Speas, P. & C. Tenny. 2003. Configurational properties of point of view roles. In A. M. Di Sciullo (ed.), *Asymmetry in Grammar*, 315–344. Amsterdam: John Benjamins.
- Travis, L. D. 2010. *Inner Aspect*. Dordrecht: Springer.
- Tsimpli, I.-M. 2005. Peripheral positions in early Greek. In *Advances in Greek Generative Syntax: In honor of Dimitra Theophanopoulou-Kontou*, 179–216. Amsterdam: John Benjamins.
- Tsimpli, I.-M. 2014. Early, late or very late?: Timing acquisition and bilingualism. *Linguistic Approaches to Bilingualism* 4(3). 283–313.
- Vainikka, A. & M. Young-Scholten. 2011. *The Acquisition of German: Introducing Organic Grammar*. Berlin: De Gruyter Mouton.
- van Kampen, J. 2009a. The Non-Biological Evolution of Grammar: Wh-Question Formation in Germanic. *Biolinguistics* 3(2-3). 154–185.
- van Kampen, J. 2009b. What's so difficult about a wh-pronoun? Poster presented at *Relating Asymmetries between Speech & Comprehension in the Acquisition of Language* (Groningen), 24 January.
- van Kampen, J. 2010. Typological guidance in the acquisition of V2 Dutch. *Lingua* 120(2). 264–283.
- Vila, I. 1990. *Adquisición y desarrollo del lenguaje*. Barcelona: Graó.
- Villa-García, J. 2012. Recomplementation and locality of movement in Spanish. *Probus: International Journal of Romance Linguistics* 24. 257–314.
- Wiltschko, M. 2014. *The Universal Structure of Categories: Towards a Formal Typology*. Cambridge: Cambridge University Press.

- Wiltschko, M. 2021. *The Grammar of Interactional Language*. Cambridge: Cambridge University Press.
- Wurmbrand, S. 2013. Tagalog infinitives: Consequences for the theory of phases, voice marking, and extraction. Unpublished Ms., University of Connecticut, Storrs.
- Wurmbrand, S. 2014. The Merge Condition: a syntactic approach to selection. In P. Kosta, L. Schürcks, S. Franks & T. Radev-Bork (eds.), *Minimalism and Beyond: Radicalizing the Interfaces*, 139–177. Amsterdam: John Benjamins.

APPENDIX: ADDITIONAL GRAPHS

Catalan

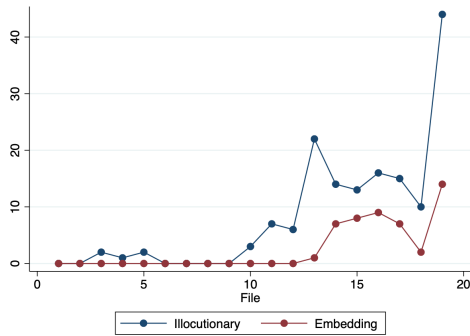


Figure 5 Laura's development.

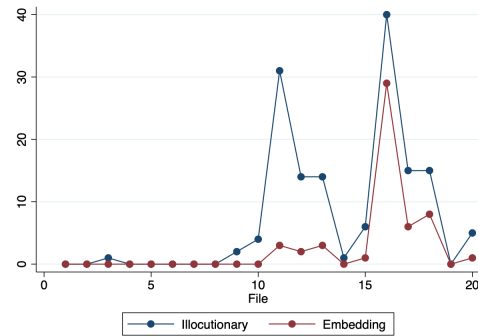


Figure 6 Gisela's development.

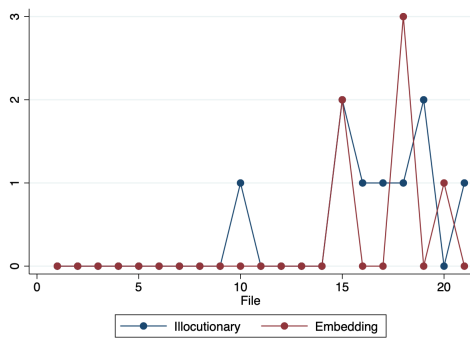


Figure 7 Àlvar's development.

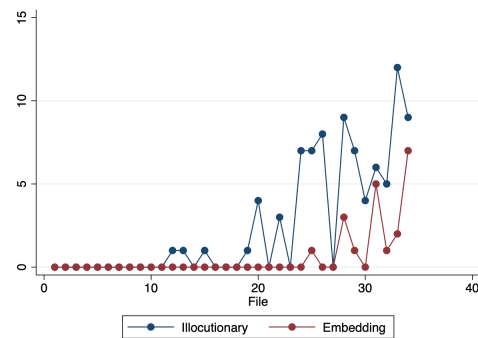


Figure 8 Guillem's development.

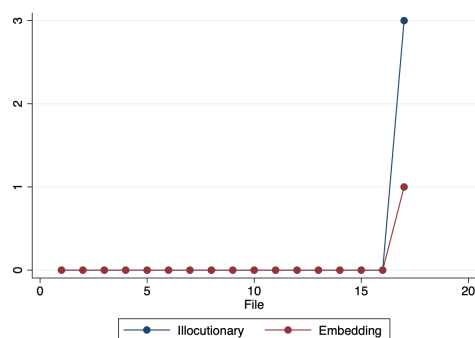


Figure 9 Júlía's development.

Spanish

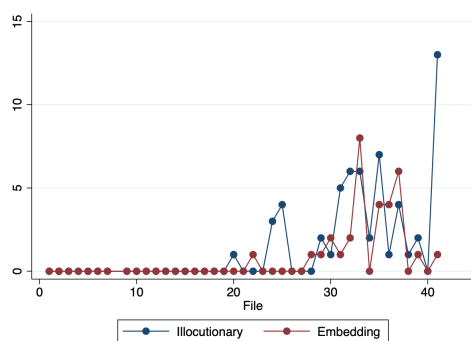


Figure 10 Irene's development.

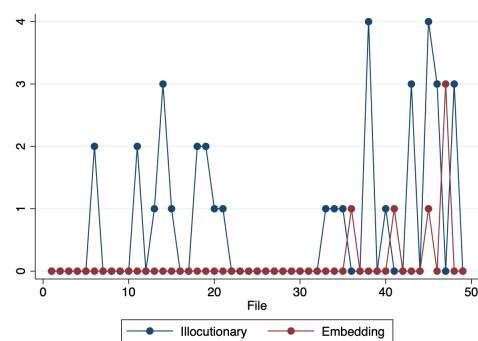


Figure 11 Yasmin's development.

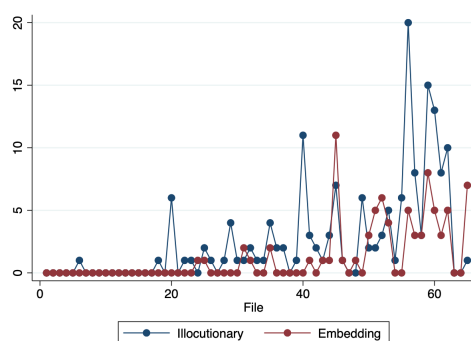


Figure 12 Juan's development.

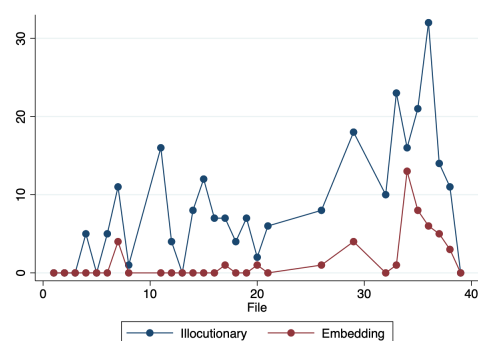


Figure 13 Magín's development.

Not All Complementisers Are Late

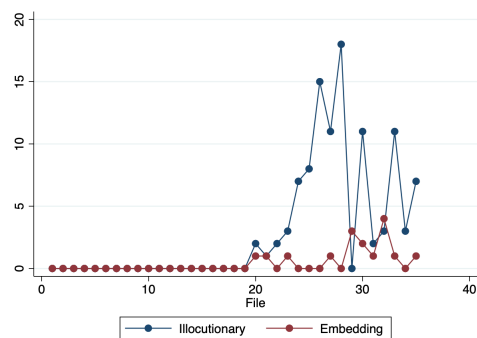


Figure 14 Emilio's development.

Núria Bosch
University of Cambridge
nb611@cam.ac.uk