UNIVERSITY OF CAMBRIDGE

Emergent Syntax and Maturation

A neo-emergentist approach to syntactic development

Núria Bosch

Supervisor: Dr Theresa Biberauer

Theoretical and Applied Linguistics
University of Cambridge

This dissertation is submitted for the degree of
Master of Philosophy

St John's College

June 2023

Declarations

Authorship. This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text.

Statement of Length. This thesis is 30,391 words in length. Excluded from this word-count are the title page, declarations, abstract, table of contents, list of figures, list of tables, list of abbreviations, figures, tables, glosses and translations, mathematical formulae and diagrams, references and the appendices.

Núria Bosch
June 2023

Abstract

This thesis explores the acquisition of functional categories from a neo-emergentist perspective (following Biberauer, 2011, *et seq.* and Biberauer and Roberts, 2015). The central proposal is that syntactic development is neither guided by a maturational program nor by innate categories. This line of thinking, which impoverishes Universal Grammar, is not just conceptually desirable (per Chomsky, 2005, *et seq.*), but enriches the explanatory power of the resulting theory in significant ways.

Firstly, the thesis presents a corpus study on ten children across five languages (Catalan, Italian, Spanish, German and Dutch) in the CHILDES database and it aims to probe the acquisition of the CP-, TP- and 'Split CP'-domains. Two key generalisations reveal themselves from this investigation: whilst CP-structures emerge early (either simultaneously with, or less clearly, earlier than TP-structures), further internal elaboration within CP (a Split CP-domain) systematically emerges late. I argue that these empirical generalisations mean no extant theory of syntactic development assuming innate categories is adequate for this dataset. This is because the findings generate theoretically contradictory requirements: they show that some representation of the CP must emerge early, but it cannot be formally cartographic-type until a later stage. Their simultaneity is inherently incompatible with contemporary maturational (and continuity) approaches, which posit universal categorial sequences of fixed (often cartographic) granularity (e.g., Friedmann et al., 2021).

I interpret this result to stress the analytical strengths of ontologically more 'flexible' approaches, which pursue emergent categories and allow for changes in mental granularity (in the spirit of Song, 2019). In light of this, I develop a neo-emergentist account of the developmental patterns: the suggested framework probes Biberauer and Roberts's (2015) emergent categorial hierarchy in the context of the acquisition of functional categories, whilst incorporating theoretical insights from Biberauer's (2019) Maximise Minimal Means (MMM) model and Bosch's (2022) adaptation within Dynamical Systems Theory. I show that this approach explains the patterns, and is both more theoretically parsimonious and more empirically restrictive relative to competing hypotheses, the latter of which tend to undergenerate or overgenerate.

The picture is completed by a formalisation of Biberauer and Roberts (2015) and categorial development, with the aid of the mathematical fields of Category Theory and Dynamical Systems Theory, adapting Ehresmann and Vanbremeersch's (2007, 2019) model of Evolutive Systems to syntactic acquisition. I suggest that such a framework provides a useful interdisciplinary bridge with the study of evolving complexity in natural systems and that some of the theoretical and neo-emergentist results revealed by the corpus study are expected given the adoption of Evolutive Systems.

This empirical and theoretical exploration endorses the pursuit of the idea that syntactic categories (including cartographic ones) are *emergent*. Insofar as neo-emergentism is shown to meet the challenge presented by the empirical paradigm in this thesis, a comprehensive treatment of emergent category formation might begin to be within sight. All that is necessary for these otherwise problematic patterns to fall into place is the adoption of a neo-emergentist approach to categories, and the abandonment of maturation and innate categories. I maintain this is the way forward, with interesting ramifications following as a result.

Acknowledgements

Being the sole author of this thesis feels like a bit of a crime; the least I can do is (attempt to) express my gratitude for the help and genuine kindness of the people that I have had the pleasure to meet over the years.

First and foremost, an immense, heart-felt thank you goes to my supervisor and former Director of Studies, Theresa Biberauer, who has done more for me and my academic growth than any stream of words could ever express. Having had her as my DoS and supervisor has been a huge stroke of luck. Since day 1 of my undergraduate studies, she has shown a constant belief and faith in me, even when I often didn't. She has poured an astonishing amount of energy and time in teaching me, very often well beyond her call of duty, and she has proved a source of support and encouragement, both academic and pastoral, that I have cherished throughout. Theresa deserves the kind of thanks that I could not fit in these paragraphs even if I tried, but if one thing is clear, it is that this thesis and whatever else linguistic I find myself producing wouldn't be possible without her. You are an amazing person and a great inspiration, Theresa: *baie, baie dankie*!

I would also like to thank the rest of my undergraduate supervisors and lecturers, who have made my linguistics training very stimulating. Here, I also owe particular debt to Ian Roberts, whose syntax lectures in my first and second years and talks/seminars at the department greatly helped spark my interest in syntax and shape my thought. Big thanks also go to the SyntaxLab group, for filling my Tuesday afternoons with fun syntactic puzzles over the past four years and for the occasional post-talk pub trip. Special thanks to Dora Alexopoulou, for raising some important questions during my Lent term talk that set me on the right track, and to Fanghua Zheng and Liu (Willow) Yang, for their additional questions.

I am also grateful to Giuseppe Cappelli, Naama Friedmann, Adam Ledgeway and Laurel Perkins, who helpfully answered my questions via email, and to Jeffrey Lidz, Colin Phillips and Juan Uriagereka, for discussing some of the topics in this thesis. Petra Mišmaš's Cartography course at the EGG 2022 summer school also provided me with useful references and food for thought; thanks are due to her and everyone else who attended the school. Ashton Brown's statistics courses via the SSRMP and her prompt email replies to my questions were also highly helpful. I am also very thankful for my email exchanges with Itamar Shatz, who kindly set aside time to help me with some of my statistics doubts whilst I was writing up the thesis. Thanks, as well, to Sara Cardullo, Caja Göbel and Juliette van Steensel for proof-reading the Italian, German and Dutch examples in the thesis. Of course, no-one mentioned here is responsible for any remaining inaccuracies, which are entirely my own.

Unfortunately, 2022 and 2023 have also been the toughest years I've seen, having had to contend with the ongoing cognitive and physical consequences of Long COVID since January 2022. I'm greatly thankful here to my neurologist back home, Dr Munmany, for helping me stay more or less sane throughout this year-and-a-half. It goes without saying that this thesis would almost certainly not have existed had we not mitigated my symptoms from the start. Thanks also to Theresa for bearing with me while I've been juggling this with my work, and to my family for crucial support at multiple points. My heart goes out to other people facing this poorly understood condition — praise be to the few medications that do part of the trick and here's hoping things get better for us all.

This thesis was generously supported by a Hayes Scholarship from St John's College and a Cambridge Opportunity Master's Studentship from the Cambridge Trust, which I was extremely lucky to receive. I would like to extend my gratitude to my college, for constant financial support from the very start of my undergraduate degree, and my appreciation to the community and staff at St John's, for their friendliness throughout my four years here.

A warm thank you to my friends and other fellow linguists for making my time at Cambridge so enjoyable and energising. Big thanks also to those in Catalonia, and to those elsewhere. Finally, thank you to my parents and family, for their constant support and belief in me and for always prioritising my health and education above everything else, even when the latter unexpectedly took me a bit far away from home. *Moltes gràcies!* And thanks to everyone else, friends, former teachers, and any others, who have lent a helping hand and placed confidence in me at some point, big or small.

Table of contents

List of figures

List of tables

List of abbreviations

◯	Undivided categorial space
1	1st person
2	2nd person
3	3rd person
ACC	Accusative
AUX	Auxiliary
B&R	Biberauer and Roberts (2015)
CAS	Complex Adaptive System
CFC	Core Functional Category
CL	Clitic
CLD	Contrastive Left-Dislocation
CLLD	Clitic Left-Dislocation
COND	Conditional
CONJ	Conjunctive
DAG	Directed Acyclic Graph
DAT	Dative
DIM	Diminutive
DMS	Developmental Minimalist Syntax
DO	Direct object
DST	Dynamical Systems Theory
E&V	Ehresmann and Vanbremeersch (2007)
Embed	(Finite) Embedding markers
EP	Extended Projection
EPP	Extended Projection Principle
ES	Evolutive System
EXCL	Exclamative
FE	Feature Economy

FEM	Feminine
FG	Feature Generalisation
Foc	Foci/Focalisations
FUT	Future
GER	Gerund
HES	Hierarchical Evolutive System
HTLD	Hanging Topic Left-Dislocation
IG	Input Generalisation
Illoc	Illocutionary complementisers
IMP	Imperative
IMPF	Imperfective
IMPRS	Impersonal
INF	Infinitive
INT	Interrogative
IO	Indirect object
L1	First Language
LOC	Locative
MASC	Masculine
MLU	Mean Length of Utterance
MMM	Maximise Minimal Means
NOM	Nominative
PART	Partitive
PASS	Passive
PERF	Perfect
PERIPHR	Periphrastic
Ph	Phase
ϕ**-features**	Person, number and gender features
PL	Plural
PLD	Primary Linguistic Data
PRES	Present
PRFV	Perfective
PRON	Pronoun
PST	Past
PTCP	Participle
QUOT	Quotative
REFL	Reflexive

RI	Root Infinitive
SBJV	Subjunctive
SCH	Strong Continuity Hypothesis
SG	Singular
SID	Southern Italian Dialect
Split	Split CP
TMA	Tense, Modality and Aspect
Top	Topics/Topicalisations
UG	Universal Grammar
V1	Verb Initial
V2	Verb Second
V3	Verb Third
Wh-Q	Wh-questions
Y/N-Q	Yes/no questions

1 | *Introduction*

In their analysis of the acquisition of English and Hebrew passives, Borer and Wexler (1987) established a direct relationship between biological maturation and syntactic development. In analogy with the biological programs that guide the development of, for instance, physical changes (e.g., the onset of puberty), the suggestion is that linguistic principles and categories, such as clausal spines, 'mature' in the same way, also guided by an underlying hard-wired biological mechanism. Their development is thus not contingent on learning: the principles are *not* learned and so are not directly dependent on children obtaining evidence. So long as these linguistic properties are part of our innate competence, the hypothesis that they mature can reasonably be upheld.

If, however, these primes are eradicated from the faculty of language (as appears to be the trend within the Minimalist Program; see, particularly, Hauser et al., 2002; Chomsky, 2005, *et seq.*), then the contentfulness of 'maturation' is in reality appreciably moot, casting initial doubt on the foundations of several contemporary acquisition hypotheses (viz. Friedmann and Reznick, 2021; Friedmann et al., 2021; Heim and Wiltschko, 2021; Rakhlin and Progovac, 2021). These matters are fundamentally conceptual in nature. As such, they have to be weighed against the empirical data. Therefore, the theoretical necessity of maturation, if even required at all, and its contentfulness must be established based not just on the biological and ontological foundations of language assumed in present-day Minimalism, but most importantly also on the explananda emerging from the crosslinguistic empirical data. What is, however, less often appreciated is the possibility that the developmental data itself can shed some light on questions surrounding the ontological bases of linguistic knowledge.

Against this background, this thesis studies aspects of the acquisition of functional categories, paying particular attention to maturational approaches and juxtaposing them to so-called *neo-emergentist* proposals. This attempt is instigated by several bigger-picture questions:

1. Do children draw upon a prior (innate) ontology of syntactic categories or are these constructed *ab nihilo* during the learning task?

2. If innate, to what extent is syntactic category formation directed by a biological program (maturation)?

3. If emergent, what developmental stages do these categories transit through and are innate templates nonetheless required as the basis for emergent categories (as in Ramchand and Svenonius, 2014; Wiltschko, 2014)? Additionally, how and when do children acquire the rich functional hierarchies posited typologically (e.g., Rizzi, 1997; Cinque, 1999; Svenonius, 2006, 2008)?

4. How can the empirical data give clues as to the answers to (1-3)?

Through a corpus study, I present a re-examination of the role syntactic maturation should play in the context of Chomsky's (2005) Three Factors approach. I hold that, on the basis of the developmental patterns observed, maturation[1] (and continuity) approaches to acquisition fall foul of accounting for the data — because they make unduly rigid assumptions about the ontological flexibility of syntactic categories (see also Song, 2019). I suggest, then, that, on both empirical and conceptual grounds, maturation and innate categories appear unfeasible. In line with neo-emergentism, empirical gain is achieved by reducing Universal Grammar's role, with crucial aspects of categorial development transpiring outside its domain.

In **Chapter 2**, I begin by introducing the theoretical background of this thesis. Subsequently, I present the neo-emergentist proposal I will advocate, which probes the learning path in Biberauer and Roberts (2015) (see also Biberauer, 2019) and incorporporates insights from Dynamical Systems Theory (DST) as outlined in Bosch (2022). I finish the chapter by outlining the main predictions of each approach to be tested in the corpus study. **Chapter 3** presents the corpus study, summarising its results and their wider theoretical significance. In **Chapter 4**, empirical and syntactic concerns meet with mathematical formalisation as carried out in branches of Category Theory and DST (Ehresmann and Vanbremeersch, 2007, 2019), with the aim of ratifying the acquisitional model proposed from the wider perspective of natural and biological systems. This final biolinguistic exploration reveals that predictions made by DST-based neo-emergentism suggestively converge with Ehresmann and Vanbremeersch's (2007) model of Evolutive Systems. A combination of Biberauer and Roberts's (2015) and Biberauer's (2019) results with Evolutive Systems is shown to further elucidate our view of developing grammars, categorial systems and their distinctive traits in a neo-emergentist context. **Chapter 5** concludes the discussion.

[1] *Syntactic* maturation as first introduced in Borer and Wexler (1987) needs to be distinguished from the broader *cognitive* maturation. The former is the present focus.

2 | *Approaches to syntactic development*

Traditionally, approaches to syntactic development have been divided into two main kinds depending on the amount of innate linguistic structure assumed to be accessible to the child at the start of the learning path: (i) *continuity* approaches, claiming that the functional structure of children's initial grammar is not significantly different from adults' grammars and (ii) 'incomplete structure' or *maturational* approaches, arguing that children's grammatical knowledge is not (fully) available initially and that grammatical categories and principles appear gradually. Universal Grammar (UG; or some equivalent biological mechanism) specifies the order in which innate syntactic projections mature.

This chapter is intended as a selective overview of generative approaches to the acquisition of syntax and their theoretical assumptions about clause structure and the nature of early child grammars. Although both continuity and maturation-based proposals will be introduced (§2.1 and §2.2), maturation will be our focus, for reasons to be made clear.

In §2.3, however, I outline the first motivations for scepticism towards maturation as a biolinguistically likely approach in a Three Factors context (Chomsky, 2005). §2.4 introduces neo-emergentist approaches as alternatives to maturation, which posit instead emergent categories. There, attention will centre on Biberauer's (2011, *et seq.*) Maximise Minimal Means model and the emergent categorial hierarchy proposed in Biberauer and Roberts (2015). §2.5 presents the approach I will argue in favour of — Dynamical Systems-based neo-emergentism —, which integrates Biberauer and Roberts (2015), Biberauer (2019) and Bosch (2022) and is developed further in §4. These final three sections foretell some of the theoretical and empirical arguments to follow in subsequent chapters, thereby setting the foundations for this thesis' main contentions. Finally, §2.6 summarises the differences between the approaches presented and their predictions.

2.1 Continuity

Continuity proposals claim that the functional structure of children's initial grammars is continuous with adult grammars. Versions of this perspective can range from very strong to weak depending on how identical children's and adults' syntactic representations are hypothesised to be.

On the one hand, the Strong Continuity Hypothesis (SCH) argues that a child's earliest grammar contains the full inventory of functional categories (broadly, *v*P, TP and CP), irrespective of whether the function morphemes and operations associated with them surface in production (i.a., Pinker, 1984; Hyams, 1992, 1994; Verrips and Weissenborn, 1992; Poeppel and Wexler, 1993). Despite highlighting that a 'small clause' and an exclusively lexical (VP) stage analysis (Radford, 1988, 1990) does not do justice to data from several 2-year-olds (e.g., Hyams, 1986; Guasti, 1993), it was soon noted that SCH makes too categorical a claim about the initial status of children's representations. Specifically, SCH is at odds with several experimental and corpus studies indicating selective and gradual development of (at least some) functional heads throughout the first years of life, with lexical acquisition often playing a central role (see Serratrice, 2000; Tsimpli, 2005, for some discussion). In practice, then, SCH typically still needs to resort to some additional mechanism (e.g., 'underspecification' of functional heads in Hyams, 1996) to account for selective unavailability of certain structures early in development. Additionally, their absence often cannot be ascribed to phonological reduction or prosodic licensing (see Mitrofanova, 2018; Tedeschi, 2009, on comprehension deficits in prepositional constructions in Russian and clitic omission in Italian, respectively). In this latter respect, then, SCH ends up converging with maturational proposals: both recognise that there are qualitative *syntactic* differences between adult and child grammars, and that these discrepancies cannot easily be accounted for by assuming a fully adult-like spine from the get-go.

To address some of these shortcomings, a weaker version of Continuity, the Lexical Learning Hypothesis (Clahsen et al., 1994, 1996), claims, instead, that, although functional categories are in principle available throughout acquisition via UG (thus retaining the essence of a continuity approach), this continuity is only 'weak', insofar as the instantiation of these functional categories is *not* independently licensed (unlike in SCH). It is only granted once children acquire the words, features or morphemes associated with those projections in their L1s. Therefore, the projection of functional categories and the specification of category features in given language grammars are seen as input-driven, stepwise processes (Paradis and Genesee, 1997, p. 94).

Because of the aforementioned issues with SCH, and because current literature centres primarily on testing the predictions of 'incomplete structure' perspectives, attention here will be devoted to maturation. In §3.3, though, I will come back to Weak Continuity.

2.2 Maturation

2.2.1 Bottom-up maturation

The dominant perspective in contemporary and also traditional maturational proposals is a bottom-up approach. Current models inherit logics akin to earlier proposals, such as Radford's (1988, 1990) Small Clause Hypothesis and Rizzi's (1993) Truncation Hypothesis, which traditionally account for non-target-consistent structures in child language such as root null subjects, lack of auxiliaries and root infinitives, by assuming the earliest stages contain only a *v*P/VP. It is only in subsequent developmental stages that TP and, later, CP mature.

Recently, this perspective has been endorsed by the so-called *Growing Trees Hypothesis* (outlined in Friedmann et al., 2021, and Friedmann and Reznick, 2021). Because these accounts presuppose a rich Universal Grammar, often containing a cartographic spine, the reasoning maintains that lower layers in the syntactic structure cannot be omitted if higher ones are present. 'Dangling' representations with only projection A and projection C, missing UG-specified intermediate layer B, as in **Figure 2.1**, are therefore unavailable (where C dominates domains A and B in the adult grammar):

Figure 2.1 An impossible dangling representation (Friedmann et al., 2021, p. 2)

Growing Trees specifically suggests (unlike its precedent, Truncation) that higher layers of the tree are *not* yet available at earlier stages (under Truncation, higher layers of the structure were available, though not always used)[1]. It is thus maturational mechanisms that dictate the projections available at each point in development, and these stages follow the geometry of a UG-given adult-like syntactic spine, proceeding in a bottom-up manner, as shown in **Figure 2.2**. At Stage 1, only IP/TP and VP are available to the child, permitting, for instance, the production of inflection and A-movement in SV orders with unaccusatives.

[1] Rizzi's (1993) Truncation is maturational insofar as the axiom that root clauses must be full CPs is assumed to be initially unavailable due to maturation (see Rizzi, 1993, fn. 4 for discussion on how maturation is integrated into the model). This means the IP/TP- and CP-layers are not always fully projected and so are *optionally* absent at early stages.

A later Stage 2 sees the development of the lower left periphery, up to QP (hosting, i.a., wh-questions). The entire cartographic hierarchy, from TopP until ForceP, becomes available at Stage 3, when, e.g., topicalisations and embedding are expected to appear[2].

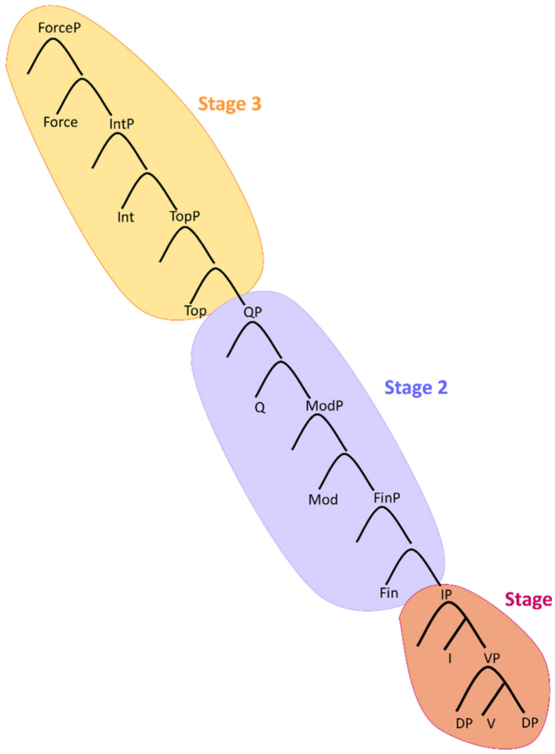


Figure 2.2 Stages of acquisition of the clausal domain in the Growing Trees Hypothesis (Friedmann et al., 2021, p. 12)

Diercks and Bossi (2021) and Diercks et al. (2023) outline a similar proposal — *Developmental Minimalist Syntax* (DMS). The general model proposed is one where structurally lower elements are acquired before structurally higher ones. DMS has at its core an interpretive proposal whereby the syntactic analysis of adult syntactic structures has implications for our understanding of language development: there is a correlation with the bottom-up minimalist derivation of syntactic trees and the timeline with which structures become acquired in child L1 acquisition (see 1). In this case, however, development is *non*-maturational: categories are *not* innate, instead minimalist derivations of sentences 'replay' the acquisition

[2]Unlike Radford (1988, 1990), Stage 1 already contains functional categories (IP/TP). Friedmann et al. (2021, p. 27) consider the possibility of an earlier phase with a bare VP lacking IP, ultimately concluding there is no evidence for it in their data as inflection is present (age range 1;6-6;1). Note, too, that in the full clausal spine in **Figure 2.2** only the left periphery is cartographic. How the child would acquire more elaborate versions of the Tense-Aspect-Mood domain (see also Cinque, 1999, cf. too De Lisser et al., 2017) or of the *v*P-domain and event-structure (e.g., Ramchand, 2008) remains unaddressed.

of those structures, predicting a general *v*P/VP → TP → CP acquisition timeline. Including DMS here is nonetheless important, as we will come back to DMS in §3.

(1) **Developmental Minimalist Syntax (an interpretive principle):**
The Minimalist derivation of adult language structure recapitulates the ontogenetic (i.e., organism-internal) development of those same syntactic structures (Diercks et al., 2023, p. 13).

A range of evidence has been argued to support bottom-up development since the 1980s, when Radford's (1988, 1990) Small Clause Hypothesis first emerged (see also Vainikka and Young-Scholten, 2011, on Organic Grammar, and Rakhlin and Progovac, 2021 on the Gradual Emergence theory of syntax). Most recently, empirical data from Hebrew and Brazilian Portuguese (Friedmann and Reznick, 2021; Meira and Grolla, 2022), which suggests that the main bulk of A-movement structures are acquired before most A-bar structures, has been interpreted as support for the Growing Trees Hypothesis. Equally, however, a collection of data that is not fully consonant with the approach has also been noted. For instance, De Lisser et al. (2017) show that the ordering of emergence of cartographic heads in the Tense-Aspect-Mood field in Jamaican Creole does not transparently correspond to a bottom-up maturational pathway and Tsimpli (2005) provides data from Greek that problematises bottom-up proposals. The perspectives outlined in §2.2.2 and §2.4 discuss similarly contradictory empirical data.

The non-conforming empirical conclusions beg the question of what the root of these discrepancies is: for example, if it comes down to the different diagnostic criteria and/or theoretical assumptions used in various studies, theoretical proposals that hold over distinct stages in acquisition (e.g., with respect to emergence or mastery), studying children at distinct age ranges and thus, possibly, different syntactic stages, or something else entirely. It is still an open question, then, if bottom-up maturation is empirically validated and, additionally, if this perspective holds true for emergence (first appearance of structures), mastery/stabilisation (as in Friedmann and Reznick, 2021), both or neither.

2.2.2 Inward maturation

Inward-growing approaches have received much less attention than the dominant bottom-up approaches. Simplifying grossly, these can be defined as approaches that assume development to begin at *v*P/VP and CP (the clausal 'edges', from the perspective of the CP>TP>*v*P functional sequence), before maturation of the TP.

Arguably, the first maturational approach that, in retrospect, argues for this progression can be traced back to Galasso's (2003) 'Empty Middle' approach for child English, which proposes a revision of Radford's Small Clause hypothesis (CP>VP to CP>IP>VP). The approach assumes a VP structure for declaratives (with Radford), CP>VP for basic interrogatives at his stage 1 and, finally, the full blown CP>IP>VP at stage 2, when inflection, auxiliaries and inversion start being acquired (notwithstanding possible underspecification; Galasso, 2003, p. 156). Roeper and Rohrbacher (1994) also proposed an analysis of some early wh-questions with root infinitives in English which reflects a somewhat similar logic to Galasso's. Partly incorporating Rizzi's (1993) Truncation mechanism and, particularly, Speas's (1994) Economy of Projection, they suggest that examples such as *what doing?* or *what you doing?* (Adam.07) require a CP>VP analysis, as the only projections satisfying Economy of Projection are CP and VP. Although neither of these proposals was widely adopted, previous literature does recognise that some child utterances might necessitate a departure from the now-standard bottom-up derivation of clauses. §2.4 will briefly return to non-maturational approaches that also advocate early acquisition of discourse and interactional knowledge.

More recently, however, Heim and Wiltschko (2021) outline an explicit inward-growing maturational approach to the acquisition of the universal and interactional spine in Wiltschko (2014, 2021). Heim and Wiltschko (2021) argue that the interactional and classification-based layers (the latter corresponding, roughly, to *v*P) mature first. This hypothesis draws on two associated models of clausal organisation: the Universal Spine Hypothesis and the Interactional Spine Hypothesis, which we present separately below (following and adapting the exposition in Bosch, accepted, pp. 14-15). These are jointly displayed in **Figure 2.3**.

First, Wiltschko (2014) posits a universal spine of functional categories. Each layer is associated with a universal function, vital for establishing reference and propositional meaning, but the substantive content of each category is language-specific. The lowest layer, *Classification*, 'catalogues' events and individuals into subcategories (e.g., telic *vs* atelic events; mass *vs* count nouns). *Point of View* (PoV) associates the classified event or individual with a particular perspective (e.g., in Indo-European languages, PoV is substantiated by aspect). *Anchoring* relates the event or individual to the utterance context (e.g., via tense or, in other languages, person or location; see Ritter and Wiltschko, 2014). Lastly, *Linking* connects the anchored event or individual with the discourse (hosting, among others, wh-elements, topics and other elements associated with the CP-domain). Broadly, this *Universal Spine Hypothesis* defends a *substantivist* view: while grammatical categories are built on a language-specific basis, the spine restricts the types of categories that languages construct and their hierarchical ordering.

The universal spine has recently been developed in Wiltschko (2021) to incorporate a syntactic encoding of speech-act information, following neo-performative approaches (Ross,

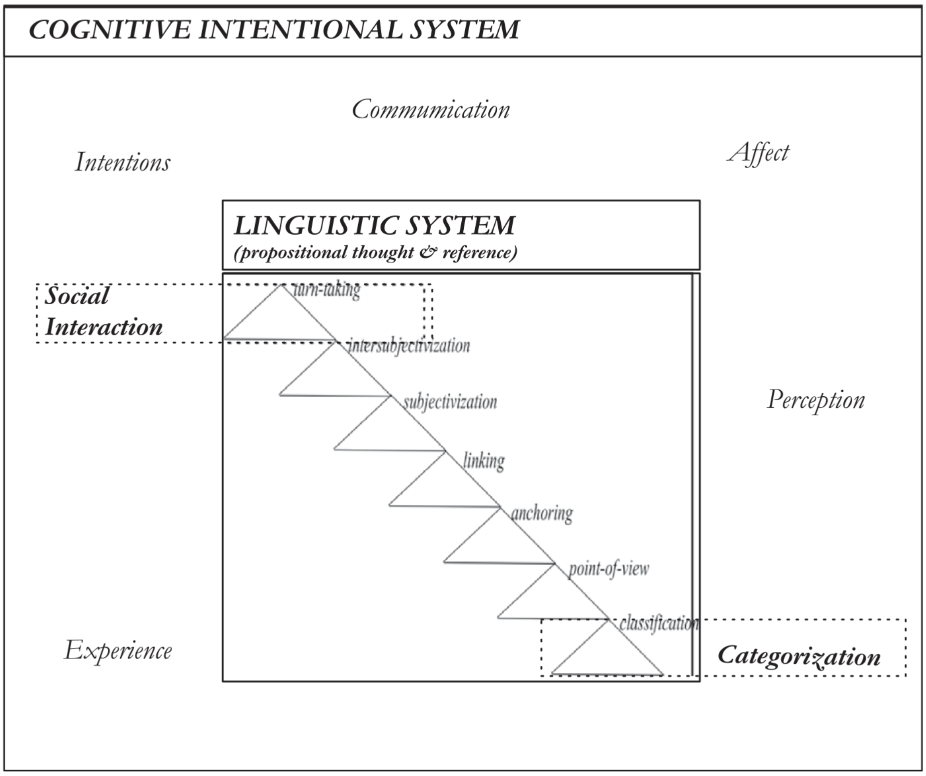


Figure 2.3 The Bridge Model (Hinzen and Wiltschko, 2023, p. 75).

1970, and subsequent work). This model is referred to as the *Interactional Spine Hypothesis*. The framework extends the universal spine by appending two additional syntactic domains dominating propositional structure (i.e., above Linking, Anchoring and Classification; from CP to *v*P). Within this higher conversational domain, a lower *Grounding/Grouping* layer, divisible into $\text{Ground}_{Addressee}$ and $\text{Ground}_{Speaker}$, and a *Responding* layer are postulated (the intersubjectivisation, subjectivisation and turn-taking layers in **Figure 2.3**, respectively). The Grounding layer allows "the speaker to configure the propositional content of the utterance so that the addressee can update their knowledge state to include it" (Wiltschko, 2021, p. 72). Finally, the higher Response layer manages "the moves that serve to synchronize the interlocutors' knowledge states" (Wiltschko, 2021, p. 72), charting the interactive component among participants in conversation.

Against this background, Heim and Wiltschko (2021) propose the *Inward Growing Spine Hypothesis*, according to which the interactional and universal spine develop inwardly, starting from the edges of the spine and developing towards the centre (see also Wiltschko, 2021). Hinzen and Wiltschko (2023) present a model of language and cognition, the Bridge Model, which views linguistic cognition as hinging on two partially pre-linguistic pillars (perceptual categorisation and socio-communicative interaction); grammar acts as a bridge that mediates both (**Figure 2.3**). In this context, then, the thematic and categorisation-based domain (Wiltschko's, 2014, Classification head) and the interactional and speech-act-based domain (Wiltschko's, 2021, Responding head) are hypothesised to develop first. Development and maturation start in these two philogenetically-prior domains. Subsequently, Linking (hosting, e.g., wh-words and complementisers) matures. A third step involves the development of GroundP, with Anchoring maturing to accommodate tense and other deictic elements (e.g., demonstratives for the nominal domain). At a final stage, GroundP is subdivided into $Ground_{Addr}$ and $Ground_{Spkr}$, and Anchoring into Anchoring and PoV.

Data on the acquisition of English *huh* is presented as preliminary evidence for part of Heim and Wiltschko's (2021) Inward Growing Spine Hypothesis. Initially, the children studied harness *huh* only to request a response (RespP; from ∼ 2;05). A second stage (from 3;05) contains both RespP and GroundP, when *huh* also marks a given proposition *p* as grounded. The final stage 3 unfolds once *huh* further distinguishes between speaker/addressee ground (from 4;05), being able to confirm the speaker's ground and/or the addressee's. The resulting picture is, they suggest, one of 'fractalian' development[3], with maturation enacting not just an inward-growing spine, but also one which successively subdivides some projections into subtyped ones (e.g., GroundP into $Ground_{Addr}$ and $Ground_{Spkr}$).

The proposal is programmatic and only the predictions regarding interactional language and associated projections have been tested. Additionally, its generalisability to very early developmental stages remains open: Heim and Wiltschko's (2021) dataset focused on late MLU (Mean Length of Utterance; typically $> 2.0 - 2.5$ MLU) and age ranges, and included the well-known late-talkers, Adam and Sarah (Brown corpus). Nonetheless, given the data in Heim and Wiltschko (2021), which is consistent with the Inward Growing Spine Hypothesis, and the earlier-mentioned evidence in Galasso (2003) and Roeper and Rohrbacher (1994), some version of an inward maturational pathway, as an alternative to a bottom-up approach, is clearly worth considering.

Since the present study will not investigate interactional language directly, the Inward Growing Spine Hypothesis will not be tested. However, in §2.6, I formulate a more general inward-growing hypothesis that abstracts away from interactional language for present purposes and whose predictions will be analysed.

[3]Two senses of 'fractalian' are intended here: namely, in the successive-division sense above (*not* the standard mathematical sense of a fractal), but also in a mathematically-inspired interpretation, which assumes parallel development across syntactic domains (verbal and nominal; see Wiltschko, 2014).

2.3 Against maturation

An appealing strength of maturational approaches lies in their ability to account for successive developmental stages, typically a challenge for Strong Continuity. Yet, whether this superficial advantage is achieved in a way that is both principled and free from stipulations is debatable (as Serratrice, 1996, already notes).

The theoretical and empirical consequences of assuming biological maturation alongside an (often very rich) set of syntactic categories become apparent in the context of Chomsky's (2005) Three Factors approach and, more generally, in light of the Minimalist desideratum of minimising the explanatory burden placed on domain-specific sources. Chomsky (2005) distinguishes three factors contributing to the growth of language systems: the genetic endowment (Factor 1), the Primary Linguistic Data (PLD) or intake (Factor 2) and the third Factor, which is understood as, i.a., "principles of data analysis [...] and efficient computation" (p. 6), such as learning biases. In making this tripartite contrast, the aim is to shift the explanatory burden onto Factor 3 and the crucial *interaction* between the factors (cf. Lewontin, 2000), in turn shrinking the content of Factor 1 and the technology postulated as language-specific. Factor 1 is assumed to contain, for example, only Merge in Hauser et al. (2002) or Merge, Agree and, possibly, some notion of formal feature in Biberauer (2019, *et seq.*). If UG is evolutionarily recent and poor (Chomsky, 2004, *et seq.*), the result is that it cannot contain totally-ordered sequences of tens of functional categories nor innate sets of parameters (Biberauer and Roberts, 2015; Biberauer, 2019; Roberts, 2019; see also Leivada and Murphy, 2022).

Not only does maturation presuppose innate categorial priors, it also requires a hard-wired biological mechanism specifying the order of acquisition of those representational primes (though cf. Gibson and Wexler, 1994, for a different implementation), placing even more burden on Factor 1. If this minimalist perspective on language as a 'biolinguistic organ' along with other cognitive systems is seriously heeded, then, it renders maturational hypotheses potentially obsolete, by virtue of eliminating the very foundations on which maturation rests (innate categories). This therefore invites a reappraisal of the premises on which current acquisitional hypotheses rely. A second argument is one of parsimony: if the developmental data can be explained without postulating maturation of UG-internal principles, then there is no need to do so and it suggests the need for a deeper explanation. Maturation shouldn't be the null hypothesis. This Three Factors perspective, however, presents us with a further, arguably more interesting question, namely whether 'dropping' innate categories and, thus, the maturation assumption helps us explain any empirical facts that remain otherwise elusive under both.

Empirical arguments specifically against maturation and approaches assuming innate ontologies have not been at the centre of the acquisition literature (though see Paradis and Genesee, 1997, for one exception regarding the former). Typically, the implicit assumption has been that an account of the patterns relying on innate categories is always available (e.g., some version of bottom-up maturation, continuity or, more recently, inward maturation). The emphasis has been placed on determining which one is on the right track. In §3, I will challenge the innocence of this premise and present a novel empirical argument *against* maturation and *for* emergent syntactic categories.

With these initial caveats in place, I now go on to summarise already-existent neo-emergentist approaches, which are non-maturational in nature and instead argue for emergent categories.

2.4 Neo-emergentism: a generative approach to emergent categories

In this thesis, I investigate instead what I see as a way forward for the theory of syntactic variation and language acquisition, namely one which pursues Chomsky's (2005) Three Factors approach and the desideratum of making minimal assumptions about the inventory of UG. This section, therefore, considers and summarises the possibility that development is best modelled via *emergent* syntactic categories, specifically as advocated in so-called *neo-emergentist* approaches to syntactic development (Biberauer's, 2011, *et seq.*, Maximise Minimal Means model) and in Biberauer and Roberts (2015).

Capitalising on the consequences and opportunities that a minimal UG brings, Biberauer (2019) endorses a Three Factors approach. She proposes an explicit third-factor principle and neo-emergentist model, Maximise Minimal Means (MMM), which aims to make utmost use of minimal resources. MMM is suggested to have (minimally) two independent linguistic manifestations, which regulate the postulation of emergent formal features ([F]s):

1. **Feature Economy** (FE; generalised from Roberts and Roussou, 2003)
   Postulate as few [F]s as possible to account for the PLD.

2. **Input Generalisation** (IG; adapted from Roberts, 2021; termed *Feature* Generalisation in Biberauer, 2020)
   Maximise available [F]s.

As a general-cognitive principle, however, its reflexes are also proposed to surface across multiple cognitive domains (Biberauer, 2019, 2020; Biberauer and Bosch, 2021).

In this connection, Biberauer (2019) suggests that $[F]$s are postulated if they fulfil a *contrastive* role in the system and highlights the role of 'systematic departures from Sausurean arbitrariness' (that is, departures from the straightforward one-to-one mappings between form-meaning that underpin core lexicon) in signalling $[F]$s. These $[F]$-cuing 'departures' or contrasts can include systematic silence (e.g., null arguments/exponents, where there appears to be meaning, despite absence of form), multifunctionality (one form, with place-dependent multiple meanings) and movement (where displacement often gives rise to additional interpretations), among several others (see Biberauer, 2019, pp. 54-57). This 'contrastivity' requirement follows previous work in phonology (originating in Jakobson, 1941; Jakobson and Halle, 1956), according to which features are postulated if, and only if, they are required to capture a contrast in the system, see Dresher's (2009, 2014) Successive Division Algorithm, Hall's (2007) Contrastivity Hypothesis and Cowper and Hall's (2014) *Reductiō ad discrīmen.*

The role of MMM within the broader Three Factors model is schematised below (adapted from Biberauer, 2019, p. 49):

(2) (Maximally poor) Universal Grammar + Primary Linguistic Data (PLD) + MMM $\longrightarrow$ Adult grammar

Given (2) and the adherence to Chomsky (2005), neo-emergentism retains several elements of a standard minimalist toolkit, shared with modern maturational approaches like Friedmann et al. (2021). For instance, the derivational machinery (Merge, Agree) and the representational formalism, such as cartography (at least as set out in (3) below), albeit the latter hypothesised to be emergent. Additionally, neo-emergentism takes parametric variation to be encoded as featural specifications on functional heads (see the Borer-Chomsky Conjecture in Baker, 2008). In these respects, then, the approach remains *generative* in essence and furthermore keeps some nativist assumptions (a maximally poor UG, with Merge, Agree and some notion of $[F]$). These theoretical assumptions, which make systems $[F]$- and Merge-regulated, set the MMM approach apart from traditional emergentist approaches such as Construction Grammar and usage-based theories in non-negligible ways, hence the prefix *neo*-emergentist (see Biberauer, 2019, p. 59).

A goal of this thesis is to contribute to this line of theoretical inquiry. Therefore, I will adopt a similarly Three-Factors- and MMM-inspired viewpoint, which assumes emergent representational primes (following, i.a., Mielke, 2008; Dresher, 2009; Ramchand and Svenonius, 2014; Biberauer and Roberts, 2015; Biberauer, 2019; Cowper and Hall, 2022; Samuels et al., 2022; Scheer, 2022; Ramchand, 2023). Most relevantly for us, I will test Biberauer and Roberts's (2015) perspective on *emergent categories*, which I summarise now.

Firstly, according to Biberauer and Roberts (2015), crosslinguistically, clauses can be analysed at different levels of 'magnification' or 'granularity'. In the interests of clarity, I define granularity below, following Song (2019, p. 21):

Definition 1 ((Syntagmatic) granularity)**.** A sequence of combinatorially oriented categories may be collapsed into a single category (e.g., $X - Y - Z \longrightarrow A$), and a single category may be split into a sequence of categories (e.g., $X \longrightarrow X_1 - X_2 - ... - X_n$).

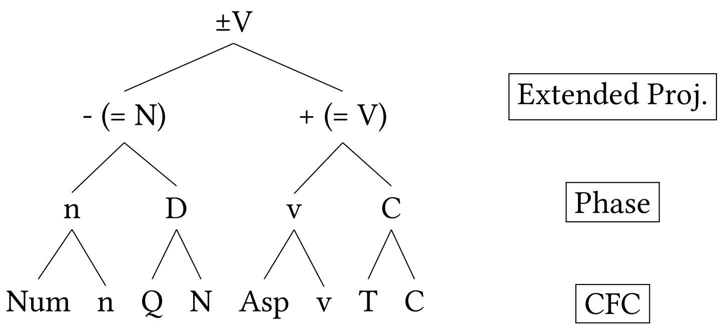
For example, the relation between the Core Functional Category C (Chomsky, 2000, 2001) and the cartographic sequence Force–Topic–Focus–Fin (Rizzi, 1997) is one of (syntagmatic) granularity; the former can be split into the latter and, vice versa, the latter can be collapsed into the former. The major granularity levels I assume throughout this thesis are the following (Biberauer and Roberts, 2015, p. 4), from less to more fine-grained (as the nominal domain is orthogonal to the present study, I restrict exposition to the clausal domain):

(3) **Extended Projection** (V) > **phase** (C, v) > **Core Functional Category** or **CFC** (C, T, v) > "**cartographic field**" (e.g. Tense, Mood, Aspect, Topic, Focus) > **semantically distinct head** (as in Cinque, 1999; Speas and Tenny, 2003; Frascarelli and Hinterhölzl, 2007; Haegeman and Hill, 2013, among others).

At the highest level of organisation, we find the Extended Projection (EP) in the sense of Grimshaw (1991, *et seq.*), referring to a maximal projection which projects from a category that it shares categorial features with. Thus, the whole clausal EP is C–T–V, with all categories sharing the features $[+V, -N]$. Here, the only featural contrast in the system may be between verbs and nouns ($\pm V$ or $\pm N$). At a second granularity level, we encounter Phasal distinctions (C, v; Chomsky, 2008). These are located at a coarser-grained level than the three Core Functional Categories C, T and v (Chomsky, 2000, 2001), given that T is not inherently a phase. According to Biberauer and Roberts (2015), the first three layers (EP, Phase and CFC) are characterised by *formal* features, like the clause-typing $[C]$ for the Phase and CFC C-domain or [TENSE] for T (in a language that grammaticalises tense, that is; cf. Biberauer and Roberts, 2015, p. 13 and Ritter and Wiltschko, 2014). Grammaticalised *semantic* features, on the other hand, typify the finer-grained cartographic fields (e.g., Force–Topic–Focus–Fin). These represent iterations of C-positions (i.e., sequences of heads with $[C]$), each associated with a grammaticalised semantic feature (e.g., Topic with [TOPIC]). Semantic features also characterise the heads in works such as Cinque (1999), which make even subtler distinctions (e.g., the syntactic and semantic distinction between Mood$_{Evaluative}$ and Mood$_{Irrealis}$), hence the 'semantically distinct' label in (3).

In turn, (3) can be reformulated as an emergent categorial hierarchy, which unifies and identifies overall commonalities between independently proposed hierarchies and categorial sequences in the literature. The hierarchy thus begins at an initially 'undivided' categorial space and gradually increments the level of grain with which the system is analysed (Biberauer and Roberts, 2015, p. 6):

(4)



The key for present purposes is that (4) can be conceptualised as a learning path. The emergent hierarchy is regulated in the contrastivist manner outlined above, with acquirers postulating categorial distinctions once contrastive features are detected in the input. According to Biberauer and Roberts (2015) and subsequent work in Biberauer's (2019) MMM model, the child is expected to first make a basic predicate/argument (or 'archi' N/V) distinction (see Douglas, 2018; Song, 2019; Biberauer, 2019, on 'archi-V' and 'archi-N' features and categories), which then aids them in making a representational division into the verbal and nominal EPs. Afterwards, the learning path makes a coarse-grained subdivision between discourse (CP-internal) material (containing, e.g., early wh-elements) and thematic (*v*P-internal) material (the Phasal stage in 4). At this point, (at least some) CP-structures are expected to emerge. This can also include some early interactional language (of the sort in, i.a., Speas and Tenny, 2003; Haegeman and Hill, 2013; Wiltschko, 2021; Heim and Wiltschko, 2021), encoded at this stage as part of this coarse-grained CP and not distinguished as a separate Speech-Act domain until a later developmental stage (see 3). At a later stage, this representational 'scaffolding' is refined to incorporate the addition of the 'anchoring' or TP-domain and thus the formation of Core Functional Categories layer (Chomsky, 2000, 2001). Crucially, however, and although not represented in (4) for simplicity, it is not until a later stage, after the emergence of the CFCs, that the approach expects cartographic heads and semantically distinct heads (as in Rizzi, 1997; Cinque, 1999, respectively) to emerge and further subdivide each of the CFCs.

Biberauer and Roberts (2015) converge, then, with the already-discussed inward-growing approaches (§2.2.2) and other approaches which have also, in various forms, argued for early emergence of the CP-domain and peripheral positions, and against a strictly bottom-up approach to development. For instance, Tsimpli (2005) provided data from child Greek arguing in favour of early emergence of peripheral (CP) positions and van Kampen (2010) has similarly supported early emergence of V2 in Dutch, before inflectional and TP-based knowledge. Biberauer (2018) also clearly highlighted the important learnability role of structural edges and speaker-hearer-oriented material in her Peripheral Speaker-Hearer Hypothesis. The approach argues that *phasal* edges constitute the locus of here-and-now and speech-act-oriented items and predicts interactional and speaker-hearer-related elements to be acquisitionally advantaged as a result of these edges' key domain-signalling function (see Bosch, submitted, for a review of approaches arguing for early emergence of a CP and/or Speech-Act-related domain, and also §3.3.2).

Altogether, the above clarifies the particular sense in which neo-emergentist approaches actively disprefer maturation. They allow for greater ontological and categorial malleability throughout development, facilitated by the adoption of emergent categories, refined in the 'increasing granularity' manner introduced, and constrained by the interaction between the three factors (on the ways in which this flexibility lies within principled bounds, see Ramchand and Svenonius, 2014; Biberauer and Roberts, 2015; Biberauer, 2019; Song, 2019). Neo-emergentism thus adumbrates the possibility that development could be modelled with the more parsimonious emergent categories, a line of generative theorising arguably still underinvestigated as far as language development is concerned.

In line with the arguments raised in §2.3, I follow this line of thinking and I proceed in the next section to detail one proposal that probes Biberauer and Roberts's (2015) emergent categorial hierarchy, summarising its key notions and its proposed scientific potential.

2.5 A neo-emergentist and systems-theoretic perspective

Henceforth, I propose to adopt Biberauer and Roberts's (2015) perspective on emergent categories as a testable hypothesis for the acquisition of functional categories, incorporating as well some of the conceptual results from Dynamical Systems Theory (DST) in Bosch (2022).

The predictions for clausal development in this neo-emergentist and DST-driven approach, then, will naturally remain identical to those in Biberauer and Roberts (2015). I will also assume, with them, that clauses can be crosslinguistically analysed making use of the granularity levels in (3). The predictions will be summarised in §2.6 and will be compared with the maturational proposals under consideration; as will be shown, these diverge in important ways. As a result of following Biberauer and Roberts (2015), the approach places emphasis on various interrelated processes in acquisition that are important from a neo-emergentist perspective, one being the already-mentioned *increasing mental granularity*, crucial in Biberauer and Roberts (2015). However, I introduce two further notions and processes — *structural homology* and *softly-assembled development* —, which highlight key aspects of neo-emergentist systemic development.

These latter two terms are co-opted from Dynamical Systems Theory, a mathematical framework of analysis of the development of complex systems, which is interlinked with acquisition concerns in Bosch (2022). There, I suggest that the properties of so-called *Complex Adaptive Systems* (CASs) converge with some of the developmental expectations by MMM. Briefly, CASs represent mathematical models of developing systems, in particular of emergent systems with multiple components that are in constant flux as a result of environmental interactions. Biological systems and units like cells, the brain and cognitive systems are examples of CASs. Ultimately, the central proposal is that incorporation of DST-derived notions in an MMM-based model may have fruitful interdisciplinary potential.

Going back to our starting point, structural homology refers to the property exhibited by CASs, whereby systemic reorganisations are always composed of structures present at earlier stages. Later emergent forms are thus built up based on earlier ones (Tucker and Hirsh-Pasek, 1993). In the current context, these emergent properties entail child grammars do not simply 'recapitulate' the layers in an *adult-like* spine (which are, by hypothesis, not UG-given and thus unavailable to the child). Innate adult-like categories are not drivers of learning, nor is development pre-wired. Instead, grammatical growth is emergent or *softly-assembled.* As subtler contrasts are detected, categories increase in granularity, going from maximally general or coarse-grained stages to later stages which progressively approximate the representational sophistication of adult grammars by refining developmentally-prior items. Systems thus proceed from a state of relative globality or generality towards specificity and differentiation, as in dynamical systems more generally (see also Rice and Avery, 1995; Harley and Ritter, 2002; and Biberauer and Roberts, 2015, on the NO>ALL>SOME learning path).

The rationale underpinning this neo-emergentist and systems-theoretic perspective is one of 'interdisciplinary crossfertilisation'. It aims to endorse neo-emergentist convictions from extralinguistic vantage points. Indeed, the developmental predictions by DST and neo-emergentism converge: structural homology coincides with the 'category refining' dynamics in Biberauer and Roberts's (2015) emergent categorial hierarchy, with later stages refining developmentally-earlier categories. CASs also proceed from states of generality to specificity, again exhibiting synergies with the coarse-to-fine-grained trajectory in Biberauer and Roberts (2015). Perhaps most critically for present purposes, DST, as a theory of *emergent* systems, rejects hard-wired development, favouring instead softly-assembled development. This makes DST inherently incompatible with maturation (hard-wired development *par* *excellence*), like Biberauer and Roberts (2015). DST-based neo-emergentism then keeps the analytical and predictive toolkit in Biberauer and Roberts (2015), but augments it with compatible conceptual notions and arguments from Dynamical Systems Theory (e.g., structural homology, softly-assembled development), ultimately aiming to strengthen the points in the former. Hereafter, I will view neo-emergentist proposals, such as Biberauer and Roberts (2015) and Biberauer (2019), from this DST-inspired perspective.

This approach will be developed further with category-theoretic and systems-theoretic formalisation in §4, which will complement neo-emergentist thinking. I leave this discussion to §4. For now, since category-theoretic and DST-related notation will not become relevant until §4, this theoretical introduction suffices.

2.6 Architectural differences and predictions of each approach

Summarising the preceding discussion, the architectural differences of the approaches introduced can be succintly broken down in terms of three parameters codifying their assumptions:

(i) Biological locus of syntactic categories: innate *vs* emergent.

(ii) Postulation of maturation.

(iii) The directionality of this maturation: bottom-up or inwardly (if applicable).

These are compared in **Table 2.1**, which incorporates the approaches introduced thus far.

Table 2.1 Architectural differences between approaches to syntactic development

Approach	Locus	Maturation?	Directionality
Continuity	Innate	✗	N/A
Growing Trees	Innate	✓	Bottom-up
Inward Growing Spine	Innate	✓	Inwardly
Neo-emergentism	Emergent	✗	N/A

Finally, the predictions with regard to the development of functional structure based on contemporary theoretical accounts are described below. As mentioned earlier, a more general version of an inward-growing approach is adopted here in line with the assumptions made regarding clausal structure (see 3) and given that the focus is not on interactional language, as in Heim and Wiltschko (2021). Therefore, I test an inward-growing approach assuming an innate cartographic hierarchy (of sort in **Figure 2.2**), according to which the spine matures inwardly, starting at both edges. Although (ii) has not been proposed in the literature (and due to its toy nature, remains generic and underdeveloped), it is still important to consider at least some version of an inward-growing account, as a competing maturational hypothesis to Biberauer and Roberts (2015) (which also assumes early development of CP) and as an alternative to a bottom-up approach (in the event that a maturational approach were most satisfactory but the empirical data favoured an inward-growing framework).

(i) **Bottom-up maturation** (as in Friedmann et al., 2021) predicts knowledge of functional structure to be acquired bottom-up, such that the emergence of *v*P-structures precedes the emergence of TP-structures, which themselves are developmentally-prior to CP-structures. It does not predict subprojections within a single domain (e.g., cartographic heads) to emerge later across-the-board: as these are provided by UG, they should become available as the spine matures bottom-up, with cartographic heads in the lower left periphery becoming available before heads in the upper left periphery (in two stages, Fin-to-QP and Top-to-Force, see **Figure 2.2**). All other things equal, co-occurrence of multiple CP-structures (e.g., topics preceding wh-elements) is thus not expected to be delayed relative to single CP-structures once the left-periphery has developed in full.

(ii) **Inward-growing maturation** (generalised) predicts an initial stage where only *v*P- and CP-structures are available, which precedes the emergence of TP-structures. It does not predict subprojections within a single a domain (e.g., cartographic heads) to emerge later across-the-board: as these are provided by UG, they should become available as the spine matures inwardly. Broadly, this approach would thus predict *v*P-structures and most CP-structures (e.g., wh-elements, topicalisations, V2) to emerge before TP-elements (e.g., auxiliaries, verb-raising, subject-raising, productive tense-agreement inflection). It would additionally anticipate evidence of multiple left-peripheral heads (e.g., co-occurrence of wh-elements with topics) to surface with the emergence of CP-structures.

(iii) **Neo-emergentism** (following Biberauer and Roberts, 2015) predicts an initial stage where only *v*P- and (at least some) CP-structures are available, which precedes the emergence of TP-structures. In contrast to (i) and (ii), evidence for multiple projections within each of these tripartite domains (thematic, tense/inflection and discourse) *is* predicted to develop later as part of the emergence of increasingly finer-grained syntactic categories. Co-occurrence of multiple CP-structures should emerge significantly later than individual CP-structures and after TP-structures have emerged.

I put these predictions to the test in §3, where I present a novel corpus study that examines their predictive differences in the domain of developmental directionality and regarding changes in children's representational granularity. I show that, despite its exploratory nature, the data may provide significant insights into the feasibility of syntactic maturation, the ontological and biological foundations of syntactic categories and the nature of their development in acquisition more generally.

3 | *The emergence of clausal spines: a corpus study*

This chapter describes the corpus study. **Section 3.1** introduces its methodology, outlining the corpora and structures analysed. **Section 3.2** presents the results of the study and discusses each developmental stage detected. Finally, a commentary on the theoretical implications of the data and crosslinguistic comparison to other acquisition literature is offered in **Section 3.3**.

3.1 Methodology

3.1.1 Children studied

We performed a longitudinal analys is of the spontaneous productions of ten typically-developing children across five languages: Catalan, Italian, Spanish, German and Dutch. Two children per language were selected from the CHILDES database (MacWhinney, 2000). The period in which children's production was analysed varies from corpus to corpus. Overall, it ranged from 0;11 to 5;02. We calculated the word-based MLU for each file. 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 the children studied, including their MLU and age range and the number of files analysed in their corpora, are reported in **Table 3.1**. Henceforth, all examples and tables will be presented in the order of languages and children given below.

Table 3.1 Children studied in the CHILDES database and summary information (Serra and Solé, 1989; Cipriani et al., 1989; Llinàs-Grau and Ojea, 2000; Montes, 1987; Miller, 1979; Bol, 1995; van Kampen, 2009)

Language	Corpus	Child	Files analysed	Age range	MLU range	Total words	Total utterances
Catalan	Serra-Solé	Laura	19	1;07-4;00	1.03-3.47	16651	7576
		Gisela	21	1;07-4;02	1.02-3.51	11132	4753
Italian	Calambrone	Martina	13	1;07-2;07	1.26-2.69	7569	3586
		Rosa	21	1;07-3;03	1.27-3.24	14659	6614
Spanish	Llinàs-Ojea	Irene	59	0;11-3;02	1.0-5.13	38234	11908
	Montes	Koki	13	1;07-2;11	1.96-3.61	11773	4276
German	Miller	Kerstin	37	1;03-3;04	1.09-2.89	28609	16686
		Simone	50[1]	1;09-2;09	1.52-4.89	45407	21371
Dutch	Groningen	Josse	28	2;0-3;04	1.2-4.01	30066	11120
	van Kampen	Sarah	50	1;06-5;02	1.07-6.07	49465	17408

3.1.2 ‘Stages, not ages’: the role of MLU

As Paradis and Genesee (1997), Caprin and Guasti (2009), Friedmann et al. (2021) and others have noted, age is to be avoided in both intralinguistic and crosslinguistic comparisons, to evade the high variability in linguistic development that is observed among children of the same age range. MLU will be adopted instead as the guiding metric for syntactic development to facilitate crosslinguistic and inter-speaker comparisons, as well as the replicability of the present findings.

Specifically, I will adopt word-based MLUs throughout. The choice of word-based MLUs over morpheme-based metrics follows from various considerations: firstly, not all corpora in the CHILDES database are morphosyntactically tagged, meaning morpheme-based MLUs are not available in practice. Secondly, the calculation of morpheme-based MLUs is highly prone to both annotation errors and subjectivity: they require the researcher or annotator to establish which morphemes the child uses productively, a process that may lead to arbitrary and error-prone decisions (Ezeizabarrena and Garcia Fernandez, 2017). We thus take word-based MLUs as a better overall developmental measure, balancing effectiveness and ease of application for early child data. Ages will often be reported nonetheless, but they will not be used to make any analytical decisions and generalisations.

[1]For Simone, the first 50 files in the corpus were analysed, out of the 73 recordings in total. For the rest of the children, the entire corpora were studied.

3.1.3 Primary corpus study: structures analysed

Three types of structures were searched for: (i) CP-related elements, (ii) TP-related elements and (iii) elements that suggest *further internal organisation* within CP, e.g., a 'Split CP', which allows at least two different projections. Diagnostics probing specifically for cartographic-type structures have rarely been adopted in previous studies (some exceptions being Westergaard, 2009; Moscati and Rizzi, 2021). As will be argued, however, they will play a key role in analysing the feasibility of proposed theoretical models.

This study does not test for the thematic *v*P/VP-domain. I take this domain, which, among others, hosts argument-structural relations, to be available very early on (see Lidz, 2022, for a recent review of extremely early development of argument-structural relations)[2].

3.1.3.1 CP-structures

Evidence for knowledge of the CP-domain will be gathered from six main sources, namely wh-questions, yes/no questions (for Germanic languages only), topicalisations/focalisations, finite embedding markers, illocutionary complementisers (for Romance) and knowledge of V2 (for Germanic).

Evidence for V-to-C movement (or a distributional distinction between finite/non-finite verbs in V2/V-final position) in Dutch and German children will be taken to indicate an emerging CP-domain. More precisely, the study investigates evidence in the same recording of both non-finite verbs in a non-raised (or base) position with OV-order as well as of finite verbs which can be shown to have raised outside of the verbal domain. If the initial constituent is the subject or some non-object topic, these display surface VO-order when the verb is finite. This is illustrated below, where (5a, 5c, 5d) exemplify verb-raising of *haben/lust/eet*, shown by the absence of surface OV-order, and (5b) displays a combination of a finite modal in its V2-position (*willen*) and a non-finite verb *trinken* following its object *Kola*[3]:

(5) a. Sie haben Glück (German)
they have.3PL luck
'They are lucky.'

b. Ich will Kola trinken
I want.1SG Cola drink.INF
'I want to drink coke.'

[2] Therefore, I will also remain agnostic as to whether there is a 'small clause' stage in acquisition (i.e., an exclusively VP and lexical stage; Radford, 1988, 1990) and, indeed, whether it is detectable exclusively on the basis of production data.

[3] Unless otherwise noted, the examples provided are mine. For Italian, German and Dutch, they were also checked with native speakers.

c. Ik lust geen kaas (Dutch)
I desire.1SG no cheese
'I don't like cheese.'

d. Nu eet ik een boterham
now eat.1SG I a sandwich
'Now I eat a sandwich.'

(5a-5c) instantiate cases of subject-initial structures with finite verbs. These structures have been analysed as involving raising of the subject and verb to the T-domain only (Zwart, 1993, 1997, *et seq.*). I will, nonetheless, include subject-initial declaratives as evidence of knowledge of V2 (following, i.a., den Besten, 1983; Schwartz and Vikner, 1989, 1996; van Craenenbroeck and Haegeman, 2007), on the grounds that exclusion of subject-initial sentences would effectively reduce the V2-diagnostic to the (non-default) topicalisation diagnostic. This is undesirable, as it would not control for a scenario where knowledge of the V2-system emerges *before* the first productions of topicalised non-subject elements. This point also becomes important in light of literature noting an initial stage in the acquisition of V-to-C in Germanic where children omit the first-position element, yielding surface V1 (Evers and van Kampen, 2001; van Kampen, 2010). Note, though, that this decision does not itself rule out the analytical possibility that the earliest subject-initial declaratives are encoded by the child as involving raising to the TP-domain or a projection distinct from CP. I will return to this point in the discussion of the empirical data in §3.2.3.1 and §3.2.3.2, where I will reject this possibility on various grounds.

Some examples of wh-questions, where the wh-elements are raised to SpecCP, alongside subject-auxiliary inversion, are given in (6).

(6) a. **Quan** comença el partit? (Catalan)
when begin.3SG the match
'When does the match start?'

b. **Perché** l'hai cambiato? (Italian)
why CL.DO=AUX.HAVE.2SG change.PTCP
'Why did you change it?'

c. **Como** cantó María? (Spanish)
how sing.PRFV.3SG Mary
'How did Mary sing?'

d. **Wo** hast du gegessen? (German)
where AUX.HAVE.2SG you eat.PTCP
'Where have you eaten?'

e. **Wat** doen we vanavond? (Dutch)
what do.1PL we tonight
'What are we doing this evening?'

Although analysed and counted as CP-diagnostics, wh-less questions of the sort in (7), attested in Germanic languages, are not counted as sufficient evidence for emergence of the CP-domain. That is, we required other CP-diagnostics to have *also* emerged for us to conclude that the CP-domain might be in place (this decision only became relevant in one child, Sarah)[4]. Questions with a wh-word but no verb were excluded.

(7) Macht das Pferd? (German)
do.3SG the horse
'What is the horse doing?' (Schmerse et al., 2013, p. 658)

In Germanic languages, yes/no questions were included, given these involve verb-raising to the C-domain. In contrast, polar interrogatives in Romance are typically marked via intonation only, with inversion having been lost for most modern Romance varieties (Giurgea and Remberger, 2016). Note that VS-order as a marker of polar interrogatives has been retained in Spanish, coexisting with intonation-only marking (e.g., *¿Está María en casa? vs ¿María está en casa?* 'Is Mary at home?'; Giurgea and Remberger, 2016, p. 865). However, this order is syntactically ambiguous between a structure with raising to C (inversion properly understood) and one where the verb does not move from T and the subject remains in a lower position. For this reason, too, I exclude all polar interrogatives in Romance.

Topicalisations and focalisations are exemplified in (8-9). In Romance, these also include Clitic Left-Dislocation (CLLD), where the topic is accompanied by a clitic (8a-8b), and other cases of contrastive or informational foci (Laka, 1990; Rizzi, 1997), where a clitic is not required (8c).

(8) a. **Questa** la compro io (Italian, CLLD)
this CL.DO= buy.1SG I
'This one I'm buying.'

b. I **a mi** em donaràs un regal? (Catalan, CLLD)
and to me CL.IO= give.FUT.2SG a present
'And will you give ME a present?'

c. **Arroz** he comido (no pasta) (Spanish, Contrastive focus)
rice AUX.HAVE.1SG eat.PTCP not pasta
'I have eaten RICE (not pasta).'

(9) a. **Das** glaube ich nicht (German, Topicalisation)
that believe.1SG I not
'That, I don't believe.'

[4]This requirement may be unduly conservative. Santelmann (2003) argues for wh-less questions in early Swedish that they display evidence of CP-level structure (see also Evers and van Kampen, 2001; van Kampen, 2009, on Dutch).

b. **Nu** heb ik mijn bal vergeten (Dutch, Topicalisation)
now AUX.HAVE.1SG I my ball forget.PTCP
'Now I have forgotten my ball.'

Knowledge of embedding was diagnosed via finite embedding structures containing overt subordination markers. These involve complement and relative clauses headed by a subordinating marker (Cat./Sp. *que*, It. *che*, Ger. *dass* and Dutch *dat*, 'that')[5], relative clauses introduced by relative pronouns (i.a., German *der* or *die*, 'who.NOM.MASC.SG' and 'who.NOM.FEM.SG', respectively), conditional clauses with a conditional marker (e.g., Cat/Sp. *si*, It. *se*, Ger. *wenn*, Dutch *als*, 'if') and any other clauses headed by a subordinating marker (e.g., temporal, causal or purpose clauses). Non-finite embedded clauses were not counted as evidence for CP. Embedding markers are abbreviated in the tables in §3.2.1 and the Appendices as 'Embed'. Examples from Catalan and German are given in (10):

(10) a. Hi havia una caputxeta vermella, **que** anava (Catalan)
CL.LOC= be.IMPF.3SG a little.riding.hood red that go.IMPF.3SG
a casa d'un llop
to house of-a wolf
'There was a little red riding hood, who was going to a wolf's house.'

b. **Wenn** du diese Taste drückst, singt das Spielzeug (German)
if you this button press.2SG sing.3SG the toy
'If you press this button, the toy sings.'

Finally, a further type of main-clause complementisers tracked in the Romance languages are so-called *illocutionary* complementisers, present primarily in Ibero-Romance (see Corr, 2016). Illocutionary complementisers involve instances of main-clause complementisers (such as Cat./Sp. *que*) that do not function as subordinators heading a complement clause, and instead surface in main-clause contexts with a range of illocutionary functions (Corr, 2016, 2022). I will focus on Spanish and Catalan, given the child data to be considered. Adopting Corr's (2016) terminology, the item *que* in (11a) generates an exclamation (EXCLAMATIVE *que*); a quotation in (11b) (QUOTATIVE que); and in (11c), it contextualises the preceding utterance information, resulting in a conversational move (CONJUNCTIVE *que*).

(11) a. ¡**Que** casi sube y rompe la estantería! (Spanish)
that.EXCL nearly climb.3SG and break.3SG the shelves
'He/she nearly climbed and broke the shelves!'

[5] German, however, allows asyndetic (complementiser-less) embedded clauses, yielding V2-order, as a more restricted alternative to canonical V-final complement clauses introduced by *dass* (e.g., Schwartz and Vikner, 1996). No clear-cut instances of embedded V2 before the emergence of *dass* were found in the German data, so I will not comment on these further.

b. **Que** no lo pienso hacer
that.QUOT not CL.DO= think.1SG do.INF
'(I've said) I'm not going to do it.'

c. Assenta't **que** et pentinaré (Catalan)
sit.down.IMP.2SG=CL.REFL that.CONJ CL.IO= comb.FUT.1SG
'Sit down, I'll brush/comb you.'

Collectively, these are referred to as ILLOCUTIONARY *que*, to distinguish these uses from the well-established role of *que* as a subordinating complementiser. To these, we also add the possibility in Standard Eastern Ibero-Romance of introducing matrix polar interrogatives with *que*, as in (12) (INTERROGATIVE *que*). This means interrogative *que* is sanctioned in Catalan, but generally disallowed in Spanish.

(12) **Que** em pots ajudar? (Catalan)
that.INT CL.IO= can.2SG help.INF
'Can you help me?'

Although most robustly attested in Ibero-Romance, some main-clause complementisers are also attested in Italian. Many Italo-Romance varieties, particularly in central and northern Italy, generally present a more restricted system of illocutionary complementisers compared to Ibero-Romance. Broadly, they only allow conjunctive uses of *che*, *che* in jussives/hortatives with subjunctive mood and also interrogative uses of *che* (the latter depending on the variety; see Cruschina, 2012; Giurgea and Remberger, 2016). Conjunctive *che* and *che* in jussives/hortatives are exemplified in (13a-13b). Italo-Romance dialects also present constructions where the complementiser is preceded by an adjective/adverb (13c), a pattern attested across Romance (Cruschina and Remberger, 2018).

(13) a. Dammene un altro **che** lo mettiamo qui (Italian)
give.IMP.2SG=CL.IO=CL.PART a other that.CONJ CL.DO= put.1PL here
'Give me another one, we'll put it here.'

b. **Che** ciascuno scelga una carta!
that each choose.SBJV.3SG a card
'Everyone choose a card!' (Giurgea and Remberger, 2016, p. 874)

c. Certo **che** ha capito!
certain that AUX.HAVE.3SG understand.PTCP
'Of course he/she understood!' (Cruschina and Remberger, 2018, p. 1)

Therefore, although less frequent than in Ibero-Romance, we will also search for illocutionary complementisers in the Italian data[6]. These are abbreviated in the tables as 'Illoc'.

3.1.3.2 TP-structures

TP-level knowledge was tested for by probing for the presence of the following structures: Subject-Negation-Verb orders, Subject-Adverb-Verb orders, Subject-Clitic-Verb orders and auxiliaries. The majority of the diagnostic criteria for TP tracked word-order patterns to identify movement of a subject to a domain above *v*P and below CP.

This included instances where an adverb, negation or, in Romance, a clitic was placed between the subject and the verb. Some examples from Catalan, Italian and Spanish are given in (14):

(14) a. **La mama el** va comprar per a l'aniversari (Catalan, S-Cl-V)
the mum CL.DO= AUX.GO.3SG buy.INF for to the-birthday
'Mum bought it for the birthday.'

b. **Giulia non** vuole giocare con me (Italian, S-Neg-V)
Giulia not want.3SG play.INF with me
'Giulia doesn't want to play with me.'

c. **Jaime ya** volvió de la escuela (Spanish, S-Adv-V)
James already come.back.PRFV.3SG from the school
'James already came back from school.'

Subj-Neg-V, Subj-Adv-V and Subj-Cl-V configurations are taken to indicate that the subject has raised out of the thematic domain, to a higher clausal position, namely TP. This is under the assumption that the negation marker is placed at the edge of *v*P in Germanic (Zeijlstra, 2004) and within TP in Romance (Pollock, 1989), and that we can also distinguish between VP- or TP-adjoined adverbs. We also assume that clitics left-adjoin to T (following Kayne, 1991, *et seq.*; see also Pollock, 1989; cf., i.a., Uriagereka, 1995; Raposo and Uriagereka, 2008; Manzini, 2023, for other analyses). Importantly, Subj-Neg-(V) and Subj-Adv-(V)[7] diagnostics were only used for Germanic when the subject can be shown *not* to be in the CP-domain, such as in the sentences below, where yes/no questions and topicalisation signal that the subject is not located in SpecCP:

[6]Note that several Southern Italian Dialects (SIDs), like Calabrian, display a more expansive system of illocutionary complementisers (see Giurgea and Remberger, 2016). The two Italian children studied grew up in Tuscany, so they were not exposed to varieties with robust illocutionary complementisers.

[7]With V in parentheses in Germanic as it will often have raised to the C-domain, thus outscoping both Subj and Adv/Neg.

(15) a. Bist **du heute** laufen gegangen? (German, S-Neg-V)
AUX.BE.2SG you today run.INF go.PTCP
'Have you gone running today?'

b. Dit boek heb **ik niet** gelezen (Dutch, S-Adv-V)
this book AUX.HAVE.1SG I not read.PTCP
'This book, I haven't read'.

Non-topical subjects are excluded from these two diagnostics as well, on the grounds that only topical subjects have been argued to raise outside the *v*P in Germanic, with non-topical subjects remaining low (Diesing, 1992; Fernald, 2000; Mohr, 2005; Kratzer and Selkirk, 2007).

These diagnostics crucially assume that overt subjects are placed in the TP-domain. For Romance, this is not uncontroversial, particularly in light of approaches that analyse overt subjects in Romance null-subject systems as being topical and thus located within CP (e.g., Barbosa, 1995; Ordóñez, 1997; Alexiadou and Anagnostopoulou, 1998; Raposo and Uriagereka, 2008). I pursue a conservative route in this thesis and treat overt subjects as being in the SpecTP position, in line with previous acquisition literature (e.g., Guasti, 1993; Bel, 2003; Villa-García, 2014), but the consequences of this assumption are worth exploring in future research (building, for instance, on Grinstead, 2004; Villa-García, 2014). For Germanic, this assumption also raises questions about the status of the Extended Projection Principle (EPP), which is, again, amply debated (see, e.g., Biberauer, 2010, on its (non)obligatoriness). I assume, nonetheless, that, whenever the subject has raised above negation or adverbs and below the CP-domain, these subjects are located in the 'anchoring' or inflectional domain. The diagnostics adopted here for Germanic crucially *can* remain agnostic as to the status of the EPP in Germanic and allow for a scenario where the EPP is simply optional.

Lastly, we also tracked instances of auxiliaries (including past-tense, aspectual, future auxiliaries, as well as progressive[8] and passive auxiliaries). This was complemented with a quantitative analysis of tense-agreement morphology across all ten children. Full details can be found in **Appendix C** and it will be summarised in §3.1.5. The implications will be foregrounded in §3.2.

From a theoretical perspective, it is also worth remarking that there is *a priori* little reason to assume that Germanic auxiliaries, and any systems that distinguish HAVE/BE auxiliaries, such as Italian, should emerge at the same time as T-elements proper. Systems with split auxiliary selection are commonly analysed in such a way that the allomorphs of HAVE and BE both realise a node in the aspectual/*v*P domain (see, i.a., Adger, 2003;

[8]The Dutch progressive construction *zijn* 'to be' + *aan het* 'on the' + V_{Inf} is excluded as a TP-diagnostic, as it is assumed to be hosted in the *v*P (see Bogaards et al., 2022).

D'Alessandro and Roberts, 2010; Bjorkman, 2011). Therefore, it is worth bearing in mind that the auxiliary diagnostic for German, Dutch and Italian may, at least partially, be targeting non-TP-material, i.e., material located within the aspectual domain, e.g., Asp/Perf. This caveat is important, particularly for future work that wishes to tease apart the emergence of TP vis-à-vis AspP. Since this is not the present aim, I will set it aside. §3.2 will show, too, that this potential syntactic difference does not seem to impinge on the results, as all auxilaries emerge at comparable times crosslinguistically[9].

Finally, a clear gap is the absence of V-to-T diagnostics. Their omission is due to various factors. Firstly, a by-now frequent analysis of Germanic V2 assumes that verbs do not transit through T, instead moving from V to C directly (see Holmberg and Platzack, 1995; Vikner, 1995, 2001, 2005). Under this analysis, V-to-T diagnostics for Germanic become meaningless. For Romance, V-to-T raising could be diagnosed via surface Verb-Adverb orders (following Pollock, 1989). However, these kinds of constructions with adverbs are known to be very rare in early child speech (e.g., Platzack, 1992; Armon-Lotem, 1997; Schaeffer, 2000). Since the robustness and frequency of V-to-T diagnostics appear weak and they are furthermore unbalanced across language groups (being likely futile in Germanic), no V-to-T diagnostics were adopted.

3.1.3.3 Split CP-structures

As Moscati and Rizzi (2021) note, the most direct evidence for knowledge of a fine-grained CP-domain would involve cases of co-occurrence of left-peripheral elements. This is the only diagnostic adopted here for these structures and we take them to provide evidence for a CP-system that is additionally populated by a set of separate projections.

In Romance, these will most frequently involve instances of topics preceding wh-phrases (16) or cases of 'recursive' topics (17). A possible structural representation of (16a) is given in (16b).

(16) a. **La Júlia**, **on** ha anat? (Catalan)
the Júlia where AUX.HAVE.3SG go.PTCP
'Júlia, where has she gone?'

b. $[_{TopP}$ La Júlia $[_{Q/FocP}$ on ha $[_{TP}$...$[_{VP}$ anat]]]]

(17) a. **Questo**, **a te**, ti spaventa (Italian)
this to you CL.IO= scare.3SG
'This, it scares YOU.'

[9]On Biberauer and Roberts's (2015) hierarchy, at least, this is not unexpected: T and Asp are parallel in their system.

b. **Aquí**, **esto** no lo pones (Spanish)
here this not CL.DO= put.2SG
'Here, this, you're not putting it.'

Other instances of cartographic-type structures will include cases of complementisers or subordination markers being followed by a topic or a wh-element, such as those in (18a), and, in Ibero-Romance, instances of quotative complementisers preceding wh-elements (18b). In Catalan, furthermore, interrogative complementisers can be *preceded* by topics (18c):

(18) a. Me dijo **que el puzle** no lo quería comprar (Spanish)
CL.IO= tell.PRFV.3SG that the puzzle not CL.DO= want.IMPF.3SG buy.INF
'He/she told me that he/she didn't want to buy the puzzle.'

b. ¿**Que cuánto** te han costado estas bambas?
that.QUOT how.much CL.IO= AUX.HAVE.3PL cost.PTCP these trainers
'How much have you said these trainers have cost you!?'

c. I **en Joan**, **que** no vindrà a sopar? (Catalan)
and the John that.INT not come.FUT.3SG to dine.INF
'Is John not coming to dinner?'

In Catalan and Spanish, *sí que* 'yes that' and *que sí que* 'that yes that' constructions, the former shared with Italian *sì che*, are also analysed (see 19). The examples below are accompanied by the (adult-like) cartographic structures assumed in Villa-García and González Rodríguez (2020) (the label of XP is left open to debate, see p. 26). Although I remain agnostic as to exact implementation of the cartographic analysis of these structures, the key assumption that I adopt from this work and others below is that these constructions require more elaborate structure than a basic CP (in the CFC sense)[10].

(19) a. Pablo **sí** $\boxed{\text{que}}$ vino (Spanish)
Paul yes that come.PRFV.3SG
'Paul did come.'

b. Creo $\boxed{\text{que}}$ **sí** $\boxed{\text{que}}$ ganaron la carrera
think.1SG that yes that win.PRFV.3PL the race
'I think that they did win the race'.

[10]These constructions resemble recomplementation structures in Ibero-Romance, where a topic is 'sandwiched' between two complementisers. Recomplementation has often received a similar cartographic analysis (Villa-García, 2012, *et seq.*; though cf. Villa-García and Ott, 2022 for a paratactic account). Although recomplementation structures were initially investigated, only 1 non-clear-cut example was found in the Ibero-Romance children (Irene). Therefore, I will set recomplementation aside in what follows.

c. [$_{ForceP}$ (que) [$_{TopicP}$ [$_{XP}$ sí [$_{X}$ que [$_{TopicP}$ [$_{FocusP}$ [$_{\Sigma P}$ [$_{TP}$ ...]]]]]]]]

As far as Germanic languages are concerned, however, the V2-system makes a Split CP-domain substantially harder to detect and, indeed, not all approaches agree that the left periphery of Germanic V2-languages is 'cartographic' in the traditional sense (cf. Giorgi and Pianesi, 1997; Biberauer and Roberts, 2015; Cormany, 2015; Walkden, 2017; Hsu, 2017, 2021, for various views that assume a 'weakened' left periphery in V2-languages, e.g., via 'conflation' or 'bundled' functional heads). However, limited instances of V3-orders are present in Germanic languages (these are more widespread in all colloquial and some vernacular varieties of German, although less so in Dutch; see Freywald et al., 2015, though cf. Meelen et al., 2020, on Dutch urban youth varieties). These surface V3-orders, apparent violations of the V2-rule, are often modelled via an additional projection above the landing site for V2-subjects (e.g., Haegeman and Greco's, 2016, 2018, FrameP). This additional projection accommodates these restricted V3-patterns and helps capture why a restricted set of elements would seemingly be invisible to the general Germanic V2-rule.

Structures generating surface V3 include frame-setters, Hanging Topic Left-Dislocation (HTLD) and Contrastive Left-Dislocation (CLD). Frame-setters are information-structural categories that provide specifications that restrict the domain in which the proposition is valid (such as temporal or local adverbs and phrases). HTLD permits DPs in first position to correspond with a resumptive personal pronoun in the V2-clause, often (though not invariably) in the middle field. CLD instead involves an utterance-initial XP with contrastive topicalisation or focus alongside a resumptive pronoun (typically a *d*-pronoun), which is preferably located in the left periphery (again, not obligatorily; see Frey, 2004, for a review of the syntactic differences between HTLD and CLD). Some examples of these are offered in (20).

(20) a. **In alle geval**, ik had het niet verwacht (Dutch, Frame-setter)
in any case I AUX.HAVE.PST.1SG it not expect.PTCP
'Anyway, I had not expected it.' (Haegeman and Greco, 2020, p. 65)

b. **Dem Alex**, ich habe ihm geholfen (German, HTLD)
the.DAT Alex I AUX.HAVE.1SG him.DAT help.PTCP
'Alex, I helped him.'

c. **Diesen Kuchen hier**, den möchte ich probieren (German, CLD)
the.ACC cake.ACC here PRON.ACC want.1SG I try.INF
'This cake here, I want to try.'

Lastly, another kind of V3-orders in Germanic include conditional and temporal clauses followed by a resumptive adverb (Dutch *als… dan* and German *wenn… dann*, 'if/when…then'). These are illustrated in (21) and will also be analysed (following Haegeman and Greco, 2020 and Meklenborg, 2020, who also analyse them cartographically; see also De Clercq et al., 2023, for recent discussion):

(21) a. **Wenn** du ausrutschst, **dann** kannst du dich verletzen (German)
when/if you slip.2SG then can.2SG you CL.REFL= hurt.INF
'When/if you slip, then you can hurt yourself.'

b. **Als** het niet zo warm is, **dan** ga ik naar buiten (Dutch)
when/if it not so hot be.3SG then go.1SG I to outside
'When/if it isn't so hot, then I'll go out.'

A final note regarding the status of V3-orders is in order. Given that V3-orders in Germanic, and Germanic V2 more generally, can be modelled in a non-cartographic manner (see Walkden, 2017, and also references above), it is thus justified to wonder to what extent these V3-structures can properly be called 'cartographic'. This raises issues about what these kinds of V3-orders are diagnosing in our context and its formal nature. As noted earlier, however, I simply take these 'Split CP'-diagnostics (including V3-orders) to be diagnosing *further internal elaboration* within the CP-domain. The *exact richness* or *granularity* of this structural elaboration (e.g., if cartographic in the sense of Rizzi, 1997; Cinque, 1999; and subsequent work or 'weaker', as in Giorgi and Pianesi, 1997; Cormany, 2015; Walkden, 2017, i.a.) is secondary so long as this granularity is greater than that for C as a Core Functional Category. For this reason, I will continue to name these diagnostics 'Split CP'- or 'cartographic-*type*' diagnostics, *not* 'cartographic' diagnostics.

3.1.4 Summary of diagnostics, criteria for emergence and utterances excluded

Table 3.2 presents a summary of all the structures searched for in each recording sample.

Table 3.2 Summary of structural diagnostics used

Functional domains	**Structures analysed**
TP-domain	• Auxiliaries • Subject-Neg-(Verb) (when $\text{Subj}_{Topical}$ not in C in Germanic) • Subject-Adv-(Verb) (when $\text{Subj}_{Topical}$ not in C in Germanic) • Subject-Clitic-Verb (Romance only)
CP-domain	• V-to-C movement (Germanic only) • Wh-questions • Yes/no questions (Germanic only) • Topicalisations and focalisations • Illocutionary complementisers (Romance only) • Finite embedding markers

Split CP	• Top > Wh (Romance only) • Top > Top/Foc (e.g., recursive topics; Romance only) • Comp > Wh/Top (Romance only) • QUOT *que* 'that' > Wh (Ibero-Romance only) • Top > INT *que* 'that' (Catalan only) • *Sí que/sì che* 'yes that' and *que sí que* 'that yes that' structures (Romance and Ibero-Romance only, respectively) • Frame-setters, HTLD, CLD (Germanic only) • Conditional/temporal clauses with resumptive *dann/dan* 'then' (Germanic only) • Other possible combinations of the above

For each diagnostic in §3.1.3, we tested whether the structure appeared at least once in each file and collected this information in a separate table for every child. A given functional domain is assumed to have emerged after the first use of a structure corresponding to that domain (e.g., CP or TP) and followed soon after (in the next ~ 2 files) by other uses of the same or other structures searched for in that domain. That is, isolated instances of a structure were not deemed sufficient evidence for emergence, and no conclusive statements were made in such cases. Therefore, once a domain fulfilled this criterion, the relevant cells in the tables to be presented were shaded in a given colour, to signify their emergence. If a structure in a cell is attested but not shaded, it represents an isolated case, judged insufficient for emergence. The full tables are compiled in **Appendix A**.

The emergence criterion proposed above is necessarily somewhat arbitrary. There is no justified motivation for why ~ 2 files should be the relevant 'window' of consideration. However, to avoid reaching theoretical conclusions on the basis of isolated examples, it is methodologically important to adopt the more conservative measure of emergence above (as done, for example, in Villa-García, 2014), *modulo* the fact that it is not fully satisfactory. Additionally, depending on the recordings' length for a given child, this criterion had to be weakened. This was only the case for the Spanish child Irene. Her recordings at the earliest stages are extremely brief, often including < 50 words by the child, meaning they contain very few to no verbs and, therefore, often none of the structures analysed. It was also relaxed for the Split CP-diagnostics, which, by their very nature, are less frequent than the CP-structures out of which they are formed.

Identical repetitions of an adult utterance were excluded. We also did not count utterances in which the child was singing or reciting lyrics, as well as unintelligible utterances (either due to pronunciation or transcription). Cases of possible right-dislocations (with potential A-bar movement) were not counted. This is because they can be ambiguous between base-generation structures, genuine right-dislocations or, in some cases, afterthoughts. Crucially, for many children, recordings are unavailable to help distinguish these possibilities, making their disambiguation increasingly arbitrary.

3.1.5 Supplementary corpus studies: Appendices B and C

To complement the results of the primary corpus study just outlined and further probe their empirical robustness, two additional corpus studies were carried out *a posteriori*. Firstly, a quantitative analysis was conducted to count the frequency of occurrences of all CP-structures for every child (**Appendix B**). This latter study also examined the variety in types for each of the structures (e.g., the types of wh-questions used, verbs with which V2-order was found, among others). This was necessary to determine the productivity of the earliest CP-structures produced and to detect and discard possible instances of rote-learned formulae (e.g., for the earliest wh-questions).

Notably, too, the diagnostics described in §3.1.3 are entirely structural/syntactic, and do not incorporate any morphological considerations. To control for a scenario where agreement and tense-marking are acquired before structural TP-diagnostics become apparent, we also quantified children's use of inflectional morphemes (**Appendix C**). This secondary analysis of the development of verbal morphology was also performed to establish to what extent (if any) structural diagnostics converged with morphological acquisition. We identified frequency of (i) person-number combinations, (ii) tense, aspectual and mood morphology and (iii) imperatives produced in each file. As criteria for emergence, a morpheme was considered productive once it has been used with at least three different verbs (excluding copulas and the Catalan/Spanish impersonal verbs *haver-hi/haber* 'there is/are', which only conjugate with 3rd person), partly following the criteria in Caprin and Guasti (2009).

Space constraints preclude a complete description of these complementary studies, but see **Appendices B** and **C** for greater detail. §3.2.3 will predominantly focus on the empirical patterns arising from the primary corpus study. However, the ramifications of both additional analyses for the overall developmental patterns will be discussed where appropriate.

3.2 Results

The analysis of the spontaneous production of the ten children revealed, in most instances, a clear developmental trend regarding the gradual emergence of the various structures analysed. I begin by presenting an outline of the *surface* findings that concern order of emergence of functional structures. Three clear stages will emerge from this production data. Thereafter, I consider the connection between MLU – interpreted as a sign of syntactic development – and the structures produced by the children. A more in-depth commentary and analysis of the structural patterns in the data, both across languages and across children, will follow in §3.2.3. In this section, I will incorporate the results obtained from the supplementary quantitative and morphological analyses in **Appendices B** and **C**. From these, I will conclude that the first developmental stage detected receives weak support; this will help qualify the empirical results and their theoretical implications.

3.2.1 The order of emergence of functional structure

Based on the diagnostic criteria outlined in §3.1.3, the findings preliminarily reflect an order of appearance of structures that can be analysed in terms of three distinct stages across the ten children. The stages can be embedded into one of the three following descriptions. These exclude a period, in some children, where none of the structures are attested or, if some are attested, they represent isolated cases with unclear status.

1. The first CP-structures emerge. These typically included wh-questions, V2 (for Germanic), illocutionary complementisers (for Romance) and, more rarely, topicalisations/focalisations (typically MLU $\leq$ 1.5).

2. A second stage with CP- and TP-structures. Virtually no Split CP-structures are attested (between MLU 1.5 – 2.5).

3. A final stage with CP- and TP-structures, as well as the emergence of Split CP-structures (after MLU $\sim$ 2.5).

In this section, to facilitate readability, I present the results of the corpus study via summarised tables, which indicate the point of emergence of these stages and the structures attested (per the criteria outlined earlier). **Table 3.3** summarises the data obtained for the

Catalan, Italian and Spanish children and **Table 3.4** outlines the results for German and Dutch. TP-diagnostics are presented first, shaded in orange, followed by CP-diagnostics, in blue, and finally Split CP-diagnostics, in green. Note that the MLU ranges for each Stage are determined based on the first file in every child that satisfies the criteria for emergence for each domain, as defined above. In Germanic, where no wh-words are attested in a given period, but wh-less questions are present, this is noted as 'wh-less'. The full tables, with the structures attested in every file, are available in **Appendix A**.

Table 3.3 Summary of production data by the Romance children

Language	Child	MLU	S-Neg-V	S-Adv-V	S-Cl-V	Aux	Wh-Q	Top/Foc	Illoc	Embed	Split CP
Catalan	Laura	1.03-1.15									
		1.15-1.44					✓		✓		
		1.44-2.54	✓	✓	✓	✓	✓	✓	✓	✓	
		2.54[11]-3.18	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Gisela	1.04-1.58							✓		
		1.58-2.61	✓	✓	✓	✓	✓		✓		
		2.61-3.41	✓	✓	✓	✓	✓	✓	✓	✓	✓
Italian	Martina	1.26-1.57				✓					
		1.57[12]-1.66					✓		✓		
		1.66-2.69	✓			✓	✓	✓	✓	✓	
		2.69-2.55	✓		✓	✓	✓	✓	✓	✓	✓
	Rosa	1.27-1.78				✓	✓	✓			
		1.78-2.5	✓		✓	✓	✓	✓	✓	✓	
		2.5-3.24	✓		✓	✓	✓	✓	✓	✓	✓
Spanish	Irene	1.42-1.32[13]									
		1.32-1.61					✓				
		1.61-2.95	✓			✓	✓	✓	✓	✓	
		2.95-3.38	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Koki	1.96-2.69	✓		✓	✓	✓	✓	✓	✓	
		2.69-3.38	✓	✓	✓	✓	✓	✓	✓	✓	✓

[11] The first Split CP-structure in Laura is attested at a file with MLU 3.47. However, since the rest of the files in that period include substantially lower MLUs (around 2.5) and MLU 3.0 is not reached until later files, I adopt the next file's MLU value as the value of emergence for Split CP-structures in Laura. These are the values cited in any future tables and summaries.

[12] The same approach as for Laura (see footnote 11) was adopted for the first CP-structure in Martina (attested at a file with MLU 1.9, but surrounded by files at MLU ~ 1.5).

[13] The first file of this period has a slightly higher MLU than the final file, hence the apparently antichronological ordering. The same remark holds for other apparently antichronological timings.

Table 3.4 Summary of production data by the Germanic children

Language	Child	MLU	S-Neg-V	S-Adv-V	Aux	V2	Wh-Q	Y/N-Q	Top/Foc	Embed	Split CP
German	Kerstin	1.09-1.28					✓				
		1.28-1.59				✓	✓	✓	✓		
		1.59-2.32	✓	✓	✓	✓	✓	✓	✓	✓	
		2.32-2.89	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Simone	1.54-1.94				✓	✓		✓		
		1.94-2.78	✓	✓	✓	✓	✓	✓	✓	✓	
		2.78-3.46	✓	✓	✓	✓	✓	✓	✓	✓	✓
Dutch	Josse	1.2-1.69				✓	✓				
		1.69-3.57	✓	✓	✓	✓	✓	✓	✓	✓	
		3.57-3.2	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Sarah	1.12-1.09					Wh-less				
		1.09-1.68				✓	Wh-less				
		1.68-3.52	✓	✓	✓	✓	✓	✓	✓		
		3.52-4.92	✓	✓	✓	✓	✓	✓	✓	✓	✓

More concretely, these first-pass findings allow us to describe these preliminary three stages as follows.

There is an initial developmental period, let us call it **Stage 1**, where children appear to have a **'coarse-grained' discourse/interactional domain** (CP), hosting early wh-words, finite verb-movement in V2-languages, topics and illocutionary complementisers, as well as a thematic domain (*v*P)[14]. This preliminarily corresponds to the phasal layer in Biberauer and Roberts (2015), insofar as the system does not present evidence for either TP-based structure or cartographic-type constructions. However, this does not imply the bulk or entirety of CP-structures also emerge at this stage (or, indeed, that this coarse-grained CP is adult-like). In fact, as the tables above show, structures such as embedding or, in some children, topicalisation/focalisation typically emerge later than the earliest CP-structures (this will become important in §3.2.3). The key is that some CP-knowledge has emerged by this point. Stage 1 is absent in Gisela and Koki (the latter because her recordings start at a later developmental stage, MLU 1.94)[15], and is generally short in duration, two points which we will come back to.

Subsequently, in **Stage 2**, the **TP layer** seems to emerge. This is shown by the appearance of auxiliaries and word-order patterns that indicate a higher projection above *v*P but below the CP-domain. The structures present in Stage 1 continue being attested and being consolidated. At this stage, we hypothesise that the child is at the Core Functional Category level in Biberauer and Roberts (2015). There is no clear evidence yet that the child

[14] The latter is presupposed, as this study does not test for thematic, *v*P-internal developments (§3.1.3).

[15] Although her first file only contains CP-structures, I do not draw any conclusions from this, given the absence of recordings from earlier periods. I group this file with the next recordings as a single stage.

is operating on two or more separate projections within CP and, therefore, a 'Split CP' level of organisation.

Finally, in **Stage 3**, **subdivisions within the thus-far coarse-grained CP-domain** emerge. Production data that suggests at least two separate projections are available within CP appears at this stage, but not before. Again, the description in terms of emergent categories as in Biberauer and Roberts (2015) provides an account for this, as these cartographic-type structures lie at a higher level of granularity, predicted to be acquired later. Therefore, Stage 3 is the stage at which cartographic-type divisions appear to be made by the children studied; they are virtually unattested before this stage. Some counterexamples to this latter generalisation are attested in Stage 2, however. §3.2.3.3 comments on them, concluding that their status remains unclear and thus that they do not pose a problem for the broad empirical patterns at stake.

Sketched pictorially, the resulting learning path is the following (to be qualified):

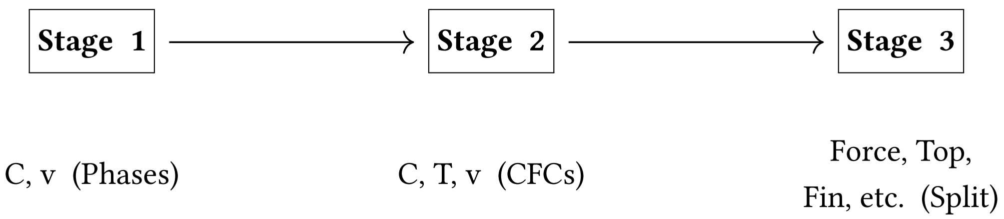


Figure 3.1 Surface stages observed in the data

Taken at face value, the tables above and in **Appendix A** generate a clear ordering of the emergence of the different functional domains analysed, namely CP → TP → Split CP. These developmental stages are examined in greater detail in §3.2.3, where I introduce the results from the supplementary studies in **Appendices B** and **C**. I will conclude that, whilst the possibility of Stage 1 being representationally realistic is not out of the question, the production data does not offer strong support in favour of it, even if it is certainly *compatible* with the existence of a Stage 1. At the same time, however, I will argue there is no support for a stage where TP emerges *before* the first CP-material and that the data is furthermore *incompatible* with it. In contrast, the additional quantitative analyses confirm that Stages 2 and 3 as currently characterised are strongly supported in all children. I will continue to separate the data in terms of these three stages, even if the supporting evidence for Stage 1 will be shown to be inconclusive. Before turning to these points, the next section discusses the predictive role of MLU in delimiting the stages, compared to age.

3.2.2 The effect of MLU and age

Further compressing the tables above into stages as ranges of MLUs, the ten children display the following stages ('✗' denotes the relevant stage is not detectable in the child data; '—' indicates absence of the relevant stage because the recordings start at more advanced developmental periods):

Table 3.5 MLU values of emergence across the three stages

	Stage 1	**Stage 2**	**Stage 3**
Laura	1.15	1.44	2.54
Gisela	✗	1.58	2.61
Martina	1.57	1.66	2.69
Rosa	1.27	1.78	2.5
Irene	1.32	1.61	2.95
Koki	—	1.96	2.69
Kerstin	1.28	1.59	2.32
Simone	1.54	1.94	2.78
Josse	1.2	1.69	3.57
Sarah	1.09	1.68	3.52

Overall, the three stages in **Table 3.5** emerge on average at the following MLUs. The average lengths (in months) of Stages 1 and 2 are also noted:

1. **Stage 1 – Discourse and thematic domains only**: MLU 1.3 (range 1.09 – 1.57; average length of 3.4 months).
2. **Stage 2 – Emergence of TP-structures**: MLU 1.69 (range 1.44 – 1.96; average length of 7.8 months).
3. **Stage 3 – Emergence of a Split CP-domain**: MLU 2.82 (range 2.32 – 3.57).

The analysis of the emergence of the various structures studied indicates that there exists a relationship between the MLU of the child (understood as a metric of syntactic development) and the types of structures produced, with a clear relative order among those structures. The MLU variances of all three stages also appear fairly homogenous (with the variances for Stages 1, 2 and 3 being 0.0298214, 0.0261567 and 0.1755122, respectively). Overall, the MLU values for each stages are therefore stable. However, in both Gisela and

Koki, Stage 1 is absent. This is for distinct reasons: in Gisela, it is possible that this stage is truly non-existent, but its absence could be an artifact of the near-absence of recordings between MLU 1.5 – 2.5, with most recordings involving values above 2.5. In Koki, this is simply because her recordings begin at a later developmental stage. Josse and Sarah's Stage 3 begins comparatively late. I comment on this in §3.2.3.3. Finally, Wilcoxon Signed Rank tests comparing the MLU at the point of emergence of the three stages output statistically significant differences for Stage 2 *vs* Stage 3, as well as Stage 1 *vs* Stage 3 (Stage 1 *vs* Stage 2, $z = -0.5879995$, $p = 0.56$; Stage 2 *vs* Stage 3, $z = -3.902964$, $p < 0.001$; Stage 1 *vs* Stage 3, $z = -3.702658$, $p < 0.001$). The lack of significant results for Stage 1 *vs* Stage 2 will align with the empirical discussion in §3.2[16].

In contrast, age is not a reliable predictive factor. **Table 3.2.2** displays the age of emergence of the three sets of structures analysed. Age of emergence presents high variance. The variances for each variable (Stage 1, 2 and 3) when measured in terms of age (in months) vary substantially (being 5.928571, 13.55556 and 21.73333, respectively). Few correlations can be made with respect to age and syntactic knowledge.

Table 3.6 Age of emergence across the three stages

	Stage 1	**Stage 2**	**Stage 3**
Laura	1;10.22	2;04.11	3;03.21
Gisela	✗	2;04.25	2;08.00
Martina	1;08.02	1;10.29	2;04.13
Rosa	1;07.13	2;04.29	2;10.14
Irene	1;04.16	1;06.16	1;11.13
Koki	–	1;07.20	2;04.18
Kerstin	1;10.03	2;01.01	2;09.11
Simone	1;09.11	1;10.28	2;06.23
Josse	2;00.07	2;02.08	2;11.09
Sarah	1;10.05	2;00.17	3;00.19

The age at which children produce specific structures can thus vary significantly from child to child. There is vast variation: at 2;00, Josse is still at Stage 1, while Irene reaches

[16]The non-parametric Wilcoxon Signed Rank test was used for this statistical analysis, and other statistical analyses in this thesis, over paired-samples t-tests, as the t-tests' normality assumption was not met in the datasets. This was always checked with a Shapiro-Wilk test for normality, which, in this case, returned significant results ($p = .01$). If a t-test had been used, however, it would also have returned similar results for all comparisons (Stage 1 *vs* Stage 2, $t(17) = -0.4364$, $p = 0.67$; Stage 2 *vs* Stage 3, $t(19) = 10.7$, $p < 0.001$; Stage 1 *vs* Stage 3, $t(17) = 5.977$, $p < 0.001$). All tests were carried out in R (R Core Team, 2022).

Stage 3 at 1;11. Similarly, Rosa and Laura reach Stage 2 by the same age (2;04), but Laura reaches Stage 3 around half a year later (2;10 *vs* 3;03). The most obvious 'outlier' is Irene, who reaches Stage 3 by 1;11, well before most children reach Stage 2 or even Stage 1. Echoing previous literature (§3.1.2), these results further endorse the need to make crosslinguistic comparisons and predictions with another less variable metric, such as MLU.

3.2.3 The data and stages in detail

We now discuss the three stages in more detail and give some examples of the kinds of structures attested in each of them. Some theoretical discussion is offered, but the implications of this data are expanded on in §3.3.

3.2.3.1 Stage 1: Discourse comes first

Stage 1 is a structurally fairly limited stage, owing to the short length of children's utterances (typically between MLU 1 – 1.5). Nonetheless, most children appear to produce some CP-structures at this first stage.

As for Romance, the earliest recordings contain wh-questions, illocutionary complementisers and some topics/foci. Some examples of these early CP-structures are given in (22):

(22) a. **Què** és? (Laura, MLU 1.3)
what be.3SG
'What is it?'

b. Ai, **que** crema! (Laura, MLU 1.35)
ouch that.EXCL burn.3SG
'Ouch, it's burning!'

c. **Que** cau! (Laura, MLU 1.3)
that.EXCL fall.3SG
'It's falling!'

d. **Dov**'è? (Rosa, MLU 1.39)
where-be.3SG
'Where is it?'

e. **Pello** oio io [: quello voglio io] (Rosa, MLU 1.75)
that want.1SG I
'That, I want.'

f. **Ove** va chetto? (Martina, MLU 1.57)
where go.3SG this
'Where does this one go?'

g. **Che** gira (Martina, MLU 1.9)
that.EXCL stir.3SG
'He/she stirs (it).'

h. ¿**Quién** es? (Irene, MLU 1.32)
who be.3SG
'Who is it?'

i. ¿**Qué** hay ahí mamá? (Irene, MLU 1.69)
what have.3SG there mum
'What is there mum?'

The illustrations in (22) reveal that the nature of these early CP-elements is rather limited, mostly restricted to copular wh-questions and illocutionary complementisers. The function of (22g) is ambiguous based on the surrounding context[17]; it appears to indicate either surprise or exclamation, but it might also have an imperative interpretation (see footnote 17 for conversational context)[18]. The above examples represent a near-complete list of the wh-questions uttered by Laura and Irene. However, wh-questions are less sporadic in the Italian data: Martina produces 6 wh-questions (3 being non-copular, of which 2 are non-finite). In Rosa's case, these are abundant from Stage 1 and involve a variety of wh-words (55 examples are found, all copular; see **Appendix B**, **Table B.1.1**). 3 topics/foci are also apparent, see (22e), albeit right before the start of Stage 2. There are no instances of topicalisations/focalisations at this stage in the other Romance children. Some evidence for a C-head, however, potentially comes from early illocutionary complementisers, with 4 examples in Laura (and one isolated instance in Gisela at MLU 1.13). I come back to this discussion below (**Table 3.8**), concluding that the evidence for a Stage 1 in the Romance children exists but remains weak at best.

German and Dutch children likewise present evidence of CP-structures at this stage, and to a greater extent than Romance children. Firstly, there is evidence of early wh-questions and some topicalisations, as shown in (23).

[17]Example (22g) was uttered in the following conversational context:

*MOT: devo girare il caffé ('I have to stir the coffee')
*CHI: **che gira** ('It/he/she stirs')
*MOT: giri te? ('Are you stirring (it)?')
*MOT: gira ancora, gira ('Stir (it) again, stir (it)')

[18]This apparent illocutionary complementiser by Martina is, according to three informants, ungrammatical in Tuscan Italian (her variety), but generally licit in Catalan/Spanish (my own judgements). This observation is not restricted to this example or child, but there is no space to expand on its implications (see Bosch, accepted, for discussion on Catalan, Spanish and preliminary Italian data).

(23) a. **Wo** ’s den Nina? (Kerstin, MLU 1.28)
where be.3SG the.ACC Nina
‘Where is Nina?’

b. **Wat** doet ie nou? (Josse, MLU 1.59)
what do.3SG he now
‘What is he doing now?’

c. **Jetzt** kommt ein Affe (Kerstin, MLU 1.76)
now come.3SG a monkey
‘Now a monkey comes.’

d. **Blume** eßt nicht (Simone, MLU 1.69)
flower eat.3SG not
‘Flower, he/she doesn’t eat.’

e. **Bau(e)n** will ich (Simone, MLU 1.71)
build.INF want.1SG I
‘To build (something), I want.’

Like the Romance children, however, the number of wh-questions produced at this stage is highly reduced: (23b) is one of the three wh-questions Josse produces at Stage 1 and Sarah produces no wh-questions (setting aside 4 and 1 wh-less questions, respectively). Kerstin and Simone, on the other hand, produce 3 wh-questions and 16 wh-questions, respectively. Although more plentiful than in the Dutch children, Kerstin’s questions are copular, and so do not display lexical variety. More critically, all of Simone’s wh-questions take the form *Wo ist. . . ?* (‘where is...?’), suggesting they could be rote-learned forms (as also noted in Clahsen et al., 1993, p. 419). In this respect, the data for wh-questions alone is not sufficient to warrant the claim that Dutch/German children have CP at Stage 1. Nonetheless, some topics have clearly emerged in both Kerstin and Simone by their Stage 1 (6 and 3 are attested, respectively) and one yes/no question is present in Kerstin (*Siehste?*, a colloquial contraction of *Siehst du?* ‘you see?’). Topicalisations and yes/no questions are absent in the Dutch children.

Setting this aside, there are, however, noticeable signs of a distinction between non-finite verbs (which remain in-situ and yield OV-order where an object is present) and finite verbs (which can be shown to have raised outside the *v*P)[19].

(24) a. Ich will de(r) Ball (Kerstin, MLU 1.58)
I want the ball
‘I want the ball.’

[19]Note that several examples at Stage 1 (and also 2) feature Root Infinitives (RIs) (this is also apparent from **Appendix C**). It is beyond the scope of this work to provide an analysis of RIs crosslinguistically.

b. Wieder aufräumen (Kerstin, MLU 1.34)
again up.tidy.INF
'To tidy up again.'

c. Ma(ch) ma(l) Schuh ab (Kerstin, MLU 1.53)
make.IMP.2SG once shoe off
'Take off (the) shoe.'

d. Mag nicht Kuche(n) backe(n) (Simone, MLU 1.62)
like.1SG not cakes bake.INF
'I don't like to bake cakes.'

e. Bluse auszieh(e)n (Simone, MLU 1.62)
blouse off.take.INF
'To take off (the) blouse.'

f. Komt ie (Josse, MLU 1.2)
come.3SG he
'He comes.'

g. Koffie drinken (Josse, MLU 1.44)
coffee drink.INF
'To drink coffee.'

h. Doet niet meer (Josse, MLU 1.55)
do.3SG not more
'He/she doesn't do (it) anymore.'

i. I(k)moe(t) ete(n) (Sarah, MLU 1.16)
I-must.1SG eat.INF
'I have to eat.'

j. I(k)wi(l) farkj [: varkentje] (Sarah, MLU 1.25)
I-want.1SG pig.DIM
'I want piglet.'

k. Boekje lezen (Sarah, MLU 1.68)
book.DIM read.INF
'To read book/booklet.'

Stage 1 thus already contains evidence of children's early deployment of the V2-rule in Germanic (more generally, early acquisition of V2 has been noted in, i.a., Clahsen, 1986; Clahsen and Penke, 1992; Verrips and Weissenborn, 1992, for German, and Jordens, 1990; van Kampen, 2010, for Dutch). The examples in (24) show that finite verbs such as *mag* 'may/be allowed to' or *will* 'want' are raised to the V2-position. This is shown by the fact that these verbs move beyond negation (24d, 24h) and also do not present the OV-order expected if the verb were in its base-position (cf. 24g and 24k). In contrast, non-finite verbs present the opposite pattern: they do not raise beyond adverbs or negation (24d) and present surface OV-orders (as in *Bluse ausziehen*, 24e). A full analysis of the number of the verbs in V2/V1 and V-final position in the first files of Stages 1-2 is available in **Appendix B**, which, i.a., also shows V2 is almost error-less from the start.

Additionally, the V2-position is not constrained to specific classes of verbs (cf. Jordens, 1990, who suggested that no activity verbs initially appear in V2-position in Dutch and de Haan, 1987, who proposed instead that only modals and auxiliaries occupy this position). Instead, activity, accomplishment and achievement verbs (Vendler, 1957) are attested in V2/V1-position (**Table 3.7**). This suggests the V2-rule is being applied productively in the child grammar, without any lexico-syntactic conditioning factors (see also Verrips and Weissenborn, 1992, pp. 290-293):

Table 3.7 Classes of verbs in V2/V1-position at Stage 1

	Verbs in V2/V1-position
Kerstin	• *Statives*: *sein* ‘to be’, *es gibt* ‘there is/are’, *gehen* ‘to be working/functioning’, *haben* to have’. • *Activities*: *runtergehen* ‘to go down’, *gehen* ‘to go’, *machen* ‘to make’, *gucken* ‘to look’, *aufmachen* ‘to open up’, *lassen* ‘to let’, *fahren* ‘to drive’, *halten* ‘to hold’. • *Achievements*: *werden* ‘to become’, *fallen* ‘to fall’, *kommen* ‘to come’. • *Modals*: *wollen* ‘to want’.
Simone	• *Statives*: *sein* ‘to be’, *gehen* ‘to be working/functioning’, *es gibt* ‘there is/are’, *haben* ‘to have’. • *Activities*: *suchen* ‘to search for’, *gehen* ‘to go’, *machen* ‘to make’, *weinen* ‘to cry’, *schlafen* ‘to sleep’, *bauen* ‘to build’. • *Accomplishments*: *essen* ‘to eat’. • *Achievements*: *kommen* ‘to come’, *passen* ‘to fit’. • *Modals*: *mögen*, ‘to like/may’, *wollen* ‘to want’.

Josse	• *Statives*: *zijn* 'to be', *kennen* 'to know'. • *Activities*: *kijken* 'to look', *gaan* 'to go', *doen* 'to do'. • *Achievements*: *kommen* 'to come'. • *Modals*: *mogen*, 'may/be allowed to', *moeten* 'to have to/must', *willen* 'to want'.
Sarah	• *Statives*: *zijn* 'to be', *hebben* 'to have'. • *Activities*: *kijken* 'to look', *doen* 'to do'. • *Modals*: *moeten* 'to have to/must', *kunnen*, 'can', *willen* 'to want'.

As a result, it seems reasonable to suggest that some knowledge of the V2-system has emerged by this point and that the verb can be shown to raise to a higher projection, which, we suggest, could be (phasal) C. Importantly, the knowledge of the V2-system established at both Stage 1 and Stage 2 (described in the following section) involves at most 'basic' V-to-C. The specific type of V2-system (e.g., 'Force-V2' or 'Fin-V2' in Poletto's, 2002, and Wolfe's, 2015, typology; see also Biberauer and Roberts, 2015) is expected to be consolidated at Stage 3, when cartographic divisions are seemingly effected. It follows, too, that V3-configurations requiring finer-grained cartographic-type structures (e.g., those introduced in §3.1.3.1) are expected not to emerge until Stage 3. This is borne out and will be discussed in §3.2.3.3.

Reiterating the above, however, while there is some evidence for CP-based material, particularly in Germanic children, it can be shown to rest on weak foundations in most children. The CP-structures found during Stage 1 across the ten children are summarised in **Table 3.8** (based on **Appendix B**). The first value in parentheses next to the number of wh-questions indicates the proportion of copular wh-questions and the second one in German and Dutch signals the proportion of wh-less questions. As noted in the previous section, Gisela presents no evidence for Stage 1, besides an isolated illocutionary complementiser at 1.13 MLU. However, this is at least partly due to the near-absence of recordings between 1.5 – 2.5 MLU. Koki also displays no Stage 1, but this is simply because her recordings start at a later stage (MLU 1.94; Stage 2).

Table 3.8 CP-structures produced at Stage 1 and its length

	V2	Wh-Q	Y/N-Q	Top/Foc	Illoc	Embed	Length
Laura		1 (1)		0	4	0	1;10.22-2;04.11 (MLU 1.15-1.44)
Gisela							✗
Martina		6 (3)[20]		0	1	0	1;08.02-1;10.29 (MLU 1.57-1.66)
Rosa		55 (55)		3	0	0	1;07.13-2;04.29 (MLU 1.27-1.78)
Irene		2 (1)		0	0	0	1;04.16-1;06.16a (MLU 1.32-1.61)
Koki							–
Kerstin	✓	3 (3, 0)	1	6		0	1;10.03-2;01.01 (MLU 1.28-1.59)
Simone	✓	16 (16, 0)	0	3		0	1;09.11-1;10.28 (MLU 1.54-1.94)
Josse	✓	7 (0, 4)	0	0		0	2;00.07-2;02.08 (MLU 1.2-1.69)
Sarah	✓	1 (1, 1)	0	0		0	1;10.05-2;00.17 (MLU 1.09-1.68)

Altogether, evidence of CP-knowledge at Stage 1 does exist (especially regarding knowledge of V2, early wh-questions/topics in Italian, topics in German and illocutionary complementisers in Laura), but it is limited and weak in many children. It remains an open question whether Stage 1 truly exists in most of the children studied, but in §3.3.2 I come back to behavioural evidence in favour of extremely early acquisition of wh-dependencies. This will strengthen the point of needing to develop our knowledge of this stage (MLU $\sim 1.0 - 1.5$), especially via comprehension-based techniques. It is these latter techniques that can elucidate most clearly whether the earliest CP-related productions have to be interpreted as mere rote-learned and unproductive forms, or whether some principled syntactic process already subserves them.

Note that an alternative analysis exists whereby the earliest CP-related structures provide evidence for a projection distinct from CP. For example, Clahsen (1990) argues

[20] 2 wh-questions are non-finite.

for a FiniteP (FP)[21] analysis of early V2, before subject-verb agreement and TP have been acquired (see also Diercks et al., 2023). The projection cannot be equated with TP/IP or AgrP, as the specifier position is not restricted to subjects; it can also accommodate topics, as we have seen for Simone and Rosa. Instead, an FP proposal suggests that only finite verbs can occur in the head position of FP and it cannot be filled with lexical complementisers and wh-elements. Later FP develops into CP (see also Clahsen and Penke, 1992; Clahsen et al., 1993). This analysis would not hold up to scrutiny for Stage 1 on both empirical and conceptual grounds, particularly for Romance languages. Firstly, some wh-questions are attested at this stage (not all of which are formulaic questions), and so do not fulfil Clahsen's criterion for an FP analysis. Furthermore, and most seriously, in the Catalan and Italian data, some early illocutionary complementisers are also apparent, again undermining the possibility that it is a projection whose head is restricted to [+finite] verbs which is responsible for the patterns in Stage 1.

Besides, one can also level theoretical and ontological arguments against an FP analysis. If assuming a UG-given spine, FP would presumably instantiate an 'extraneous' or 'placeholder'-like projection, not directly encoded in the UG-given spine (e.g., the traditional CP>TP>*v*P sequence or the more elaborate version in **Figure 2.2**). This would defeat entirely the point behind maturational approaches, which make predictions that rest directly on structural height in a universal spine (FP would not form part of it, thereby opening the door to further stipulations; this also holds for continuity). In a neo-emergentist approach to categories, the FP analysis would carry even less power: if FP is a precursor to the adult-like CP (Clahsen's suggestion), then it *would* be an emergent (coarse-grained) CP displaying some of the characteristics of adult-like CPs. The virtue of a neo-emergentist approach is precisely to allow for structurally-homologous development into ever more fine-grained and adult-like categories. Thus, I discard FP and TP analyses of early CP-related structures.

It is also worth noting that the examples in (22-24), which we suggest might present some evidence for early CP-knowledge, clearly also display (correct) agreement-marking. This appears to undermine the claim that TP has not emerged by this stage, if we take agreement to require licensing via a T-mediated Agree relation. These examples could open the door to considering a TP-before-CP acquisition timeline, in view of the weak support for Stage 1 just outlined. This would be a hasty conclusion, however: the results from the additional morphological study in **Appendix C** revealed that subject-verb agreement morphemes become productive (i.e., are used with ≥ 3 verbs; recall §3.1.5) across the three persons at around the same time as the structural diagnostics emerge, namely, the period thus far classed as Stage 2. At Stage 1, only 3sg is used productively for most of the children,

[21]This FiniteP is distinct from Rizzi's (1997), FinP, which postdates Clahsen's work.

with few signs that 1sg, and especially 2sg, have been acquired by this point; primarily 1sg reaches productivity at Stage 2, but in 4 children only. There is only one child that presents evidence of productive subject-verb agreement with all three persons at Stage 1, Rosa[22]; however, productivity is attained right before Stage 2 (2;04.23 *vs* 2;04.29). **Table 3.9** summarises the development of subject-verb agreement morphology across all ten children (note that Gisela lacks a Stage 1, hence the 'Stage 0' label, corresponding to the stage where no structures are attested; recall also that Koki's recordings start at Stage 2).

Table 3.9 Onset of productivity of verbal morphology during Stages 1 and 2

	3sg	**1sg**	**2sg**	**Other**
Laura	1.15 MLU	1.35 MLU[23]	1.64 MLU	1pl: 2.42 MLU (Stage 2)
	Stage 1	Stage 1	Stage 2	
Gisela	1.5 MLU	2.32 MLU	Not productive	
	Stage 0	Stage 2		
Martina	1.57 MLU	1.57 MLU	1.66 MLU	
	Stage 1	Stage 1	Stage 2	
Rosa	1.41 MLU	1.75 MLU	1.54 MLU	3pl: 2.78 MLU (Stage 2)
	Stage 1	Stage 1	Stage 1	
Irene	1.57 MLU	2.28 MLU	Not productive	
	Stage 1	Stage 2		
Koki	1.96 MLU	1.96 MLU	2.54 MLU	3pl: 2.47 MLU (Stage 2)
	Stage 2	Stage 2	Stage 2	1pl: 2.51 MLU (Stage 2)
Kerstin	1.76 MLU	1.58 MLU	1.72 MLU	3pl: 1.67 MLU (Stage 2)
	Stage 1	Stage 2	Stage 2	1pl: 1.65 MLU (Stage 2)
Simone	1.62 MLU	1.71 MLU	2.09 MLU	3pl: 2.52 MLU (Stage 2)
	Stage 1	Stage 1	Stage 2	1pl: 3.35 MLU (Stage 2)
Josse	1.55 MLU	1.94 MLU[24]	2.14 MLU	3pl: 2.8 MLU (Stage 2)
	Stage 1	Stage 2	Stage 2	1pl: 2.46 MLU (Stage 2)
Sarah	1.68 MLU	1.68 MLU	2.11 MLU	3pl: 2.46 MLU (Stage 2)
	Stage 2	Stage 2	Stage 2	1pl: 2.66 MLU (Stage 2)

[22]This remark becomes less contentful with Dutch, given its impoverished inflectional morphology. The whole (present-tense) inflectional paradigm is simply composed of -*en* (↔ [plural]), -Ø (↔ [speaker]) and -*t* (↔ [ ], the 'default' form in the language, typically for 3sg; see Don et al., 2013). Therefore, the division above between 1sg, 2sg and 3sg is more artificial and a full, future analysis of Dutch morphological development would have to take these caveats into consideration.

Furthermore, tense-marking (e.g., past and future tense) and auxiliaries are also generally absent at this stage and increase substantially in frequency at Stage 2. **Table 3.10** summarises this.

Table 3.10 Production of tensed forms and auxiliaries at Stage 1 *vs* Stage 2

	Tense-markings		**Auxiliaries**	
Child	Stage 1	Stage 2	Stage 1	Stage 2
Laura	0	15	0	34
Gisela	✗	0	✗	1
Martina	0	4	0[25]	7
Rosa	2	3	2	15
Irene	4	83	0	2
Koki	—	37	—	14
Kerstin	0	2	0	12
Simone	2	15	0	52
Josse	0	13	0	16
Sarah	0	38	0	78

Table 3.10 thus shows that morphological indicators of tense-marking are non-existent in most children at Stage 1. There are some exceptions. Irene produces 4 past-tense forms. These instantiate separate repetitions of the same three verbs (*cayó* 'fell.3SG', *marchó* 'left.3SG'and *acabó* 'finished.3SG'), also repeated abundantly in later files. Crucially, none of these verbs appear in their present-tense form at this stage. This converges with similar findings in child Greek: Tsimpli (1992, p. 60) also notes a restricted number of past-tense forms at the earliest stage in the two children studied. These are used with very few verbs, which also never appear in their present-tense forms, and, suggestively, some of these early past-tense-marked verbs coincide with those used by Irene (namely, *epese* 'fell.3SG', *teliose*

[23]However, one of the three verbs in 1SG in this recording is produced in the utterance *no sé* 'I don't know', repeated multiple times. It is possible this could be a rote-learned form, in which case 1SG does not reach productivity until MLU 1.64 (Stage 2).

[24]In Josse's case, productivity in 1SG is technically reached earlier at MLU 1.59. However, this is exclusively due to negated modals (e.g., *mag niet* 'may not'). These forms are known to be rote-learned at early stages (see Hoekstra and Jordens, 1996, for syntactic and prosodic evidence in Dutch). Importantly, no other 1SG verbs are found at or before MLU 1.59; productivity with other verbs (including positive modals) is not reached until MLU 1.94. In light of this, I excluded 1SG negated modals.

[25]Martina presents one isolated auxiliary right before the start of Stage 1. Although not included in the table above as it falls outside Stage 1, I comment on it below.

'finished.3SG'). Rosa and Simone also produce two past-tense forms (*fumava* 'it was smoking' and *era* 'was.3SG' for Rosa, and *ging* 'went.3SG', the latter repeated twice by Simone in the same file). Since the number of past-form verbs is extremely limited at this stage, these observations are likely indicators that tense is not productively used by this point. As argued by Tsimpli (1992, 2005), it is also possible these early past-tense-marked forms could signal aspectual distinctions (e.g., [±PAST] reinterpreted as completion *vs* non-completion of an event; Tsimpli, 1992, p. 67), rather than presence of the functional category T (see also the Aspect First Hypothesis, i.a., Antinucci and Miller, 1976; Bloom et al., 1980; Wagner, 2001). Altogether, it does not seem unlikely that these early past-tense-marked forms either could be rote-learned or may not instantiate tense as this relates to TP directly.

Besides past-tense marking, Rosa presents two auxiliaries in Stage 1. These involve repetitions of the same utterance (*s'è rotto*, 'it has broken'). Martina also produces an isolated auxiliary *ho rotto* ('I have broken (it)'), albeit right before the start of Stage 1 (see footnote 25). Two auxiliaries (both with the same verb) are arguably not sufficient to suggest *passato prossimo* is productive at this MLU stage (see Caprin and Guasti, 2009, for a similar results in other Italian children)[26]. The remaining children present no morphological evidence for tense-related knowledge at Stage 1. Added to the complete lack of structural indicators of TP and the visible increase in auxiliaries and past-tense markings at Stage 2, we conclude that there is no clear-cut evidence that the children are harnessing a further projection above the *v*P-domain (which is not CP) at Stage 1.

The findings of **Appendix C** parallel independent conclusions in the acquisition literature, such as Tsimpli's (1992) observations above on early Greek (see also Tsimpli, 2005). Caprin and Guasti (2009) also demonstrated that only 3SG is used productively during the MLU 1.0 – 1.5 stage (in 26.7% and 6.7% out of 15 children, for 1st and 2nd conjungation, respectively). Their criteria for productivity are either 75% of correct usage in obligatory contexts or use of a morpheme with (minimally) three distinct verbs. It is not until 1.5 – 2.0 MLU that 1SG and 2SG become productive in 47.4% and 15.8% out of 19 children, respectively. Later on (MLU 2.0 – 3.0), other morphemes start emerging (e.g., 1PL or 3PL). The present study also replicates previous findings for German: it is well-known that the 3SG *-t* ending is attested from the earliest stages (Clahsen, 1990, 1991; Jordens, 1990); however, it is also frequently noted that the complete agreement paradigm is missing early on and that agreement errors are not uncommon (Clahsen and Penke, 1992; see discussion in

[26]This is also why the two auxiliaries in Rosa, produced in two relatively contiguous files in Stage 1, were not deemed sufficient for emergence of TP (see **Table 3.3** earlier and **Appendix A**). Their identical and possibly unproductive nature favoured an adjustment of the emergence criterion in §3.1.4 and we counted instead the next file with TP-structures (MLU 1.78). The decision also aligns with Caprin and Guasti (2009), who establish that *passato prossimo* is very rare at MLU 1 – 1.5, with only 3 correct uses across 15 children.

Tsimpli, 1992, pp. 35-36). Similarly, abundance of 3SG, over 1SG/2SG, at the earliest stages is also documented for Finnish (Saikkonen, 2018) and for Catalan and Spanish (alongside the imperative form, which is often syncretic with 3SG; Grinstead, 2000; Salustri and Hyams, 2006).

Overall, I conclude, with some of the aforementioned works, that the presence (or absence) of inflectional morphemes is itself not sufficient to establish that the abstract properties and dependencies associated with this functional morpheme are unambiguously present in the child grammar. The paradigm of inflectional morphemes must be shown to be clearly productive and/or to be combined with structural evidence for those functional categories[27]. Neither of these conditions seems to be fulfilled at Stage 1. In turn, these results also provide some motivation for the claim that V2-movement with finite verbs is available independently of children's knowledge of subject-verb agreement morphology (see Verrips and Weissenborn's, 1992, Independence Hypothesis for German, and further support in Westergaard, 2009, for Norwegian and van Kampen, 2010, for Dutch; *pace* Clahsen and Penke, 1992). This may also give learnability-based support to a 'Swooping' theory of V2 (where V-to-C can be learned independently of movement to T; Holmberg and Platzack, 1995; Vikner, 1995), over a 'Stepping' approach (where V-to-C necessarily proceeds via intermediate movement through T; Platzack, 1986).

The upshot of this discussion is that some limited evidence for CP-knowledge at Stage 1 exists, yet it is fragile in several children and non-existent in one (Gisela). Additionally, the duration of this stage is generally short (around 3.4 months and 0.4 MLU points on average)[28]. At the same time, though, there is no convincing evidence to suggest that TP is accessible at this stage: while there is some agreement-marking, we have shown it to be largely unproductive (besides 3SG) and there is near-total absence of structural cues for TP. More to the point, the data does *not* support a TP-before-CP analysis and it is furthermore *incompatible* with it: some CP-knowledge is in place as soon as evidence for TP emerges, making a TP-before-CP analysis a non-starter. Therefore, based exclusively on the production data, I conclude that if any representational knowledge has emerged by Stage 1, the data speaks in favour of it being coarse-grained (phasal) knowledge of the CP-domain (besides *v*P). Despite the apparently inconclusive nature of Stage 1, the collective findings

[27] The aetiology of the correct agreement-markings (if not a T-mediated Agree relation) is still unsettled, as well as which ϕ-features children have acquired by this point. Space limitations preclude a discussion of these points, so I will simply argue that they are not evidence for TP.

[28] This is not a knock-down argument, however. Perkins and Lidz (2021) show that the make-up of child grammars appears to be able to change substantially in a very short span of time: by 15 months children grasp basic argument-structural relations, yet by 18 months they already appear to represent wh-dependencies (see also §3.3).

of Stage 1 and Stage 2, which I turn to now, will suffice to adjudicate between possible theoretical accounts.

3.2.3.2 Stage 2: The emergence of TP

Shortly after the initial block of CP-structures have appeared in the production data, these continue to be syntactically elaborated and, across all ten children, the first instances of structures at the TP-level appear.

The earliest TP-related structures in the Romance children include auxiliaries and word-order patterns indicating possible subject-raising beyond adverbs, negation and clitics (25).

(25) a. **Aquest ja** està acabat (Laura, MLU 1.98)
this already be.3SG finished
'This one is already finished.'

b. L'**he** perdut la boleta (Laura, MLU 1.78)
CL.DO=AUX.HAVE.1SG lose.PTCP the ball.DIM
'I have lost the small ball.'

c. Que **jo no** puc (Laura, MLU 2.42)
that.QUOT I not can.1SG
'(I've said) I can't.'

d. No, **jo em** vull treure els patins (Gisela, MLU 2.32)
no I CL.REFL= want.1SG take.off.INF the skates
'No, I want to take off the skates.'

e. **Hai** vitto uno gatto (Martina, MLU 1.99)
AUX.HAVE.2SG see.PTCP a cat
'You have seen a cat.'

f. **Io non** voglio, dalli (Martina, MLU 1.98)
I not want.1SG give.IMP.2SG=CL.DO
'I don't want to, give them (to me).'

g. **Io ne** vojo due! (Rosa, MLU 2.14)
I CL.PART= want.1SG two
'I want two of them.'

h. **Ese no** e(s) (Irene, MLU 1.94)
that not be.3SG
'It's not that one.'

i. **Él no** (e)stá abajo (Koki, MLU 2.51)
he not be.3SG down
'He is not under/downstairs.'

j. Mira dónde **está** bañando a nena (Koki, MLU 2.47)
look.IMP.2SG where AUX.BE.3SG bathe.GER the girl
'Look where he/she is bathing the girl.'

Structures such as (25e) and (25c) suggest that a projection distinct from both *v*P and CP is required to model these early structures. At Stage 2, children's productions can accommodate auxiliaries (25b, 25e, 25j). Here, subjects can also precede anaphoric negation (25f, 25h, 25i), adverbs (25a) or clitics (25d, 25g), indicating that a TP-projection hosting constituents like auxiliaries or raised subjects is present at this stage. Such structures are categorically absent at Stage 1. Crucially, this projection *has* to be distinct from CP. This is most clearly illustrated in *Que jo no puc* (25c), given the presence of quotative *que*, preceding the subject *jo*.

The same points apply to the Germanic examples in (26), which display the first instances of S-Neg-V, S-Adv-V orders and auxiliaries (in all examples of S-Neg-V/S-Adv-V orders the subject is *not* in the CP-domain, as per the requirement in §3.1.4).

(26) a. Will **da** [: willst du] **auch** Brot? (Kerstin, MLU 1.58)
want you also bread
'Do you also want bread?'

b. Nehmen **wir erstmal** den Ball (Kerstin, MLU 1.58)
take.1PL we first time the.ACC ball
'First we take the ball.'

c. Habe(n) **wir noch nich(t)** (Kerstin, MLU 1.65)
have.1PL we yet not
'We haven't yet.'

d. Puppa **is(t)** (ka)puttgegange(n) (Simone, MLU 1.93)
Puppa AUX.BE.3SG break.go.PTCP
'Puppa (the doll) has broken.'

e. Geht **das nich(t)** (Simone, MLU 1.93)
go.3SG that not
'That doesn't work.'

f. Daar ging **ie wel** eens mee naar bed en zo (Josse, MLU 1.69)
there go.PST.3SG he well once along to bed and so
'He used to go to bed with it and so on.'

g. Stift kan **ze niet** (Josse, MLU 1.94)
pin can.3SG she not
'As for pin/pinning, she cannot.'

h. Hij **(i)s** weggelopen (Sarah, MLU 1.88)
he AUX.BE.3SG away.run.PTCP
'He ran away.'

i. Zie **je** **wel**? (Sarah, MLU 2.11)
see.2SG you well
'See?'

j. (I)k **he(b)** nie(t) (ge)schomme(ld)! (Sarah, MLU 2.05)
I AUX.HAVE.1SG not swing.PTCP
'I have not swung!'

As with Romance, this stage features many TP-constructions. These include auxiliaries, such as (26d, 26h, 26j), and word-order patterns which suggest subject-raising to the TP-domain, either by raising above adverbs (26a, 26b, 26c, 26f, 26i) or above negation (26c, 26e, 26g). For instance, in *Stift kan ze niet* (26g), the subject *ze* is assumed to be in the TP-internal subject position, not in the CP-domain, given the presence of a topic *stift* in the left periphery. As this subject precedes negation, we have to conclude it lies in an intermediate projection between CP and *v*P, namely TP. Examples (26c, 26e) with narrative inversion and with the subject preceding negation and/or a modal particle are another clear instance of this word-order pattern.

The emergence of structural diagnostics is also accompanied by an increase in morphological tense-marking and auxiliaries during Stage 2. **Table 3.11** incorporates the earlier information from **Table 3.10** and compares it to the emergence of structural diagnostics for TP (in the two cases where 2SG does not reach productivity at Stage 2, the recording at which both 3SG and 1SG become productive is reported instead and is marked with an asterisk '*').

Table 3.11 Emergence of structural *vs* morphological diagnostics for TP

	TP-structures	**Productive agreement**
Laura	2;04.11 (file 7)	2;05.08 (file 8)
	1.44 MLU	1.64 MLU
Gisela	2;04.25 (file 9)	*2;06.23 (file 10)
	1.58 MLU	2.32 MLU
Martina	1;10.29 (file 5)	1;10.29 (file 5)
	1.66 MLU	1.66 MLU
Rosa	2;04.29 (file 9)	2;04.23 (file 8)
	1.78 MLU	1.75 MLU
Irene	1;06.16a (file 11)	*1;08.26 (file 21)
	1.61 MLU	2.28 MLU
Koki	1;09.18 (file 2)	1;09.18 (file 2)

	2.54 MLU	2.54 MLU
Kerstin	2;01.01 (file 21)	2;02.20 (file 23)
	1.59 MLU	1.72 MLU
Simone	1;10.28 (file 6)	2;02.20 (file 26)
	1.94 MLU	2.09 MLU
Josse	2;02.08 (file 5)	2;04.11 (file 8)
	1.69 MLU	2.14 MLU
Sarah	2;00.17 (file 9)	2;02.18 (file 11)
	1.68 MLU	2.11 MLU

Importantly, then, morphological and syntactic diagnostics for TP converge considerably and point to knowledge of TP-based dependencies emerging at Stage 2, with both indicators crucially being absent at Stage 1. It also coincides with an increase in non-3SG forms (principally, 1SG and 2SG; see **Appendix C** for this data).

At this stage, it is also apparent that *both* CP and TP must be at play to account for the structures produced by the ten children studied, as already indicated by some examples above (e.g., 25c, with an illocutionary complementiser, and 26g, with topicalisation). We established in §3.2.3.1 that the evidence for a Stage 1 is unsettled. What emerges from Stage 2, however, is that evidence for CP-knowledge becomes plentiful, suggesting that it *is* in play at Stage 2, alongside TP, and strengthening our earlier point that a TP-before-CP analysis is unworkable. (27) gives some examples of matrix CP-structures attested at the start of Stage 2 (see **Table 3.13** later and **Appendix B** for greater detail on their productivity):

(27) a. **Un gorro** ara trec (Laura, MLU 1.88)
a hat now take.out.1SG
'A hat, I'm taking out now.'

b. **Que** no la puc treure (Gisela, MLU 2.32)
that.QUOT not CL.DO= can.1SG take.out.INF
'(I've said) I can't take it out.'

c. **Bere** voglio (Martina, MLU 1.66)
drink.INF want.1SG
'To drink, I want.'

d. **Que'** lo l'apri? (Rosa, MLU 1.78)
that CL.DO= CL.DO=open.2SG
'That, do you open it?'

e. **Como** dijo? (Irene, MLU 2.28)
how say.PRFV.3SG
'How did he/she say?'

f. **Ése** se lo dio la tía Carol (Koki, MLU 2.47)
that CL.IO= CL.DO= give.PRFV.3SG the aunt Carol
'That one, aunt Carol gave it to him/her.'

g. **Was** soll ich denn? (Kerstin, MLU 1.86)
what should.1SG I then
'What should I do then?'

h. **Bist** du Elke? (Simone, MLU 2.31)
be.2SG you Elke
'Are you Elke?'

i. **Daar** gaat ie (Josse, MLU 2.14)
there go.3SG he
'There he goes.'

j. **Zie** je dat? (Sarah, MLU 2.11)
see.2SG you that
'Do you see that?'

At the end of this stage finite embedding markers also emerge for all children. Some instances are given in (28):

(28) 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. È l'omo **che** ccrive (Martina, MLU 2.55)
be.3SG the-man that write.3SG
'It/he is the man that writes.'

c. Deja esto ahí **para que** no se rompe (Koki, MLU 3.08)
leave.IMP.2SG this there so that not CL.REFL= break.3SG
'Leave this there so that it doesn't break.'

d. **Weil** der Tommy gebißen hat (Simone, MLU 2.52)
because the Tommy bite.PTCP AUX.HAVE.3SG
'Because Tommy bit (something/someone).'

e. Ik krijg deze **a(l)s** ik uit de crèche ben (Sarah, MLU 3.52)
I get.1SG this when I out the nursery be.1SG
'I get this when I am out of the nursery.'

The late emergence of subordination contrasts with the relatively early emergence of the other CP-diagnostics and moreover represents a theoretically consequential finding.

The simultaneous emergence of embedding markers and topicalisation in Friedmann et al. (2021) (as part of their final stage in **Figure 2.2**) is, in several instances, not replicated. At most, it only holds for some Romance children. Topicalisation emerges considerably before embedding markers in several of the children studied (**Table 3.12**). The difference is particularly stark in the Germanic children. A paired-samples t-test comparing the MLU at the point of emergence of both structures confirms that there is a highly statistically significant difference between the two ($t(19) = 7.1188$, $p < 0.001$). In order to facilitate a comparison with Friedmann et al.'s (2021) conjecture (which bears on topicalisation and embedding, *not* the left-peripherally lower focalisation), this analysis excluded fronted objects in Romance that are possible instances of foci or that are ambiguous between focalisation/topicalisation[29].

Table 3.12 Emergence of topicalisation *vs* embedding markers

	Topicalisation	**Embedding**
Laura	2;08.03	3;00.02
	1.88 MLU	2.42 MLU
Gisela	2;08.00	2;08.00 (**same file**)
	2.61 MLU	2.61 MLU
Martina	1;11.20	1;11.20 (**same file**)
	1.99 MLU	1.99 MLU
Rosa	2;04.29	2;06.29
	1.77 MLU	2.6 MLU
Irene	1;08.09b	1;09.10
	2.24 MLU	3.28 MLU
Koki	1;11.25	1;11.25 (**same file**)
	2.47 MLU	2.47 MLU
Kerstin	2;00.05	2;07.23
	1.76 MLU	2.13 MLU
Simone	1;10.20	2;04.20
	1.62 MLU	1.96 MLU

[29]Note, however, that this might be an underestimate of the occurrence of topics and does not control for a scenario where topicalisations superficially emerge later than focalisations simply because children occasionally omit their accompanying resumptive clitics (children are known to produce object clitics at early stages, but crucially also to omit them; see Guasti, 1993; Schaeffer, 2000; Wexler et al., 2004; Babyonyshev and Marin, 2006; Gavarró et al., 2010; though cf. Wexler et al., 2004; Gavarró et al., 2010, on Spanish). The data for the earliest fronted objects, either topics or foci, can be found in **Appendices A** and **B**.

Josse	2;03.28	2;09.02
	1.94 MLU	2.42 MLU
Sarah	2;00.17	3;00.19
	1.68 MLU	3.52 MLU
Average	2.0 MLU	2.54 MLU

As a result, an initial generalisation can be made from **Table 3.12**, which we will harness in §3.3 to draw theoretical conclusions:

Generalisation 1 (Topics Precede Finite Embedding)**.** Topicalisation emerges before finite embedding markers in most of the children studied.

Now, the above does not necessarily imply that this developmental pattern will always replicate itself in other languages or children (this is a separate empirical matter); the possibility remains open that topics emerge together with embedding in languages like Hebrew (as Friedmann et al., 2021, and Friedmann and Reznick, 2021, argue). What it *does* imply, however, is that topics do not consistently emerge alongside embedding markers across all languages and children, a finding which contradicts the absolute expectations of the Growing Trees Hypothesis. I come back to this point in §3.3. The timing discrepancies between Germanic and Romance in the emergence of topicalisation are also telling. This is particularly true in light of literature pointing to Germanic topicalisation being formally much simpler than the options available in (some) Romance varieties (see Biberauer and Roberts, 2015, pp. 15-18, for an explicit comment), aligning with the earlier emergence of topics in the former.

I conclude from the above, then, that CP has also emerged by this point and that a TP-before-CP approach is unfeasible. A summary of the types and tokens of CP-structures attested throughout Stage 2 is given in **Table 3.13**. Stage 2 is short in Gisela (2 files). Nonetheless, her developmental patterns are still in line with the ones reported for the rest of the children.

Table 3.13 CP-structures produced at Stages 2 and its length

	V2	Wh-Q	Y/N-Q	Top/Foc	Illoc	Embed	Length
Laura		14 (8)		4	38	4	2;04.11-3;03.21 (MLU 1.44-2.54)
Gisela		1 (0)		0	6	0	2;04.25-2;08.00 (MLU 1.58-2.61)
Martina		15 (13)		3	6	8	1;10.29-2;04.13 (MLU 1.66-2.69)
Rosa		78 (76)		9	3	8	2;04.29-2;10.14 (MLU 1.78-2.5)
Irene		16 (5)		3	10	4	1;06.16a-1;11.13 (MLU 1.61-2.95)
Koki		32 (27)		7	2	4	1;07.20-2;04.18 (MLU 1.96-2.69)
Kerstin	✓	13 (11, 2)[30]	20	21		1	2;01.01-2;09.11 (MLU 1.59-2.32)
Simone	✓	150 (138, 7)	3	102		24	1;10.28-2;06.23 (MLU 1.94-2.78)
Josse	✓	55 (36, 21)[31]	37	68		1	2;02.08-2;11.09 (MLU 1.69-3.57)
Sarah	✓	123 (99, 45)	104	116		0	2;00.17-3;00.19 (MLU 1.68-3.52)

As with Stage 1, an FP analysis of these early CP-productions is not available (*pace* Diercks et al., 2023, pp. 125-128), much less a TP analysis. This is particularly evident in Stage 2. An FP analysis of these CP-related constructions would leave unexplained why children can produce several constructions which are hallmarks of adult-like CPs, well before embedding complementisers appear, and seemingly in a highly productive manner, as **Table 3.13** and **Appendix B** indicate.

These theoretical matters aside, there are no indications at either Stage 1 or 2 that the child is operating on cartographic-type materialisations of the CP-domain. At no point can we detect a combination of left-peripheral elements of the sort in §3.1.3.3. There is,

[30] 1 wh-question is non-finite.
[31] As for footnote 30.

however, a restricted set of dubious counterexamples to this latter claim, which I turn to in the following section.

All in all, from the resulting discussion on Stages 1 and 2, the following generalisation can legitimately be made:

Generalisation 2 (Early Acquisition of CP)**.** CP-structures emerge early on in the developmental data, either simultaneously with TP-material or, less clearly, possibly shortly before the first TP-structures.

Generalisation 2 then also places emphasis on another key finding, namely that there is no evidence for a TP-before-CP stage. When evidence for TP emerges, indication for CP-knowledge is *also* in play. Whether CP truly emerges *before* TP remains open. However, I will argue in §3.3 that this generalisation, alongside **Generalisation 1**, suffices to draw some important conclusions regarding the feasibility of the theoretical approaches introduced in §2, particularly once the predictive power of these two generalisations is combined with **Generalisation 3**, introduced in the following section.

3.2.3.3 Stage 3: Refining the system — emergent cartographic divisions

It is at the final Stage 3 that a novel empirical generalisation emerges, which I will argue below is theoretically very significant. At this stage, evidence for co-occurrence of multiple CP-elements emerges robustly, something that was virtually absent during Stages 1 and 2.

Some structures with multiple elements in the left periphery attested at this stage in Catalan, Italian and Spanish are shown in (29). A very common kind of structure displaying evidence for a Split CP involves topics preceding wh-questions (29a, 29c, 29d, 29f). Example (29b) includes recursive topics. (29e) presents a quotative complementiser preceding a wh-word.

(29) a. I **el ioc'ioc**, **què** fa? (Laura, MLU 2.91)
and the ioc-ioc what do.3SG
'And the ioc-ioc (onomatopoeia), what does it do?'

b. **Ara**, **aquest**, no puc (Gisela, MLU 2.71)
now this no can.1SG
'Now, this one, I can't.'

c. **Quetto**, **cosa** fa? (Martina, MLU 2.37)
this what do.3SG
'This, what does it do?'

d. **Questo**, **dove** si mette? (Rosa, MLU 2.5)
this where CL.REFL= put.3SG
'This, where do you put it?'

e. ¿**Que** **qué** nos falta? (Irene, MLU 3.23)
that.QUOT what CL.IO= lack.3SG
'(I've asked) What are we missing?'

f. ¿**Esas bonitas cortinas**, **dónde** las compraron? (Koki, MLU 2.71)
those beautiful curtains where CL.DO= buy.PRFV.3PL
'Those beautiful curtains, where did they buy them?'

At Stage 3, at least three separate slots are also detectable in the production data. Example (30a) combines what appears to be an adverbial topic *ahir*, followed by a second topic/focus, alongside another topic (the CLLD topic *a jo*). Similarly, example (30b) displays a *que sí que* structure (see §3.1.3.3) and (30c-30d) instantiate two topics (*ahora* and *ésto*/*todos los muchachos*) preceding a wh-element (*(a) dónde*). By MLU ≥ 2.5 we conclude, then, that all children can harness two, and, based on some examples, seemingly also three, separate projections within CP.

(30) Three co-occurring left-peripheral elements

a. **Ahir**, **tu**, **a jo**, me vas dir que me (Laura, MLU 2.91)
yesterday you to me CL.IO= AUX.GO.2SG tell that CL.IO=
l'explicaries avui
CL.DO=explain.COND.2SG today
'Yesterday, YOU told ME that you would explain it to me today.'

b. **Que** **sí** **que** vol (Gisela, MLU 2.61)
that.QUOT yes that.QUOT want.3SG
'(I've said) He/she does want to.'

c. ¿Y **ahora**, **ésto**, **dónde** se mete? (Irene, MLU 4.01)
and now this where CL.IMPRS= put.3SG
'Now, this, where does it go?'

d. ¿**Ahora**, **todos los muchachos**, **a dónde** se van a (Koki, MLU 3.38)
now all the boys to where CL.REFL= go.3PL to
ir?
go.INF
'Now, all the boys, where are they going to go?'

In Germanic languages, Stage 3 involves the emergence of the restricted kinds of V3-orders discussed in §3.1.3.3. Frame-setters yielding V3-orders are illustrated in (31). (31a) and (31d) contain instances of Hanging Topics, while (31b, 31c, 31e, 31f) present cases of Contrastive Left-Dislocation.

(31) a. Und **dann das** brauch(e) ich (Kerstin, MLU 2.68)
and then that need.1SG I
'And then I need that.'

b. **Den Teddy**, ich halt **den** so (Simone, MLU 2.78)
the.ACC Teddy I hold.1SG PRON.ACC so
'The Teddy, I hold him like this.'

c. **Die Leute**, **die** ham [: haben] des (Simone, MLU 4.89)
the people they have.3PL that
puttemacht [: kaputtgemacht]
break.make.PTCP
'The people, THEY have broken that.'

d. En **toen daar** was Hanneke weggegaan (Josse, MLU 3.91)
and then there AUX.BE.PST.3SG Hanneke away.go.PTCP
'And then, there Hanneke had left.'

e. **De staart**, **die** had je goed gevonden (Sarah, MLU 3.15)
the tail that AUX.HAVE.PST.2SG you good find.PTCP
'The tail, that you found good.'

f. **Andere boekje**, **die** lees ik nooit (Sarah, MLU 2.89)
other book.DIM that read.1SG I never
'The other book/booklet, I never read (it).'

Finally, cases of conditional/temporal clauses with adverbial resumption are exemplified below. These are attested primarily in the Dutch data and are abundant in both Sarah and Josse after MLU 2.5 (36 and 12 instances are recorded, respectively). 1 clear-cut example is found in the German data (in Simone).

(32) V3-order in conditional/temporal clauses with resumptive *dann/dan*

a. **Als** ik naar de grote school gaat, **dan** komt (Josse, MLU 3.78)
when/if I to the big school go.3SG then come.3SG
sneeuw
snow
'When I go to the big school, then snow comes.'

b. **Als** je drie bent, **dan** mag je niet van de (Sarah, MLU 3.05)
when/if you three be.2SG then may.2SG you not from the
politie op school
police in school
'When/if you're three, the police don't let you into school.'

c. **Wenn** du keine Suppe aufißt, **dann** hau ich (Simone, MLU 4.89)
when/if you no soup up.eat.3SG then smack I
den dir
PRON.ACC you.DAT
'If you don't eat up the soup, I'll smack you.'

Table 3.14 contrasts the emergence of the earliest CP- *vs* Split CP-structures. It is clear from **Table 3.14** that all the relevant split-structures analysed in the ten children emerge around or after MLU 2.5, but never earlier. It is possible, nonetheless, that this impressionistic MLU ~ 2.5 'threshold' will not replicate itself in other children. The overall pattern would remain unaffected so long as Stage 3 remains a clearly separate stage.

Table 3.14 Emergence of CP- *vs* Split CP-structures

	CP-structures	**Split CP-structures**
Laura	1;10.22	3;03.21
	1.15 MLU	2.54 MLU
Gisela	2;04.25	2;08.00
	1.58 MLU	2.61 MLU
Martina	1;08.02	2;04.13
	1.57 MLU	2.69 MLU
Rosa	1;07.13	2;10.14
	1.27 MLU	2.5 MLU
Irene	1;04.16	1;11.13
	1.32 MLU	2.95 MLU
Koki	1;07.20	2;04.18
	1.96 MLU	2.69 MLU
Kerstin	1;10.03	2;09.11
	1.28 MLU	2.32 MLU
Simone	1;09.11	2;06.23
	1.54 MLU	2.78 MLU
Josse	2;00.07	2;11.09
	1.2 MLU	3.57 MLU
Sarah	1;10.05	3;00.19
	1.09 MLU	3.52 MLU

While most of the MLU values for Stage 3 in **Table 3.14** are homogenous (MLU ~ 2.5), there is a glaring exception in the Dutch children, who display emergence of Split CP-structures at MLU 3.57 and 3.52. Now, developmental mismatches in Germanic languages whose V2-systems are analysed essentially identically are unexpected: why should the Dutch children pattern so differently to the German ones? Space constraints preclude a full discussion and, additionally, two children is too small a sample to draw conclusive statements; this pattern may be coincidental. I leave this to future research.

When comparing the emergence of cartographic-type structures to other child productions, we can also note that subordination appears either before the first structures with co-occurring left-peripheral elements emerge or, in some cases, simultaneously. This is shown in **Table 3.15**:

Table 3.15 Emergence of embedding markers *vs* Split CP-structures

	Embedding	**Split CP-structures**
Laura	3;00.02	3;03.21
	2.42 MLU	2.54 MLU
Gisela	2;08.00	2;08.00 (**same file**)
	2.61 MLU	2.61 MLU
Martina	1;11.20	2;04.13
	1.99 MLU	2.69 MLU
Rosa	2;06.29	2;10.14
	2.6 MLU	2.5 MLU
Irene	1;09.10	1;11.13
	3.28 MLU	2.95 MLU
Koki	1;11.25	2;04.18
	2.47 MLU	2.69 MLU
Kerstin	2;07.23	2;09.11
	2.13 MLU	2.32 MLU
Simone	2;04.20	2;06.23
	1.96 MLU	2.78 MLU
Josse	2;09.02	2;11.09
	2.42 MLU	3.57 MLU
Sarah	3;00.19	3;00.19 (**same file**)
	3.52 MLU	3.52 MLU

That the emergence of embedding consistently precedes or coincides with the emergence of Split CP-structures is remarkable and very suggestive. At the very least, their relative timing is highly indicative of the fact that Split CP-structures are difficult to acquire, seemingly as difficult as embedding constructions or, in some children, possibly harder. This is unexpected if part of or the entire cartographic-type left periphery is accessible to children by this point (cf. Friedmann et al., 2021). Therefore, this correlation suggests that their systematic absence at earlier developmental stages is also unlikely to be accidental.

The emergence of these structures is moreover sudden and 'explosive' in the production data, such that for the vast majority of children no Split CP-structures or extremely few are detectable before MLU ~ 2.5; they abruptly increase in quantity after this point in a statistically very significant manner, as shown by the results of a Wilcoxon Signed Rank test ($z = -2.949874$, $p = 0.003$)[32].

Table 3.16 Production of Split CP-structures before and after MLU ~ 2.5

	Before MLU ~ 2.5	**After MLU ~ 2.5**	%
Laura	1	20	4.8-95.2%
Gisela	0	9	0-100%
Martina	0	5	0-100%
Rosa	1	31	3.1-96.9%
Irene	0	85	0-100%
Koki	0	41	0-100 %
Kerstin	3	4	42.9-57.1%
Simone	2	7	22.2-77.8%
Josse	1	19	5-95%
Sarah	2	51	3.8-96.2%
Total	10	272	3.5-96.5%

However, as shown in **Table 3.16**, the generalisation is not without its exceptions, and is weakest in Kerstin. A total of 10 apparent exceptions are present, which emerge before the other Split CP-structures. Most represent isolated examples, not fulfilling the emergence criterion introduced in §3.1.4; however, the possibility that they are theoretically problematic needs to be ruled out. All 10 apparent counterexamples are given in (33):

(33) a. I **això**, **què** és? (Laura, MLU 1.98)
and this what be.3SG
'And this, what is it?'

[32] A paired-samples t-test was not used for this analysis as the asssumption of normality was not met. A Shapiro-Wilk test for normality on the variable with the production of Split CP-structures indicates this, which returned significant results ($p < 0.001$). Note that, in any event, both the non-parametric test above and the parametric test return significant results, albeit much more conservative ones (with the results of the t-test being $t(19) = 2.5626$, $p = 0.02$). Crucially, however, the Wilcoxon Signed Rank test result is likely to be underpowered, due to the test's treatment of zero-differences and zero-values (see Pratt, 1959), which do not enter the computation and thus reduce the sample size (affecting, therefore, Gisela, Martina, Irene and Koki).

b. **Quello cos**'è? (Rosa, MLU 1.5)
that what-be.3SG
'That, what is it?'

c. **Des hier ich** will anziehen (Kerstin, MLU 1.86)
that here I want.1SG on.put.INF
'That here, I want to put (it) on.'

d. **Das Buch** <**das** will ich> [?] (Kerstin, MLU 1.67)
the book that want.1SG I
'The book, I want that.'

e. **Bubi**, **der** fällt runter (Kerstin, MLU 1.78)
Bubi he fall.3SG down
'Bubi, he is falling down.'

f. **Der Puppe die** geht da auf (Simone, MLU 1.93)
the doll she go.3SG there on
'The doll, it goes on there.'

g. **Die Eisebahn da** kann durchfahr(e)n (Simone, MLU 1.82)
the train there can.3SG through.drive.INF
'The train can go through there.'

h. **Nou jij** mag (Josse, MLU 2.14)
now you may.2SG
'Now you can.'

i. **Nou ik** heb is [?] (ka)pot (ge)maakt (Sarah, MLU 2.46)
now I AUX.HAVE.1SG is [?] broken make.PTCP
'Now I have broken (it).'

j. **Nou ik** wi(l) tellen (Sarah, MLU 2.47)
now I want.1SG count.INF
'Now I want to count.'

On the one hand, we observe ungrammatical sentences and errors in German and Dutch with surface V3-order and, on the other, genuine apparent counterexamples (where the relevant structure is generated via grammatical means in the corresponding language). As regards the former, both (33c) and (33g) represent ungrammatical utterances in adult German and do not instantiate any of the V3-structures in §3.1.3.3. (33f) is also ungrammatical, but in this case the structure superficially instantiates Contrastive Left-Dislocation, albeit with a gender error (*die Puppe* 'the doll' is feminine in German, and so should not be accompanied by *der*, the nominative masculine article). Since these structures are ungrammatical and their analysis is unclear (especially 33c and 33g), I will not consider examples (33c, 33f, 33g) to be serious counterexamples.

Furthermore, closer examination of genuine, grammatical counterexamples reveals that they form clear homogeneous groups. Firstly, all counterexamples from Catalan and Italian children reduce to examples with a 'that/this, what is it?' form (there are no counterexamples in the Spanish children). That no other counterexamples with a different verb or structure exist in the Romance dataset is significant. The cause behind the early emergence of these forms before the rest of the Split CP-structures is unclear, however: they could instantiate cases of more primitive topic-comment structures, of the sort described in Manfred Krifka's work, where the deictic *that* would be the topic and the wh-question, the comment (e.g., [that$_{topic}$ [what is it$_{comment}$]]; Krifka, 2008; Bambini and Torregrossa, 2010; see also van Kampen, 2010, on early topic-comment structures)[33]. It is also possible that, prosodically and syntactically, these are unconnected utterances, where the topic 'that/this' simply signals a deictic/pointing gesture to a salient element in the discourse, which is not syntactically connected to the wh-question. Both possibilities involve speculation and the latter is currently impossible to probe, given the unavailability of recordings from which their prosodic structure could be analysed. Given their homogenous nature, I set these examples aside, but why structures such as 'that/this + what + copula' should emerge before the remaining Split CP-diagnostics is an open question.

In Germanic, the counterexamples are either V3-orders with Dutch *nou* (3 cases between Josse and Sarah) or cases of CLDs. The former make up 30% of all counterexamples. Considering the adverbial nature of *nou*, it is not inconceivable that examples such as (33i) could be due to adjunction of *nou* (following Lebeaux, 1988; de Villiers, 1991; Roeper, 1992; de Cat, 2007; Biberauer, 2018, where adjunction is harnessed as a default analysis by the child), as opposed to signalling a specifier of FrameP (or equivalent). This analysis might also be applicable to the Romance counterexamples above (see, e.g., de Cat, 2007, on child French dislocation). Prosodic data would be needed to carefully test for this, however.

Finally, (33d) and (33e) are fairly clear cases of CLDs and are thus more problematic. As with cases of 'that/this + what + copula' structures, there always remains the possibility that *Bubi* and *das Buch* were produced as separate utterances which are syntactically and prosodically unconnected. (33d) is also an utterance interrupted by an interlocutor (hence the '<>' notation); this could have played a role in the repetition of the demonstrative *das*. Again, prosodic data would be needed to make conclusive statements on both fronts; however, recordings are unavailable for Kerstin and Simone.

Altogether, there are reasons to think not all counterexamples seriously defy the overall developmental norm. In some instances, these represent ungrammatical utterances, with unclear analysis. In others, they appear to form homogenous 'groups' of counterexamples,

[33]Thanks to Theresa Biberauer for suggesting this possibility.

which themselves merit future scrutiny. I have argued that their homogenous nature suggests precisely that their exceptionality lies in their shared syntactic properties, possibly without constituting clear-cut evidence for finer-grained representations. However, that 8/10 of the counterexamples should be found in Germanic, which we noted in §2 has been argued to instantiate a 'bundled' left periphery, is also suggestive and deserves more attention. All things considered, I will take these examples not to pose serious challenges for the general trend detected, which still appears remarkably robust.

Setting these ambiguous counterexamples aside, the finding from this third stage is striking in light of approaches that assume a UG-given spine to be accessible to the child (either partially at some stages, as in maturation, or totally, as in continuity). The late and, crucially, *sudden* emergence of this 'Split CP-domain' is also contrasted with the very early availability of discourse-related structures. The latter appear at the earliest stage (approximately 1-2 years in age and around 1 MLU point before Split CP-structures) and become consolidated at Stage 2. Based on the data presented, it appears that, while peripheral positions are used very early on, children operate initially on a coarse-grained discourse domain (interpreted as an emergent CP allowing a single head and specifier). Later developmental stages refine this coarse-grained CP and successively divide it into separate projections, such that this refined CP can now accommodate co-occurring CP-structures.

I therefore contend that fine-grained syntactic heads of the sort in Rizzi (1997) are *emergent* and are not available as soon as CP becomes apparent at either Stage 1 or 2, *contra* the expectations of maturational approaches (and recent cartographic continuity approaches like Westergaard, 2009; see §3.3 for discussion on the micro-cues model). Instead, they can be shown to systematically emerge at this third stage, as part of a refinement process over developmentally-earlier structures (the Phasal and CFC domains). The data in this thesis offers the first preliminary developmental evidence for this claim. Significantly, this corroborates independent syntactic and biolinguistic work advocating for an 'emergent cartography' in, i.a., Ramchand and Svenonius (2014), Svenonius (2016), Scontras et al. (2017), Mišmaš et al. (2018), Leivada and Westergaard (2019), Marušič et al. (2019), Larson (2021) and Ramchand (2023)[34]. As a result, the following empirical generalisation surfaces:

Generalisation 3 (Cartography is Emergent)**.** Evidence for cartographic-type structure within CP systematically and abruptly emerges at a later developmental stage, elaborating on developmentally-prior structure (a 'basic' CP).

As I will suggest in §3.3, the combination of **Generalisation 2** and **Generalisation 3** generates an inevitable contradiction in maturational approaches such as the Growing Trees

[34]See also Lee and Cournane (2019), Cournane and Klævik-Pettersen (accepted) for other relevant work on diachronic syntax.

Hypothesis and in any nativist approach to cartographic categories. The best resolution of this conflict, I will propose, involves dropping innate categories from our ontology, and by implication, dropping maturation as both an empirically and biolinguistically feasible mechanism. As a result of **Generalisation 3**, the strengths of a neo-emergentist approach to syntactic categories surface in a way that is novel in the context of syntactic development.

In the following discussion section, then, I will interpret the switch from Stage 2 to Stage 3 as due to a *representational difference*. However, I also explore three alternative explanations of **Generalisation 3** in Bosch (2023), which appeal to sentence planning limitations, derivational complexity and utterance length. There, I argue that, whilst these possible explanations could in principle account for *some* of the patterns described, they often make incorrect predictions. In all three cases, they fail to explain the overall pattern in its entirety, with its impressionistic 'suddenness' remaining mysterious.

3.3 Discussion

3.3.1 Theoretical implications: on the need for 'ontological flexibility'

The results of the corpus study challenge the predictions of both the bottom-up and (generalised) inward-growing maturational proposals outlined in §2. Recall that, firstly, according to a bottom-up approach, TP-structures should have emerged before the first CP-structures. We have shown this not to be borne out in the children studied. Strict bottom-up development looks inaccurate: there is no clear evidence for emergence of TP before CP. Rather, they either emerge simultaneously at Stage 2 or, less clearly, CP emerges at Stage 1 shortly before the first TP-structures at Stage 2 (**Generalisation 2**). Further, topicalisation often emerges well before embedding (**Generalisation 1**), and illocutionary complementisers are also early arrivals (see **Appendix B** and Bosch, accepted), both contradicting the categorical and bottom-up expectations of the Growing Trees Hypothesis. However, the fatal problem lies in the late emergence of cartographic-like structures (**Generalisation 3**). Approaches that assume an adult-like UG-given spine cannot accommodate this observation — the early emergence of C *versus* the comparatively late emergence of a finer-grained organisation of C is enigmatic.

Although the general inward-growing hypothesis formulated in §2 (adapted from Heim and Wiltschko, 2021) is more compatible with our data than a bottom-up approach, it is still not fully corroborated. While this inward-growing hypothesis correctly predicts CP-structures to emerge early on, its maturational nature, associated with a UG-given cartographic spine, means child categories will always have adult-like cartographic counterparts. This incorrectly predicts the availability of a cartographic-type level of organisation as soon as CP-structures emerge (and likely before TP-structures mature). Both maturational proposals tested here, then, either undergenerate (assuming *no* CP-knowledge at the first stages of acquisition) or overgenerate (assuming *too fine-grained* knowledge of the CP-domain as soon as the latter emerges). They cannot get the balance 'right' as far as granularity is concerned: they adopt UG-given spines of fixed granularity.

A similar point holds for the Lexical Learning Hypothesis in the Weak Continuity literature, briefly introduced in §2. Assuming, as we have been, that a cartographic layer of structural organisation is required to capture some crosslinguistic patterns, a 'modern' version of Lexical Learning would predict the relevant cartographic heads to be acquired as soon as a structure/feature associated with that head is learned; for example, TopP, QP and ForceP would be acquired as soon as the child learns constructions involving topicalisations, wh-questions and complementisers, respectively. However, the data presented rather suggests that these constructions (wh-questioning, topicalisation, illocutionary complementisers, etc.) are learned *before* evidence for use of multiple cartographic heads is apparent in the production data. Therefore, the Lexical Learning Hypothesis overgenerates like (generalised) inward maturation, assuming fine-grained syntactic structure well before the child is visibly operating on a Split CP.

In sum, the results stemming from **Generalisations** 2 and 3 generate an irresolvable contradiction in maturational (and continuity) approaches: the empirical data presented forces us to assume some early representation of the CP-domain (around the same time as TP emerges, at the very least), but this representation cannot be finer-grained than Chomsky's CFCs at this earlier stage, lest the theory overgenerate. Expressly, the generalisations force a theoretical approach which assumes some early knowledge of the CP-domain that does not develop into a cartographic-type left periphery until later. Development cannot simply 'recapitulate' tree geometry for this effect to be generated. It would necessitate categorial systems and their granularity to be 'malleable' during development, an analysis which is definitionally incompatible with proposed maturational approaches assuming exclusively innate categorial priors, such as a cartographic spine. We conclude, then, that a maturational approach is too rigid. Notice, additionally, that non-cartographic maturational approaches do not technically face this contradiction, by virtue of assuming more impoverished clausal spines (such as Heim and Wiltschko's, 2021, Inward Growing Hypothesis and the earlier 'Empty Middle' approach by Galasso, 2003). However, they arguably do not possess the tools to account for the contrast between Stages 1-2 and Stage 3. While not necessarily ruled out as possible developmental theories, a separate explanation would be required to accommodate the patterns in Stage 3. Put differently, the three problems laid out in what follows would still apply to them. This also carries over to other emergentist approaches that make developmental predictions based instead on 'derivational ordering or timing', such as Diercks and Bossi's (2021) and Diercks et al.'s (2023) Developmental Minimalist Syntax. To be clear, the problem is *not* restricted to maturational approaches, but its paradoxical nature becomes most severely engraved in this type of theory.

If the interpretation of the patterns is on the right track, it follows that at least three interrelated (but dissociable) problems have to be addressed by any theoretical approach to development:

1. **The Directionality Problem**
   Developmental stages cannot straightforwardly be characterised as recapitulating a UG-given tree, either bottom-up or inwardly.

2. **The Granularity Mismatch Problem**
   Child grammars and adult grammars do not (always) operate at the same level of grain. Acquisition may be characterised by *increasing granularity*, thus bridging the initial existing 'gap' with the adult system.

3. **The Comparability Problem**
   The 'fine-tuning' and structural homology patterns detected are not readily accounted for by innate categorial ontologies. Child categories may not always have direct counterparts in an adult-like UG-given spine.

I propose here that these problems are inexorable in existent maturational approaches, especially those of a cartographic nature, and can only be resolved by dropping two assumptions in turn: firstly, foregoing an adult-like UG-given spine as the point of departure for the child and, by implication, a maturational mechanism that dictates the development of this latter spine. Doing so immediately lends us the additional ontological flexibility needed to account for (1-3): these problems dissolve with the abandonment of an innate set of categorial priors, particularly once neo-emergentism is adopted. In a neo-emergentist approach to acquisition, later developmental stages refine developmentally-prior categories (structural homology), leading to the increasing granularity patterns described. It also becomes misleading to describe development as 'recapitulating' an innate spine in any given directionality and to assume that child categorial systems always have to be in direct correspondence with the categories postulated in adult-like UG-given spines. In effect, I propose that only an approach that assumes categories to be *ontologically flexible* (i.e., truly emergent) can readily capture the patterns in **Generalisations 1-3**[35].

As a result, then, no such contradiction or problems arise with Biberauer and Roberts's (2015) successive-division learning path, which by design expects changes in representational granularity as development unfolds. The results of the corpus study are therefore consistent with it, as it correctly assumes that a coarse-grained, 'basic' representation of CP (e.g., Phasal or CFC) will emerge early, but that fine-grained knowledge requiring cartographic-type projections will *not* emerge until a later developmental stage. Its prediction regarding early emergence of CP before TP is not strongly supported, but the data does not contradict it and it is, at the very least, compatible with it. Evidence for multiple left-peripheral elements, however, consistently emerges at a later stage, typically after MLU ~ 2.5. In this respect, then, neo-emergentism converges with Friedmann et al.'s (2021) Growing Trees Hypothesis, insofar as the *cartographic* left periphery develops at later developmental stages. However, neo-emergentism emphasises there is, in fact, some basic structural representation of the CP before it is elaborated on at this third stage.

These stages, we suggest, are in line with an emergentist approach to categories, which affords them ontological flexibility (Hale, 1986). Emergent categories lend themselves more readily to the fine-tuning and structural homology patterns described. We conclude, therefore, that the current data presents strong novel support for neo-emergentist approaches to syntactic categories and, particularly, the first preliminary set of longitudinal evidence in favour of the pathway in Biberauer and Roberts (2015).

Before moving on to §4, where I propose a developmentally- and biolinguistically-oriented model of the stages in Biberauer and Roberts (2015), I turn to independent findings in the acquisition literature that support some of the patterns reported.

3.3.2 Crosslinguistic comparison

The developmental trends presented in this study share many similarities with data reported for other languages. For example, Santelmann (2003) reports very early acquisition of wh-movement in child Swedish, which appears productive by MLU 1.5 (see also Saikkonen, 2018, for relevant data for child Finnish, and Soares, 2003, for European Portuguese). Santelmann

[35]There still remains the possibility that an innate *template* of conceptual categories guides this emergence process (e.g., the $classification \rightarrow point\ of\ view \rightarrow anchoring \rightarrow linking$ template in Song, 2019, from Wiltschko, 2014, which mediates all granularity levels; see also Ramchand and Svenonius, 2014). I leave this open. Nonetheless, the point that categorial divisions would have to be emergent to capture the patterns still applies.

(2003) shows, among other things, that there is variety in the subjects used and that there is evidence that the verb has moved (cf. 35a and 35b)[36]:

(34) Variation in "what/where is this" questions

a. A e bilen? (Tor 2;2, MLU 1.5)
where is car.DEF
'Where is the car?'

b. A e pappa? (Tor 2;2, MLU 1.5)
where is papa
'Where is papa?'

(35) Questions with and without moved verb

a. Vad det **är**? (Frank 2;0, MLU 1.48)
what that is
'What is that?'

b. Va **är** denna? (Frank 2;0, MLU 1.48)
what is this
'What is this?'

In addition, Westergaard (2009) also shows that the cue for general V2 and also some yes/no questions appear to have emerged by around MLU 1.5:

(36) Early examples of V2 in child Norwegian

a. Der er stor stor Ole Brumm (Ann.01 1;8, MLU 1.4)
there is big big Ole Brumm
'There is (a) big big Winnie the Pooh.'

b. Det er vovva (Ina.01 1;8, MLU 1.53)
it is doggie
'It/that is (a) doggie.'

c. Så tegne æ mamma (Ina.02 1;10, MLU 1.59)
then draw.PRES I mommie
'Then I draw mommie.'

(37) Early examples of yes/no questions in child Norwegian

a. Er det båt? (Ina.03 1;10, MLU 1.59)
is it boat
'Is that a boat?'

Other studies also report developmental stages consistent with the data reported in this paper. Tsimpli (2005) shows for child Greek that structures involving elements in

[36] In this section, glosses and translations are taken from the literature cited.

peripheral positions are found earliest in the child data (e.g., focus, wh-questions, CLLD, clitic-doubling), whilst tense and agreement features are not acquired yet (including structures such as the unfocused preverbal subject position). Van Kampen (2010) has also explicitly advocated early representation of the discourse domain in Dutch, particularly regarding the V2-system, before the inflectional domain is clearly in place. Schelletter et al. (1997) also discuss the case of a German/English bilingual child. At 2;6, there is no evidence of tense-marking or 3rd-person marking in Sonja's English (besides the copula *be*). However, wh-movement is already attested at 2;3-2;6. A similar picture recurs for her German, which likewise lacks much person marking in verbs until 2;3, despite evidence of verb-raising to a finite position. Wh-movement, like in her English, appears very early, from 2;0 onwards. As far as interactionally-oriented language is concerned, Shirai et al. (2000) and Bosch (accepted) also report very early emergence of Mandarin Sentence-Final particles and illocutionary complementisers in Catalan and Spanish, respectively.

Beside production-based data, recent experimental and behavioural work by Perkins and Lidz (2021), which develops work by Seidl et al. (2003), also suggests that awareness of wh-movement and non-local dependencies emerges remarkably early — 18-month-olds display knowledge of the complementarity that exists between a wh-phrase and a local object of a verb (e.g., **What did the chef burn the pizza?*). This work suggests that the fundamental syntactic capacities for computing nonlocal syntactic dependencies become evident extremely early, far outstripping children's production. It additionally predicts that this emergent knowledge can be used as part of incremental parsing in other tasks, such as lexical acquisition (see Perkins et al., 2021, for some preliminary evidence that this is the case; and Lidz, 2022, for a review).

The relatively late emergence of a Split CP-domain also finds parallels in other corpus studies in the literature. Strikingly, these often converge on the MLU ~ 2.5 threshold for the emergence of cartographic-type material.

These structures have also been noted in Moscati and Rizzi (2021), who show that at least 20 examples across 11 Italian-speaking children display a structure where topicalisation or a vocative phrase co-occurs with wh-movement. All examples reported date from MLU 2.5 or later:

(38) a. Mucca, come stai? Voc > Wh (Diana, MLU 3.89)
cow how are
'Cow, how are you?'

b. Questo dove si mette? Top > Wh (Rosa, MLU 2.5)
this were CL goes
'This, where does it go?'

c. Tu dove sei stata queste due settimane? Top > Wh (Martina, MLU 2.5)
you where are been these two weeks
'You, where have you been in these two weeks?'

d. Lui Babbo Natale, cosa fa? Top > Wh (Raffaello, MLU 3.21)
he Santa Calusa what does
'Him, Santa Claus, what does he do?'

e. Pinocchio, dove vai? Voc > Wh (Rosa, MLU 3.03)
Pinocchio where go
'Pinocchio, where are you going?'

The acquisition of fine-grained prepositional structure (e.g., as in Svenonius, 2006, 2008) has also been argued to proceed from a stage where a coarse-grained prepositional category is available, but cartographic heads encoding fine-grained meaning distinctions (such as Svenonius's AxialPartP) have not yet been acquired. This is the essence of Mitrofanova's (2018) Underspecification of P Hypothesis. Mitrofanova (2018) reports a qualitative difference in both comprehension and production with respect to knowledge of PP structure, which is interpreted in a manner compatible with the 'increasing granularity' and 'structural homology' progression argued for here. The 2-year-old children went through a stage where they omitted locative prepositions at high rates and did not conform to target-like case assignment patterns in utterances with preposition omission. High rates of preposition omission were also found to correlate with impoverished comprehension of locative prepositions. These P-omission utterances are suggested to correspond to underspecified syntactic knowledge, with only a coarse-grained prepositional category being represented in the child grammar (analysed as a phonologically null and semantically underspecificed PlaceP; **Figure 3.2**).

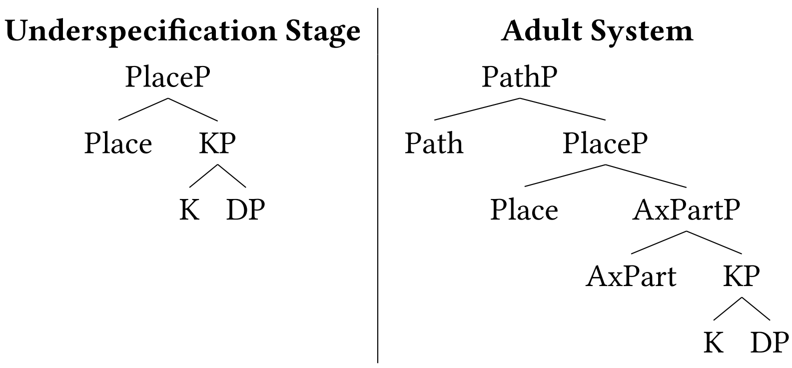


Figure 3.2 Underspecification of P Hypothesis

A recent corpus study on the acquisition of Italian prepositions in Sanfelici and Gallina (2022) also makes a similar point: for MLU-groups between 1.0 and 2.5, only prepositional structures with a single preposition morpheme appear (e.g., *a* 'at/to', *di* 'of', *con* 'with'). It is only once MLU reaches 2.55 – 2.99 (Group 4) and 3.0 – 3.49 (Group 5) that the combination of two prepositional morphemes appeared in the child production data (e.g., *dentro a* 'inside', lit. 'inside at'; *sopra di* 'above', lit. 'above of'). Finally, in the last stage with MLU-range 3.5-5.53, more complex prepositions were produced consisting of three morphemes (e.g., *di sopra di* 'above', lit. 'of above of'). As they emphasise, the length of the lexical item is not the predictor of the order of appearance, as some monomorphemic prepositions emerge later than bimorphemic ones (e.g., *tra* 'between'). As with the data in §3.2, emergence of cartographic structure such as that required for bimorphemic prepositions is 'delayed'.

Considering now the Tense-Aspect-Mood domain, De Lisser et al. (2017) also present developmental data from Jamaican Creole. Jamaican Creole is a particularly good testbed for the acquisition of fine-grained structure in the inflectional and tense layer: the language displays a rich functional hierarchy of Tense, Modal and Aspect (TMA) markers. The production data in De Lisser et al. (2017) is divided into three phases for each child, with a Phase 1 corresponding to recordings of MLU < 2.5, Phase 2 to MLU 2.5 – 3.49 and Phase 3 to MLU > 3.5. In their analysis of the cumulative development of the TMA-zone in Jamaican Creole, they note that combinations of markers in the same utterance are prevalent in the children's production. For example, and paralleling the findings for Italian (both with respect to CPs and PPs) and in §3.2, the examples reported are systematically found at either Phase 2 or Phase 3, even though TMA-markers already appear at Phase 1. These are given below (the IDs of the children are noted in all capitals):

(39) a. Ø jos don bied (COL, 2;00; Phase 2)
Ø RETRO COMP bathe
'I just finished having a bath.'

b. Yo ben a jraiv i van? (KEM, 2;11; Phase 3)
2SG PAST PROG drive DET van
'Were you driving the van?'

c. Ø jos a jum bak (RJU, 2;04; Phase 2)
Ø RETRO PROG come back
'He is just coming back.'

d. Im did kyaahn waak (RJU, 2;08; Phase 3)
3SG PAST ABL-NEG walk
'He couldn't walk.'

e. Ø mos a riid (RJU, 2;08; Phase 3)
Ø NEC PROG read
'She must be reading.'

f. A wuda afi get some jakit (ALA, 2;10; Phase 3)
1SG EPIS OBL get some jacket
'I would have to get some jacket.'

g. Mosi im a-go jraiv di kyar (ALA, 3;00; Phase 3)
EPIS 2SG PROS drive DET car
'Maybe he is going to drive the car.'

So far, however, the timing of emergence of fine-grained information-structural distinctions that characterise a cartographic-type left periphery appears to contradict the results of Westergaard's (2009) micro-cues model. Westergaard (2009) argues for sensitivity to cartographic and information-structural content from the earliest stages of acquisition, in contrast to neo-emergentism. To end this discussion section, I would like to suggest that the findings in Westergaard (2009) are actually compatible with the timeline presented here (at least preliminarily), and that this outcome surfaces once MLU is taken into consideration. **Table 3.17** summarises the micro-cues proposed by Westergaard, which children are assumed to be latching on to[37], and the first reported examples across the three children studied. In the interests of brevity, I restrict discussion to those micro-cues that involve two or more left-peripheral elements/positions and, thus, those that most directly impinge on the contrast proposed for Stages 2 and 3[38]. I include the already-discussed V2 and yes/no question micro-cues for ease of comparison.

Table 3.17 Emergence of Westergaard's (2009) micro-cues (selected)

Micro-cue	Ina	Ann	Ole
V2 in declaratives [$_{DeclP}$ XP [$_{Decl}$ V]]	1;08.20 (file 1) 1.53 MLU	1;08.20 (file 1) 1.4 MLU	1;09.10 (file 1) 1.72 MLU
Y/N questions [$_{PolP}$ [$_{Pol}$ V]]	1;10.04 (file 2) 1.59 MLU	1;08.20 (file 1) 1.4 MLU	1;10 (file 2) 2.24 MLU
Declaratives with clause-initial *kanskje* [$_{DeclP}$ *kanskje* XP… [$_{VP}$ V]]	2;02.12 (file 9) 2.64 MLU	2;06.21 (file 15) 3.55 MLU	2;06.21 (file 14) 3.69 MLU
Subject-initial declaratives with focus-sensitive adverbs [$_{DeclP}$ XP [$_{FocP}$ Foc-Adv [$_{Foc}$ V]]]	2;0.05 (file 5) 1.92 MLU	2;03.09 (file 10) 2.85 MLU	2;02.12 (file 8) 2.93 MLU
V2 with monosyllabic wh-elements [$_{IntP}$ [$_{Int}$ wh [$_{TopP}$ [$_{Top}$ V … XP$_{[+FOC]}$]]	2;01 (file 6) - with V2 2.01 MLU 1;10.4 (file 2) - w/o V2 1.59 MLU	1;11 (file 4) - with V2 1.68 MLU 2;03.09 (file 10) - w/o V2 2.85 MLU	1;09.10 (file 1) - with V2 1.72 MLU (no examples - w/o V2)

[37]Note that the micro-cues themselves are not taken to be innate, but the cartographic spine on which they piggyback is.

[38]I will assume, then, that those micro-cues with a single left-peripheral element (e.g., basic V2, yes-no questions or wh-questions) are compatible with a 'coarser-grained' analysis of them, such as one harnessing Chomsky's Core Functional Categories (as per §3.1.3.1), and are not sufficient evidence to postulate cartographic structure in children's grammars.

Closer examination of the acquisition patterns in Westergaard (2009) suggests that 'micro-cue-level' distinctions start playing a role in production, in most cases, at a strikingly similar MLU to that suggested here for the cartographic-type granularity level of the categorial hierarchy. This is shown by the micro-cue for declaratives with initial *kanskje* (emerging on average at MLU 3.29) and the micro-cue for subject-initial declaratives with focus-sensitive adverbs (average MLU 2.57)[39]. In (36-37), we presented examples of two micro-cues that 'deviate' from the norm and emerge early, namely V2 and yes/no questions (repeated in **Table 3.17**). Yet, as suggested above and in footnote 38, they are precisely cues that can easily be re-formulated in more 'macro' terms; throughout §3.1.3.1-3.2.3.1 we have understood them as providing evidence for a C-head, but not necessarily any cartographic version thereof.

The final micro-cue (word-order with monosyllabic wh-questions) is also an apparent exception, but again, this is compatible with the framework here: monosyllabic wh-questions with focused subjects generate V2-order, which can be interpreted as standard wh-movement to Spec-CP and of the verb to C. As for those with non-V2-order, Westergaard (2009, pp. 210-212) suggests that non-focused subjects in wh-questions raise to Top and cause the verb to remain in non-V2, often sentence-final, order, which she proposes is in T and then transits string-vacuously to InTop (the head of InTopP, shorthand for InnerTopic Phrase, which hosts a higher second subject position, dominating TP). Although this structure can in principle be analysed with a more elaborate left-periphery, both movements (subject to Topic and verb to InTop) are frequently string-vacuous. So, the possibility that the child simply does not move them to those positions at early stages (keeping both within TP, for example) cannot be ruled out from Westergaard's data alone, unless clear movement cues are visible. This could potentially explain their relative earliness. Neither of these kinds of wh-questions, then, would *require* a strictly cartographic analysis (see also footnote 38). Importantly, this does not necessarily mean that children are not *sensitive* to *some* of these focus-related distinctions early on, as Westergaard clearly demonstrates. However, encoding all of them as grammaticalised cartographic heads from the beginning is a very different scenario altogether (recall also the discussion around types of V2-systems in §3.2.3.1). The two require some dissociation.

Therefore, it is only those micro-cues that would force us to assume cartographic-style organisation in children's grammar which generally emerge late, entirely in line with our expectations. Westergaard's results, then, are possibly compatible with the ones here, with the proviso that the present neo-emergentist approach emphasises that there is crucial representation-building taking place in child grammars *before* the reflexes of these

[39] Note that Ina produces them earlier than Ann and Ole at MLU 1.92.

'micro-cues' become apparent in the production data. Additionally, neo-emergentism and Biberauer's (2019) MMM model expects early sensitivity to some information-structural distinctions, but not necessarily to all of them, nor does it predict that these distinctions will have been grammaticalised into cartographic heads from the start. More work is needed on the (non-)overlap between these two kinds of approaches and how the models can shed light on each other.

Taking stock, these conclusions strongly indicate that the findings reported in Stage 3 find antecedents in other acquisition literature, even if initial appearances suggest otherwise. Granted, there is no guarantee that this developmental pathway holds universally. But, given the range of languages surveyed, it is unlikely that some of these patterns are accidental. It remains for further research to determine whether these results generalise to more languages across different layers of the clausal and/or nominal domain and whether/how these timings converge with the development of Wiltschko's (2021) interactional/speech-act domain (see also Speas and Tenny, 2003; Giorgi, 2010; Haegeman and Hill, 2013; Corr, 2016, 2022; Miyagawa, 2022).

3.4 Interim summary

To recapitulate, this chapter has established a number of findings that pose problems for nativist approaches to syntactic categories, of both the maturation and continuity type.

Firstly, the crosslinguistic developmental data points to early acquisition of the CP-domain (**Generalisation 2**). This occurs either during Stage 2 or, possibly, during Stage 1, although the case for the latter remains fragile. Early wh-questions, knowledge of V2, illocutionary complementisers, topicalisations, yes/no questions, among others, have endorsed this point.

Crucially, we also ruled out the possibility of a TP-before-CP stage, on the grounds that productive tense/agreement morphology and syntactic TP-diagnostics are absent at Stage 1. Both emerge simultaneously at Stage 2, a suggestive convergence. Put together, these results favour the conclusion that CP is acquired very early, simultaneously with TP at Stage 2 or potentially even earlier, *contra* the predictions of bottom-up maturation. Furthermore, the unexpected mismatch between the emergence of topicalisation *vs* embedding, particularly evident in the Germanic children, also raises important issues for the Growing Trees Hypothesis (**Generalisation 1**).

The key adjudicating finding has come from the Split CP-diagnostic set, which indicates remarkably late acquisition of structures that require a Split CP, relative to basic CP-structures (**Generalisation 3**). This outcome points to the conclusion that there is no motivation to assume fine-grained structural representations from the start of acquisition, at least as far as this production data is concerned. More critically, however, I showed how the combination of **Generalisations 2** and **3** are mutually antagonistic in maturation and continuity approaches assuming innate categories. Ultimately, I concluded that this conflict favours the abandonment of innate categories and maturation. Heeding this result, I argued for an 'emergent cartography' and for emergent categories more generally, presenting an account in terms of Biberauer and Roberts's (2015) hierarchy.

The generalisability of these findings to other children and languages is an open research question; however, their implications are still significant. What I hope to have shown, as well, is the scientific productivity of probing acquisition through a neo-emergentist lens (which is immaterial to its likely veracity): pressing questions emerge that only surface once the explananda of this kind of approach are taken seriously.

In the following chapter, I take a step back, setting these empirical results aside and moving to more theoretical and mathematical matters. I attempt to provide a biolinguistic argument to support some of the conclusions reached and sketch an interdisciplinary bridge between neo-emergentism and Ehresmann and Vanbremeersch's (2007) model of Evolutive Systems. The effort brings together concerns at the heart of Neo-emergent Syntax, Dynamical Systems and Category Theory, ultimately aspiring to further endorse the theoretical convictions behind neo-emergentist approaches to syntactic development.

4 | *Categorial development in neo-emergent syntax: a biolinguistic rationale and formalisation*

4.1 Modelling grammatical development with 'dynamic' Categories

This final chapter expands the neo-emergentist approach outlined in this thesis by proposing a model of grammar construction which integrates two mathematical frameworks of analysis — Category Theory and Dynamical Systems Theory. The aim is to (partially and preliminarily) formalise the insights in Biberauer and Roberts's (2015) learning path, hereafter B&R. To this end, I adopt Ehresmann and Vanbremeersch's (2007) model of (Memory) Evolutive Systems[1]. Ehresmann and Vanbremeersch (2007) (henceforth E&V) introduce a hybrid systems- and category-theoretic model for natural, open and self-organising systems, such as biological, cognitive or neural systems. The key idea captured in the model is that of a mathematical structure — called a Category — which evolves with time, specifying the configuration of a system at each developmental step. The actual development and change from Category to Category (i.e., to new systemic configurations as time elapses) is modelled via standard category-theoretic tools, notably functors (mappings between Categories).

By adopting these mathematical tools, the ultimate aim is to provide an approach to neo-emergentist grammars which appeals to developmental mechanisms and tools independently harnessed outside linguistics and thus offers a developmental picture which is analogous to that proposed for other complex natural systems.

[1] I will not make use of the 'memory' component of their model, primarily due to space constraints. In what follows therefore, I will refer to these systems as Evolutive Systems, acknowledging the intentional omission of their memory aspects.

The approach advocated here bears notable similarities to the category-theoretic analysis introduced in Song (2019). However, the two approaches are distinct in their analytical focus. The emphasis in Song (2019) is primarily ontological — aiming to relate proposed functional sequences at different levels of granularity. In contrast, the emphasis in this chapter is fundamentally *developmental* — focusing on how to characterise the process of grammar construction explicitly, the developmental changes effected at various points in time and the broader system (both dynamical and representational) as a whole.

This chapter is organised as follows: in **Section 4.1.1**, fundamental definitions are introduced. **Section 4.1.2** begins by characterising the notion of 'Successive Division' as in Dresher (2009) and B&R in terms of the mathematical concept of a co-limit and so-called *patterns*. Notions of hierarchy and granularity levels are defined in **Section 4.1.3**. **Section 4.1.4** proceeds by describing increases in (mental) granularity via so-called *vertical complexification processes*. This section will also discuss the role of the 'intermediate' stages emphasised on in this thesis (e.g., Phasal or CFC) by elaborating on the 'Reductionism Problem'. The majority of this chapter is thus devoted to mathematical analysis of the *representational* stages in B&R. However, **Section 4.1.5** will (re)introduce Dynamical Systems Theory (DST) when evolutive systems and the role of the co-regulators are presented. The resulting approach will be one which marries mathematical and conceptual insights from both DST and Category Theory, over and above the groundwork outlined in Bosch (2022).

4.1.1 Category-theoretic preliminaries

Category Theory, first introduced by Samuel Eilenberg and Saunders Mac Lane in the 1940s (Eilenberg and Mac Lane, 1945), is a general theory of mathematical structures and their relations. With its flexibility and far-reaching applicability, Category Theory takes a bird's eye view of mathematics (and is, in a sense, a theory of "the mathematics of mathematics"). Many major properties of the objects modelled by these mathematical structures (the Categories) are expressed not in terms of the internal properties of the objects, but in terms of the way they interact with each other through links, termed *morphisms*. At a high level of abstraction, the theory aims to uncover common relations between different levels of analysis or across frameworks. The primary concern of Category Theory is, thus, the manner in which an object interacts with other objects through morphisms, instead of the nature of the object itself and its structural properties. It is fundamentally a theory of relations between mathematical structures, with morphisms being privileged over objects.

Here, I introduce the basic notations to be borrowed from Category Theory and associated fields, such as Order Theory and Graph Theory. Basic definitions are adopted from E&V, or occasionally also from Awodey (2006). These are very often not verbatim but adapted to the domain that concerns us (acquisition and representation), although the key insights and operations are kept intact.

4.1.1.1 Graphs, Order Theory and Functional Sequences

One kind of relation between objects introduced in E&V's model and to be discussed here is the notion of *graph*, which is specified below (E&V, p. 23).

Definition 2 (Graph)**.** *A graph G (also called a directed graph) consists of a set of objects, called its vertices (or nodes), which we denote as $|G|$, and a set of edges (or arrows) from a vertex A to a vertex B, denoted by $f : A \longrightarrow B$. We call A the source of the arrow, and B its target.*

To conform with E&V's approach, we will model categorial sequences (c-sequences), such as functional hierarchies, as graphs. Before doing so, however, it is vital to establish what order-theoretic (and thus, graph-theoretic) properties we afford these c-sequences, namely whether we take them to be partially ordered or pre-ordered, or furthermore totally ordered. The traditional view, especially in cartographic work (Rizzi, 1997; Cinque, 1999, *et seq.*), has been that c-sequences are *chains* or total orders (**Definition 3**; partly based on Schröder, 2016, pp. 1-2):

Definition 3 (Chain)**.** *A totally ordered set or chain is an ordered pair $\langle P, \leq \rangle$ consisting of a set P and a binary relation $\leq$ contained in $P \times P$ (the total order on P), satisfying the conditions for a partial order (reflexivity, antisymmetry and transitivity) plus an additional comparability (also called totality) condition, such that $\forall a, b, c \in P$:*

1. $\leq$ *is reflexive; namely $a \leq a$.*
2. $\leq$ *is antisymmetric; namely if $a \leq b$ and $b \leq a$ then $a = b$.*
3. $\leq$ *is transitive; namely, if $a \leq b$ and $b \leq c$, then $a \leq c$.*
4. $\leq$ *is comparable (or also strongly connected); namely, either $a \leq b$ or $b \leq a$.*

Furthermore, however, selection-based binary relations in classical cartography are typically assumed to be *strict* total orders, namely total orders that are *irreflexive* (meaning no category can select itself) and *asymmetric* (meaning the selection between two categories is fixed in direction; Song, 2022).

However, I will follow Larson (2021), Song (2019, 2022) and others in taking (at least some) c-sequences (including cartographic sequences) to not be totally ordered, in contrast to what has often been assumed so far. This additional order-theoretic flexibility, afforded by the fact that a maximally poor UG is agnostic as to how orderings are implemented in the system, is welcome on empirical grounds.

Weakening the cartographic assumption that heads in a spine are totally ordered (and derived via pair-wise selection) allows us to tackle so-called *transitivity failures*, where the expectation that if cartographic head X precedes Y and Y precedes Z, then X precedes Z is not met[2]. It also facilitates accommodating totality failures, where some incomparable categories can belong to the same hierarchy but never co-occur (such as 'flavoured' categories, see Folli and Harley, 2005). An overview of these totality/transitivity failures would take us too far afield; we refer the reader to Nilsen (2003), van Craenenbroeck (2006), Bruening (2019) and Larson (2021) for discussion on transitivity failures and the 'problem of plenitude', and Song (2019, 2022) for totality failures[3].

For present purposes, it suffices to say that we will take functional hierarchies to be more flexible than acknowledged in traditional work. Agreeing with some of the above-mentioned authors, I will assume that all c-sequences are minimally pre-orders (a more general relation that is only reflexive and transitive, see Larson, 2021; Song, 2022). As not all systems instantiate cases of transitivity/totality failures, however, this still leaves open the option that many or most functional hierarchies are stricter kinds of orders (e.g., partial orders, which are effectively antisymmetric pre-orders, or the more familiar linear/total orders; cf. Song's, 2022, Weak Cartographic Hypothesis)[4].

Following E&V, the relevant cognitive units (c-sequences) will be encoded as graphs and so will be represented as the graph-theoretic equivalents of the relevant orders (e.g., Directed Acyclic Graphs, or DAGs, for partial orders). For all intents and purposes, however, the use of graph-theoretic or order-theoretic representations will be treated as equivalent; the use of graphs is simply adopted to facilitate a translation of E&V's system into syntactic systems.

An example of a DAG encoding the *CFC* stage in B&R with incomparable (flavoured) categories is given in (4.1) and a connected directed graph (a chain) with the *Split* level is given in (4.2) (their equivalent ordered sets are noted above the graphs):

[2]An example of this kind is provided in van Craenenbroeck (2006). In Venetian, wh-elements obligatorily precede complementisers (WH > CHE) and complementisers in turn precede clitic-left-dislocated XPs (CHE > CLLD). By transitivity, WH should precede CLLD, when, in fact, this is not the case and the reverse order (CLLD > WH) is allowed.

[3]As noted in Cinque (1999, Chapter 5), the assumption of strict total orders is also problematic for categories such as Negation, which appear to have crosslinguistically variable heights of realisation.

[4]In other words, it is possible we require multiple order-theoretic relations to capture the categorial flexibility attested in human languages. However, it is also conceivable that some kinds of order-theoretic failures can be modelled derivationally (cf. van Craenenbroeck, 2006; Zwart, 2009, on transitivity failures).

$$V \leq \begin{matrix} v_1 \\ v_2 \end{matrix} \leq T \leq C \tag{4.1a}$$

$$\begin{tikzcd} & v_1 \arrow[dr] & & \\ v_{CFC} \arrow[ur] \arrow[dr] & & T_{CFC} \arrow[r] & C_{CFC} \\ & v_2 \arrow[ur] & & \end{tikzcd} \tag{4.1b}$$

$$V \leq Appl \leq v \leq Asp \leq T \leq Mod \leq Fin \leq Foc \leq Top \leq Force \tag{4.2a}$$

$$V \rightarrow Appl \rightarrow v \rightarrow Asp \rightarrow Tns \rightarrow Mod \rightarrow Fin \rightarrow Foc \rightarrow Top \rightarrow Force \tag{4.2b}$$

4.1.1.2 Categories and Functors

The successive configurations of the system and its syntactic representations will be characterised via the mathematical construct of *Categories*[5] and the changes among configurations via *Functors*. The notion of a Category, which develops that of a graph, will become essential in subsequent sections. A Category is defined formally as follows (adapted from Awodey, 2006, pp. 4-5):

Definition 4 (Category)**.** A Category $\mathcal{C}$ comprises:

1. A collection of *objects* A, B, C...;

2. A collection of *morphisms* (or arrows) f, g, h...;
   Each morphism f has a *domain* A (noted $A = dom(f)$) and a *codomain* B (noted $B = cod(f)$).
   The expression $f : A \rightarrow B$ would be verbally stated as "f is a morphism from A to B".

3. Given morphisms $f : A \rightarrow B$ and $g : B \rightarrow C$, that is, with $cod(f) = dom(g)$, there is a given morphism

   $$g \circ f : A \rightarrow C$$

   called the *composite* of f and g.

4. For each object A, there is a morphism

[5] To avoid ambiguity with the linguistic sense of ‘category’, the capitalised ‘Category’ will be used to refer exclusively to the category-theoretic denotation. Similarly, we will use the curly notation $\mathcal{C}$ to denote a Category and to distinguish it from typographically similar notations, such as the syntactic head C.

$$1_A : A \to A$$

called the *identity morphism* of A.

These data are required to satisfy the following laws:

- Associativity law:

$$h \circ (g \circ f) = (h \circ g) \circ f$$

for all $f : A \to B$, $g : B \to C$, $h : C \to D$.

- Unit law:

$$f \circ 1_A = f = 1_B \circ f$$

for all $f : A \to B$.

An example of a Category with four objects and some morphisms between them is depicted below[6]:

$$\begin{tikzcd}
A \arrow[loop left, "1_A"] \arrow[r, "f"] \arrow[rr, bend left, "g \circ f"] & B \arrow[loop below, "1_B"] \arrow[r, "g"] \arrow[rr, bend right, "h \circ g"'] & C \arrow[loop above, "1_C"] \arrow[r, "h"] & D \arrow[loop right, "1_D"]
\end{tikzcd}$$

Thus, a Category is formed by objects and morphisms linking them, as in a graph. The difference between the two lies in that Categories also have an internal rule to compose successive morphisms.

At a higher level of abstraction, one can also consider a Category itself as an object, allowing us to establish relations among Categories in a way that is analogous to relations between objects in a Category. These higher-level relations are called *functors* (adapted from Awodey, 2006, pp. 8-9).

Definition 5 (Functor)**.** A functor F from $\mathcal{C}$ to $\mathcal{D}$ is a mapping of objects to objects and morphisms to morphisms, such that

1. For each object A in $\mathcal{C}$, an object $F(A)$ in $\mathcal{D}$; and

[6] For simplicity, however, and as is customary in Category Theory, I will omit identity morphisms in the formalisations that follow and any morphisms that can be deduced from composition.

2. For each morphism $f : A \rightarrow B$ in $\mathcal{C}$, a morphism $F(f) : F(A) \rightarrow F(B)$ in $\mathcal{D}$.

Additionally,

3. $F(1_A) = 1_{F(A)}$; and
4. $F(g \circ f) = F(g) \circ F(f)$, whenever $g \circ f$ is defined in $\mathcal{C}$.

That is to say, F preserves domains and codomains, identity arrows and compositions. Pictorially, this results in the following commutative diagrams:

$$\mathcal{C} \xrightarrow{F} \mathcal{D}$$

$$\begin{tikzcd}
A \arrow[loop left, "1_A"] \arrow[r, "f"] \arrow[dr, "h"'] & B \arrow[loop right, "1_B"] \arrow[d, "g"] \\
& C \arrow[loop right, "1_C"]
\end{tikzcd}
\qquad
\begin{tikzcd}
F(A) \arrow[loop left, "1_{F(A)}"] \arrow[r, "F(f)"] \arrow[dr, "F(h)"'] & F(B) \arrow[loop right, "1_{F(B)}"] \arrow[d, "F(g)"] \\
& F(C) \arrow[loop right, "1_{F(C)}"]
\end{tikzcd}$$

4.1.1.3 Configuration Categories of a system

During most of this section, we will make use of Categories as a 'static' characterisation of the overall representational system, allowing us to abstract away from the dynamic aspect of time and development.

However, §4.1.5.1 will grant these Categories a more 'dynamic' perspective, by representing the successive configurations of a system via so-called *Configuration Categories*. The internal organisation, or *configuration*, of the acquisitional system (with respect to its representations) at a given time t will be modelled by a Category $\mathcal{C}$, termed the *Configuration Category* at time t. The morphisms of $\mathcal{C}$ represent the interactions at t that characterise the present organisation of the system.

Pictorially, the configuration Categories at the *Phasal* and *CFC* stages in B&R could be represented as below, with the switch from *Ph* to *CFC* being modelled as a functor f.

$$\mathcal{C}_{Ph} \xrightarrow{f} \mathcal{C}_{CFC} \tag{4.3}$$

We return to configuration Categories in more depth in §4.1.5 in the context of evolutive systems. For now, this general and conceptual overview of them suffices.

4.1.2 The Binding Problem and Decompositions: Successive Divisions in acquisition

Thus far, we have proposed to model a representational system by use of Categories; that is, a mathematical structure whose objects represent components of the system and whose morphisms encode the relations between those objects. In complex systems, however, one also distinguishes between components at various levels of complexity; for example, in a hierarchy, as in B&R. Before applying these category-theoretic tools to properly syntactic material, this section refines the components of a Category to recognise that a given object may be 'complex', in the sense of containing an internal organisation of objects which together operate similarly function-wise. This will set the basis to define mathematically the notion of 'developmentally-prior' objects and will provide a link to the Successive Division path in B&R.

4.1.2.1 Patterns and co-limits

First, let us define the concept of pattern introduced in E&V. Informally, a pattern P can be understood as a subpart of the directed graphs that model functional hierarchies (introduced in §4.1.1.1). Slightly more formally, however (Ehresmann and Vanbremeersch, 2019, p. 5):

Definition 6 (Pattern)**.** A *pattern* P in a Category $\mathcal{C}$ defines a sub-graph of $\mathcal{C}$.

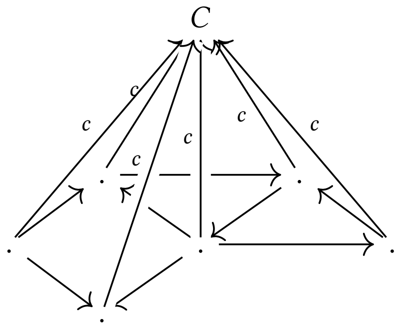
For example, in the graph introduced in (4.1), (4.4) could be a sub-graph of it, as could the independent heads T_{CFC} and C_{CFC}. The latter satisfy the definition of a graph vacuously (that single objects are graphs is vacuously true).

$$T_{CFC} \longrightarrow C_{CFC} \tag{4.4}$$

Second, let us introduce the categorical notion of a co-limit (adapted from E&V, p. 57):

Definition 7 (Co-limit)**.** An object of the Category $\mathcal{C}$ is called the *co-limit* of a pattern P in $\mathcal{C}$, denoted cP, if there exists a morphism c from the pattern P to cP.

Or pictorially (where each morphism to C 'binds' a pattern with the co-limit C; adapted from Brown, 2009, p. 275),



For instance, if $\mathcal{C}$ is a Category modelling a chemical system, a molecule is the co-limit of the pattern formed by its atoms. When looked at from a 'top-down' perspective, a pattern becomes a *decomposition* of a co-limit; the co-limit is seen as a complex object (or hyperstructure, *sensu* Baas, 1992) of which the pattern is an internal, further organisation (that is, the co-limit is in a sense 'divided' into separate patterns). A pattern P can have no co-limit, but if the co-limit of P exists, it is unique (up to an isomorphism, meaning any object isomorphic to cP is a co-limit of P). In contrast, several patterns may have the same co-limit. This last property is important in this context.

Following on from the above, I suggest that the Successive Division path sketched in B&R can be modelled using the notion of decomposition of a co-limit into a pattern. For example, taking phasal C and its successively-divided units C and T (both CFCs), C_{Ph} is the co-limit of the patterns C_{CFC} and T_{CFC}, which are therefore understood as internal organisations of the higher C_{Ph} (which is developmentally-prior in B&R). The individual heads C_{CFC} and T_{CFC} satisfy the definition of a pattern because they satisfy the definition of a graph vacuously (it is vacuously true that a graph with only one vertex is a graph) and therefore can be described as decompositions of a co-limit as done above. Successive Division *à la* B&R can thereby receive an analogous treatment to that for the biological and cognitive systems in E&V. Increases in mental granularity are understood as decomposing co-limits into patterns, thereby treating co-limits as developmentally-prior to their decompositions.

This is sketched in the diagram below for the earlier graph in (4.1).

$$\begin{tikzcd}
 & v_{Ph} \arrow[r] & C_{Ph} & \\
 & v_1 \arrow[u] \arrow[dr] & & \\
v_{CFC} \arrow[uur] \arrow[ur] \arrow[dr] & & T_{CFC} \arrow[uu] \arrow[r] & C_{CFC} \arrow[uul] \\
 & v_2 \arrow[ur] \arrow[uuu, bend right] & &
\end{tikzcd} \tag{4.5}$$

Therefore, all of v_{CFC}, v_1, v_2 are patterns that decompose the co-limit v_{Ph}. Similarly, C_{Ph} is the co-limit of T_{CFC} and C_{CFC} in this example.

Extending this analysis further, the entire hierarchy in B&R can thus be recast in the terms in **Figure 4.1**, with each subordinate level containing syntactic heads (i.e., patterns) that subdivide other functional heads (their co-limits) in hierarchically higher levels (first-pass[7]):

$$\begin{tikzcd}
 & & & & \pm V & & & \\
 & & N_{EP} \arrow[urr] & & & & V_{EP} \arrow[ull] & \\
 & n_{Ph} \arrow[ur] \arrow[rr] & & D_{Ph} \arrow[ul] & & v_{Ph} \arrow[ur] \arrow[rr] & & C_{Ph} \arrow[ul] \\
Num_{CFC} \arrow[r] \arrow[ur] & n_{CFC} \arrow[r] \arrow[u] & Q_{CFC} \arrow[r] \arrow[ur] & D_{CFC} \arrow[u] & Asp_{CFC} \arrow[r] \arrow[ur] & v_{CFC} \arrow[r] \arrow[u] & T_{CFC} \arrow[r] \arrow[ur] & C_{CFC} \arrow[u]
\end{tikzcd}$$

Figure 4.1 Emergent categorial hierarchy as a Category $\mathcal{C}$ (first-pass)

The introduction of the mathematical notions of co-limits and patterns thus provides a possible solution to the Binding Problem, as E&V note:

Definition 8 (The Binding Problem)**.**

a) (Bottom-up). The problem of yielding an object which binds (or 'glues') a pattern P into a single object with the same coherent functional role (E&V, p. 59).

[7]Flavoured categories are omitted for simplicity. Note also that this is (largely) the same diagram as that in Song (2019, p. 231), just arrived at via slightly different methodology.

b) (Top-down). The problem of finding a pattern or patterns which a complex object C 'glues' together and for which the patterns are 'subdivisions' (with C having the same coherent functional role as the patterns that subdivide it).

For language acquisition understood in the neo-emergentist manner advocated here and in B&R, I will primarily view the Binding Problem (and any subsequent problems defined in E&V) in a top-down manner (as in b)), aiming to find a mathematical characterisation of the acquisitional process that divides a 'complex' object C into finer-grained internal structures. This analytical choice is also in line with Bosch (2022), which argues that development is expected to proceed from general-to-specific or coarse-to-fine-grained from a DST perspective (see §2 and B&R's NO>ALL>SOME path). Therefore, the solution as to how to characterise both a) and b) can be understood via the new concept of co-limits: the co-limits entail a 'complex', 'bundled' or 'undivided' and developmentally-prior object which is subsequently separated into associated patterns. In language, then, I suggest a specific instantiation of co-limits may involve the hierarchically coarser-grained syntactic heads in B&R, whose patterns are their successive divisions into ever more elaborate syntactic units.

4.1.2.2 The Multiplicity Principle leads to structurally-homologous systems

A defining characteristic of Complex Adaptive Systems (CASs), briefly mentioned in §2, is not just their hierarchical structure, but also their adaptability and flexibility, which makes them remarkably robust to variation in both internal and external conditions. The concept of *degeneracy* (Edelman, 1989) has been argued to play a central role in ensuring a healthy non-conflicting balance between robustness (ability to maintain the system's configuration and functions against external and internal perturbations) and adaptability (ability to integrate structures from the environment and specialise to it). Degeneracy is the capacity of structurally different elements to perform similar functions or generate the same output. This partial overlap in the functioning of distinct components appears to be a "ubiquitous property of biological systems at all levels of organisation" (Edelman and Gally, 2001, p. 13763). Indeed, degeneracy is central to the Evolutive Systems described in E&V and, as will be argued now, also to the neo-emergentist systems expected on an MMM and DST basis (cf. §2). Degeneracy is juxtaposed to 'redundancy', which occurs when the same function is performed by identical elements.

This distinction between 'degenerate' and 'redundant' elements is captured by E&V with 'homologous' and 'connected' patterns, respectively (adapted from E&V, p. 90):

Definition 9 (Homologous and connected patterns)**.** Two decompositions P and P° of an object C are said to be *connected* if they are structurally and formally isomorphic. If C admits at least two decompositions P and P° which are not connected, we say that C is a *multifaceted* object and P and P° are *homologous* patterns (in equation: $cP = C = cP^{\circ}$).

In biology, homologous patterns emerge pervasively, such as in the Category of particles and atoms. Its objects are the elementary particles and the atoms, with an atom being the co-limit of the pattern created by its nucleus and one of its electronic orbital configurations. A well-known result from quantum physics is that each atom admits different electronic orbital configurations, leading to different energy levels (see E&V, p. 90). The patterns corresponding to these configurations are thus *homologous*, not connected, by virtue of their structural and internal dissimilarities. The atom as a whole is thus a multifaceted object. Likewise, we treat here successive divisions in the emergent categorial hierarchy as leading to *homologous* rather than connected patterns: to the extent that each functional head at every granularity level is distinguished from another via a formal or semantic feature, it follows that these heads at the same level are not connected (i.e., they are not structurally and formally isomorphic). For instance, C_{Ph} is the co-limit of C_{CFC} and T_{CFC}; however T_{CFC} is distinguished from C_{CFC} via a grammaticalised [TENSE] feature. These formal discrepancies mean C_{CFC} and T_{CFC} are not connected in the way described above; they must instead be homologous (or degenerate, in Edelman's sense).

Looked at bottom-up, 'degeneracy' reflects a flexible redundancy of function. Considered top-down, it insists on the fact that C can admit multiple realisations in structurally non-isomorphic lower-level patterns (Ehresmann and Vanbremeersch, 2019, p. 9). E&V describe the Categories which model them as satisfying the principle below, defined from a top-down perspective (from E&V, p. 92):

Definition 10 (Multiplicity Principle)**.** We say that a Category $\mathcal{C}$ satisfies the *multiplicity principle* if it admits at least two patterns which are homologous but not connected.

It is not hard to verify that this principle holds true for the syntactic system in **Figure 4.1**. All of the decompositions at the cartographic level are necessarily homologous, by virtue of being distinguished via semantic features such as [TOPIC], [FOCUS], etc.; and we have also shown that C_{CFC} and T_{CFC} are also structurally-homologous in this sense.

This notion of homologous (or degenerate) patterns and the associated multiplicity principle are important, as they formalise in a category-theoretic manner the conceptual intuition sketched in §2 and in Bosch (2022) that complex systems and, thus, developing grammars present pervasive structural homology. Additionally, the multiplicity principle forms the foundations of the development and evolution of increasingly more complex systems, from physical systems to biological organisms, up to cognitive systems (E&V, p. 90). In these complex systems, many of their signature properties hinge on the fact that a complex component C admits different decompositions which are not ‘materially’ or structurally connected, so that the resulting configuration is not redundant. The present discussion suggests language, understood neo-emergently, is no different.

There is an important caveat to note, nonetheless. The multiplicity principle also draws on an implicit notion of hierarchy (e.g., the hierarchical difference between a co-limit and its decompositions), not directly built into its formulation, but nonetheless crucial for the resulting system. Thus far, however, it is notable that **Figure 4.1** provides no explicit treatment of hierarchical systems such as the one in B&R; rather, if any notion of hierarchical inclusiveness and relations emerges in **Figure 4.1**, it does so merely implicitly. This is rectified in what follows by bringing in the notion of Hierarchical Categories in E&V.

4.1.3 Hierarchy and levels of granularity

After modelling the configuration of a system at a given time using Categories as a tool, and ‘complex’ objects via co-limits and the latters’ internal patterns, we have laid the ground to define the notion of *hierarchical Category* (adapted from E&V, p. 95).

Definition 11 (Hierarchical Category)**.** A *hierarchical Category* is a Category $\mathcal{C}$ the objects of which are partitioned into a finite sequence of (granularity) levels $0, 1, \cdots, N$ (denoted Γ), so that any object C in the level Γ_n is the co-limit in $\mathcal{C}$ of at least one pattern P included in the levels $> \Gamma_n$.

Therefore, **Definition 11** incorporates the observation that the representational system is not ‘unilevel’, but rather organised in terms of granularity levels, with finer-grained ones being developmentally-subsequent to the coarser-grained levels. We can now complete **Figure 4.2** by specifying the granularity levels that each functional sequence corresponds to (second and final pass):

$$\begin{tikzcd}[column sep=tiny]
&&&&&&& \pm V &&&&&&&&& \boxed{\Gamma_{\pm V}} \\
&&& N_{EP} \arrow[urrrr] &&&&&&&& V_{EP} \arrow[ullll] &&&&& \boxed{\Gamma_{EP}} \\
& n_{Ph} \arrow[urr] \arrow[rrrr] &&&& D_{Ph} \arrow[ull] &&&& v_{Ph} \arrow[urr] \arrow[rrrr] &&&& C_{Ph} \arrow[ull] &&& \boxed{\Gamma_{Ph}} \\
Num_{CFC} \arrow[ur] \arrow[rr] && n_{CFC} \arrow[ul] \arrow[rr] && Q_{CFC} \arrow[ur] \arrow[rr] && D_{CFC} \arrow[ul] && Asp_{CFC} \arrow[ur] \arrow[rr] && v_{CFC} \arrow[ul] \arrow[rr] && T_{CFC} \arrow[ur] \arrow[rr] && C_{CFC} \arrow[ul] && \boxed{\Gamma_{CFC}}
\end{tikzcd}$$

Figure 4.2 Emergent categorial hierarchy as a Hierarchical Category $\mathcal{C}$ (final-pass)

When co-limits emerge recursively, such that an object C in granularity level Γ_n is the co-limit of pattern P_1, which is itself the co-limit of another pattern P_2 included in a level lower than Γ_n, we describe C as the *2-iterated co-limit of the ramification* $(P_1, (P_2))$ of length 2. In the hierarchy in **Figure 4.2**, for instance, V_{EP} is the two-iterated co-limit of the patterns C_{CFC} or T_{CFC}. We call these ultimate components of the ramification *micro-components*[8].

Thus, for a given ramification of length 2, one seems to require two stages to generate the micro-components from the highest co-limit C: an initial step to decompose C (say, V_{EP}) into one or multiple patterns (v_{Ph} and C_{Ph}), and a second step subdividing C's decompositions into additional patterns (the *CFC* layer). However, one needs to rule out another logical possibility stemming from a reductionist paradigm, which involves generating the micro-components from the complex object C in a single reconstruction step. *A priori*, neither reconstruction scenario is any less advantageous. As E&V argue, however, consideration of a Hierarchical Category, and its category-theoretic properties, can help give clues as to the validity of a reductionist paradigm.

The Reductionism Problem that arises from Hierarchical Categories possessing high-level ramifications is adapted from E&V below (as per the Binding Problem earlier, a top-down and bottom-up definition is introduced; however, exposition here will still proceed guided by the top-down interpretation):

[8] cf. also B&R on the micro- and nanoparametric levels of their hierarchies. Fittingly, then, if the iterated co-limits are of order 4 or more, we might call them nano-components, analogising from B&R's nanoparameters, with levels 0 and 1 corresponding to no-choice parameters and macroparameters, respectively.

Definition 12 (Reductionism Problem)**.**

a) (Bottom-up). Can a complex object C be reconstructed from atomic components up in only one step, as the co-limit of a (perhaps) large pattern included in Γ level 0? (E&V, p. 103)

b) (Top-down). Can a series of micro-components be reconstructed in only one step as the decompositions of a co-limit in Γ level 0?

To the extent that neo-emergentism requires a restrictive theory which predicts the developmental stages attested (and rules out unattested ones), addressing this Reductionism Problem is crucial: §3 argued that the intermediate layers in B&R's hierarchy are always attested in the production data. One ideally requires a neo-emergentist approach that additionally *predicts* these 'intermediate' or 'interim' developmental stages and not merely *accommodates* them. As Theresa Biberauer (p.c.) notes, one could imagine that the conclusion reached by following a Maximise Minimal Means logic would be that these intermediate stages are a required component of all learning paths. Reaching, for instance, a Split CP stage directly via Phasal C would entail two acquisitional steps being carried out simultaneously: an initial step which recycles phasal C for reuse as a CFC and Split CP category and an additional step postulating separate semantically-specified features to divide C (e.g. [TOPIC], [FOCUS], etc.). This scenario is predicted to be dispreferred by MMM, on the grounds that Feature Economy and Feature/Input Generalisation would initially advocate a recycling step (which satisfies FE and FG/IG). Only if this fails to capture the PLD, would an additional step that postulates new features be resorted to. Thus, postulating cartographic heads right after the Phasal stage is acquisitionally more 'marked' than doing so in separate steps.

We want to capture this idea that proceeding via 'mid-way' stages is in some way more optimal than stages which 'skip' developmental steps. The next section adapts E&V's results and further argues that this conclusion appears to be the right path for linguistic systems also from a category-theoretic and biolinguistic perspective.

4.1.4 Complexification Processes and the Emergence Problem

However, in order to study the earlier problem of reduction, we first require consideration of the problem of 'construction': how can a representational system, such as the guiding example of syntactic hierarchies, be progressively constructed so that increasingly more fine-grained (and thus adult-like) layers emerge? Thus far, consideration has only been given to the entire hierarchy in B&R, as a Category, but without attention to the separate developmental steps implicit in the hierarchy, which successively build its granularity levels. We will model this emergence process with E&V's notion of complexification of a Category, which will set the basis for a more dynamic consideration of evolutive systems. These complexification processes will have clear temporal aspects to them (at least implicitly), but these will be dealt with in §4.1.5.

So-called complexification processes will model the changes in a natural system. Borrowing Thom's (1988) 'standard' changes in biology, the systems in E&V display only four kinds of changes: birth, death, confluence and scission[9]. A change in configuration is modelled by a functor p from $\mathcal{C}$ to the Category $\mathcal{C}'$ modelling the new configurations, which are effected after some of the standard changes above apply to $\mathcal{C}$ (which we will term here *options*, following E&V).

4.1.4.1 Options on a Category and Complexifications

The standard changes in Thom (1988) are translated into categorical operations on $\mathcal{C}$ as follows (from E&V, p. 118):

1. The *birth* and *death* processes depend primarily on the relations of the system with its external environment. Birth is modelled by the addition (called *absorption*) of new elements into $\mathcal{C}'$, and death by the fact that some components of $\mathcal{C}$ are no longer in $\mathcal{C}'$ (*elimination*, by loss or rejection).

2. The *scission* process decomposes some complex objects in $\mathcal{C}$ so that they break up into their separate components.

3. The *confluence* process is modelled by the integration of a pattern in $\mathcal{C}$ into a complex object (possibly emerging in $\mathcal{C}'$), which binds it by becoming its co-limit in $\mathcal{C}'$.

An *option* on $\mathcal{C}$ can therefore be seen as listing a set of objectives modifying $\mathcal{C}$ according to the changes in (1-3), adapted from E&V (p. 118):

Definition 13 (Option)**.** Let $\mathcal{C}$ be a Category. We define an *option* on $\mathcal{C}$ as a list *Op* of objectives for modifying a Category by means of items of all (or some of) the following types, to be used accordingly as described:

1. A graph U of external elements (or, alternatively, recycled internal elements) 'to be (re-)absorbed'.

[9] In cells, for instance, these correspond to endocytosis, exocytosis, synthesis and decomposition of macromolecules, respectively.

2. A set V of objects in $\mathcal{C}$ ‘to be eliminated’.

3. A set of patterns of $\mathcal{C}$ ‘to be bound’: if they already have a co-limit in $\mathcal{C}$, this co-limit has to be preserved; if they have no co-limit, they must acquire one which, depending on the case, might be an object which will emerge, or a specified object of $\mathcal{C}$ which initially is not their co-limit in $\mathcal{C}$.

4. A set of objects of $\mathcal{C}$ ‘to be divided’: if they already have a decomposition in $\mathcal{C}$, this decomposition has to be preserved; if they have no decompositions, they must acquire some via absorption of a graph U, entailing each object C now becomes the co-limit of U (Scission is thus a kind of absorption).

The Category modelling the system after these options are effected will be termed its *complexification* with respect to the option, defined with the following theorem (adapted and simplified from E&V, pp. 128-29):

Theorem 1 (Complexification Theorem)**.** Given a Category $\mathcal{C}$ and an option Op on it, a functor p from $\mathcal{C}$ to a Category $\mathcal{C}'$ can be constructed so as to satisfy the following conditions:

1. The objectives of Op are realised in $\mathcal{C}'$.

2. If there is a complex (multifaceted) object in $\mathcal{C}$ to be divided in $\mathcal{C}'$, it must acquire homologous decompositions in the complexification.

3. If $\mathcal{C}$ satisfies the multiplicity principle, so does $\mathcal{C}'$.

However, and this entails an intentional, slight deviation from E&V, not all standard changes in (1-3) are equally likely in a neo-emergentist system characterised by MMM and structural homology, which attempts to recycle and refine already-existing components to the greatest extent possible and which develops from coarse- to fine-grained. Indeed, the present neo-emergentist work suggests another Theorem is necessary to fully capture the tendencies of the system. This theorem (or even corollary) is termed ‘Death is Rare’ and is defined below. We contend, therefore, that *death* processes are expected to be particularly uncommon.

Theorem 2 (Death is Rare)**.** In structurally-homologous systems which increase in granularity (i.e., develop from ‘general’ to ‘specific’), death of some objects of $\mathcal{C}$ in $\mathcal{C}'$ is rare. The system maximises recycling throughout timescales.

Although the theorem may appear stipulative (given it is not logically derived from any category-theoretic operations introduced), its stipulative flavour fades away from a neo-emergentist perspective. In the DST and Category Theory framework adopted here for syntactic development, Death is Rare is a side-effect of the characteristics of the system. 'Discontinuation' of a category/feature in the system is predicted to be comparatively rare, on the grounds that the preferred *modus operandi* is one where structural homology is prioritised. Indeed, in the derivation of B&R's hierarchy in the next section, no death processes will be used to specify the relevant complexifications.

4.1.4.2 Complexifications in syntactic systems

In this section, as a proof of concept, I apply briefly the options and complexification processes described in E&V to B&R's hierarchy to show this category-theoretic approach can formalise the development of emergent syntactic categories. This will allow us to return to the Reductionism Problem in §4.1.4.3.

First, we construct a Category $\mathcal{C}_0$ which has for its object an undivided categorial space (represented with $\pm V$). Then, we define an option Op on $\mathcal{C}_0$ (the initial or 'starting' Category), which specifies a graph U to be absorbed and which divides $\pm V$ into the objects specified in U (in this case, N_{EP} and V_{EP}). The resulting complexification creates $\mathcal{C}_1$, with the change of configuration in $\mathcal{C}_0$ and $\mathcal{C}_1$ being modelled with a functor p:

$$\mathcal{C}_0 \xrightarrow{p_1} \mathcal{C}_1 \tag{4.6}$$

$$\pm V \qquad\qquad \begin{tikzcd} & \pm V & \\ N_{EP} \arrow[ur] \arrow[rr] & & V_{EP} \arrow[ul] \end{tikzcd}$$

The complexification step for the qualitative switch to the Phasal stage is likewise modelled via a functor $p : \mathcal{C}_1 \rightarrow \mathcal{C}_2$, with an Op specifying that the graphs containing n_{Ph} and v_{Ph} need to subdivide the co-limit N_{EP} and, similarly, that the graphs containing v_{Ph} and C_{Ph} subdivide the co-limit V_{EP} :

$$\mathcal{C}_1 \xrightarrow{p_2} \mathcal{C}_2 \tag{4.7}$$

$$\begin{tikzcd} & \pm V & \\ N_{EP} \arrow[ur] \arrow[rr] & & V_{EP} \arrow[ul] \end{tikzcd} \qquad \begin{tikzcd} & & \pm V & & \\ & N_{EP} \arrow[ur] \arrow[rr] & & V_{EP} \arrow[ul] & \\ n_{Ph} \arrow[ur] \arrow[r] & D_{Ph} \arrow[u] & & v_{Ph} \arrow[u] \arrow[r] & C_{Ph} \arrow[ul] \end{tikzcd}$$

And this emergence pathway carries on cyclically, until the system is deemed complete. The next two diagrams schematise the complexifications for the *CFC* stage and the stage subdividing *C* (focusing just on the verbal domain for simplicity)[10].

$$\mathcal{C}_2 \xrightarrow{\quad p_3 \quad} \mathcal{C}_3 \tag{4.8}$$

$$\begin{tikzcd}[column sep=small]
 & V_{EP} & \\
v_{Ph} \arrow[ur] \arrow[rr] & & C_{Ph} \arrow[ul]
\end{tikzcd}
\qquad
\begin{tikzcd}[column sep=small]
 & & & V_{EP} & & & \\
 & v_{Ph} \arrow[urr] \arrow[rrrr] & & & & C_{Ph} \arrow[ull] & \\
Asp_{CFC} \arrow[ur] \arrow[rr] & & v_{CFC} \arrow[ul] \arrow[rr] & & T_{CFC} \arrow[ur] \arrow[rr] & & C_{CFC} \arrow[ul]
\end{tikzcd}$$

$$\mathcal{C}_3 \xrightarrow{\quad p_4 \quad} \mathcal{C}_4 \tag{4.9}$$

$$C_{CFC}
\qquad
\begin{tikzcd}[column sep=small]
 & & C_{CFC} & & \\
Fin_{Split} \arrow[urr] \arrow[r] & Foc_{Split} \arrow[ur] \arrow[rr] & & Top_{Split} \arrow[ul] \arrow[r] & Force_{Split} \arrow[ull]
\end{tikzcd}$$

4.1.4.3 The Reductionism Problem again

The complexification processes described model changes in cognitive and biological systems and eventually lead to the emergence of finer-grained objects, decompositions of their co-limits. Iterating these complexification processes during development naturally leads to the emergence of hierarchical complex systems. However, is this ‘iteration’ merely illusory or is it really the case that a quantitative increase in the number of objects during successive changes correlates with a qualitative increase in their complexity order? More pointedly, “does the evolution process lead to the creation of new forms, or is everything contained in germ *ab initio* and could *emerge in only one step*?” (E&V, p. 132; emphasis mine, NBM).

This Reductionism Problem faces the question of whether it is theoretically and also developmentally possible to replace a sequence of complexifications by a unique one and get the same end-result, somehow foreseeing the intermediate stages. We argued earlier

[10] As can be seen from the complexification processes described and their respective configuration Categories, it is worth noting that, in both the present approach and in B&R, all of the granularity levels in **Figure 4.1** ultimately conform to the adult system and, therefore, earlier, coarser-grained instantiations of the system are not lost.

that dwelling on this point is rather crucial from a neo-emergentist viewpoint: it is one thing to be able to reach the highest level of granularity in B&R via the intermediate stages specified in the hierarchy, but it is another to do so in one fell swoop. Which scenario is chosen has immediate consequences for the possible developmental stages the approach could predict and so ensuring incorrect stages are ruled out is methodologically important. We saw in §4.1.3 that there is a conceptual motivation for assuming the intermediate stages are not superfluous (by appealing to MMM and the role of features and recycling). The next step is to show that this is also the only possible conclusion in the approach here and that it is paralleled by what is observed in natural systems more generally. More succinctly, I will give support to the claim that a direct path of the sort $\bigcirc \rightarrow \mathit{Split}$ (where $\bigcirc$ denotes an undivided categorial space; Song, 2019) or $\mathit{Ph} \rightarrow \mathit{Split}$ is impossible in the present system and that, e.g., the path $\mathit{Ph} \rightarrow \mathit{CFC} \rightarrow \mathit{Split}$ is the only possible way of reaching *Split* via $\mathit{Ph}/\bigcirc$. The empirical data will dictate whether this rigid theoretical assumption needs to be dropped, but, as seen in §3, the ten children studied speak in favour of it.

Take $\mathcal{C}'$ and $\mathcal{C}''$ to be two successive complexifications of the Category $\mathcal{C}$, and Op and Op' be the options associated with $\mathcal{C}'$ and $\mathcal{C}''$, respectively. The Reductionism Problem boils down to verifying whether we can formulate an option Op'' on $\mathcal{C}$ that leads to a corresponding complexification of $\mathcal{C}$ which is identical (or isomorphic) to $\mathcal{C}''$. Op'' is easy to construct if the changes involved are birth (without scission) or death (the elements to be eliminated or absorbed in Op'' would correspond to the sum of those specified by Op and Op'). The real challenge entails scission: if a series of patterns emerge in $\mathcal{C}''$ as the decompositions of a co-limit in $\mathcal{C}'$, can these patterns also be directly reconstructed as decompositions of a higher co-limit in $\mathcal{C}$, such that they yield isomorphic systems? E&V's approach (pp. 132-133) distinguishes two scenarios for reconstructing these patterns as decompositions of a co-limit in $\mathcal{C}$ (reducing the 2-iterated co-limit to a single one):

1. If a pattern P in $\mathcal{C}''$ is connected (or 'redundant'), P can be constructed as the decomposition of the higher co-limit cP in $\mathcal{C}$, in which case it will be obtained in one step.

2. However, such reduction is not possible if P is homologous. In this case, the scission of cP in $\mathcal{C}$ into P (in $\mathcal{C}''$) requires two successive steps, and there is no option Op'' on $\mathcal{C}$ allowing the emergence of P in a single step.

Again, using cells as a point of comparison, the scenario in (2) occurs in Categories modelling cells, which represent a mirror-image of the Successive Division scenario in syntax (E&V, p. 134). There, complexifications proceed bottom-up, binding patterns (e.g., amino-acids) into co-limits (e.g., compounds or bonds). The synthesis of polypeptide chains (long chains of amino-acids) necessitates three complexifications of the atomic (amino-acid) level. The initial complexification leads to its primary structure (the simplest level of protein structure, which is simply the sequence of amino-acids in a polypeptide chain). The second one creates the basic repeating configuration of the chain (wavy ribbon or helix) and the final one consists of the folding into the final conformation (the tertiary structure). This final conformation cannot be reached straight from amino-acids directly; it requires an intermediate complexification creating the basic configuration of the chain before the third complexification can apply.

This preceding result can be generalised via the multiplicity principle to a longer sequence of complexifications, deriving the following (E&V, pp. 133-134):

Theorem 3 (Iterated Complexification Theorem)**.** If the Category $\mathcal{C}$ satisfies the multiplicity principle, a sequence of complexifications of $\mathcal{C}$ cannot be replaced by a single complexification of $\mathcal{C}$; it leads to the emergence of a hierarchy of objects with strictly increasing complexity, necessitating a construction having more and more steps. This situation is described as an *emergentist reductionism.*

From **Theorem 3**, and since we established many of the patterns with co-limits in B&R are homologous, we deduce that the learning path $\bigcirc/Ph \rightarrow Split$ is indeed impossible to derive with the current theoretical machinery and should be unattested. In other words, we predict an intermediate developmental stage with C_{CFC} and T_{CFC} before cartographic heads emerge. §3 showed this to be borne out, at least for the children studied.

As a corollary of **Theorem 3**, a necessary and sufficient condition for the existence of objects of strictly increasing complexity and granularity levels is thus the multiplicity principle. In a hierarchical category in which there are no multifaceted objects, all the objects are necessarily of complexity order 0 or 1 (leading to pure methodological reductionism; see Ehresmann and Vanbremeersch, 2019, p. 10). A key feature of natural language in the present approach, then, and shared with Biberauer (2019), is its status as a natural system where formal features hierarchically interact to construct and formally distinguish between (homologous) categories (modelled here via Hierarchical Categories), which co-exist at distinct levels of organisation and are successively divided top-down (see also Biberauer, 2019, on the 'more than the sum of the parts' component of Abler's, 1989, Humboldt systems). In other words, a prime characteristic of natural language are *emergent* categories, encoded hierarchically and in the structurally-homologous manner outlined above, as **Theorem 3** anticipates.

We take the theorem above to be essential, allowing us to understand the attestation of the intermediate stages in §3 as *necessary* steps in learning paths. If the category-theoretic analysis so far is on the right track, we conclude these intermediate developmental stages, with evidence of emergent linguistic categories not directly attested in fully fleshed-out adult spines, are not mere 'odd' cases — they are a crucial back-bone of neo-emergentist grammar construction, instantiating necessary 'placeholder'-like stages for the child to reach the ultimate micro-components of the relevant hierarchical Category. Further research should determine, besides the adequacy of this category-theoretic interpretation, the pervasiveness of these interim stages, which we predict here to be abundant.

Accordingly, the preceding results in this section and the empirical data in §3, indicate the following emergence process must be taking place in acquisition, paralleling that observed for complex natural systems and for which we can also define a 'root' (E&V, p. 139):

Theorem 4 (Emergence Process)**.** The root of emergence is the existence in the system of (undivided) multifaceted objects, modelled by the multiplicity principle. If satisfied, the emerging objects and links are explicitly constructed through sequences of complexification processes.

4.1.5 Evolutive Systems: a dynamic conception of Categories

A complex system is an open system, meaning it is liable to influences from the environment, which cause the system to self-organise and develop in time. While the description so far can model the 'static' configuration of the system at a given point in time, one also needs an additional mechanism to deal with (i) the changes between configurations as the system interacts with the environment and (ii) the agents responsible for choosing the options to be effected in the system given the information in the input or PLD. This final section introduces time into the model, an agent only implicitly present thus far.

4.1.5.1 Introducing Evolutive Systems

The concept of an evolutive system fundamentally models the development of a complex system, describing its successive configurations. To this end, we will use a series of Categories indexed by time (which model the configuration of the system at a given time) with functors between them describing their changes. More precisely, an evolutive system is defined in the following manner (a graphic illustration is given in **Figure 4.3** from E&V, p. 154):

Definition 14 (Evolutive System)**.** An *evolutive system* (or ES) $\mathcal{C}$ consists of the following:

- A *time scale* T, which is an interval or a finite subset of the non-negative real numbers.

- For each instant t of T, a Category $\mathcal{C}_t$ called the *configuration Category at t*, which models the configuration of the system in the neighbourhood of t.
- For each instant $t' > t$, a functor $k(t, t')$ from $\mathcal{C}_t$ to $\mathcal{C}_{t'}$, called the transition from t to t'.

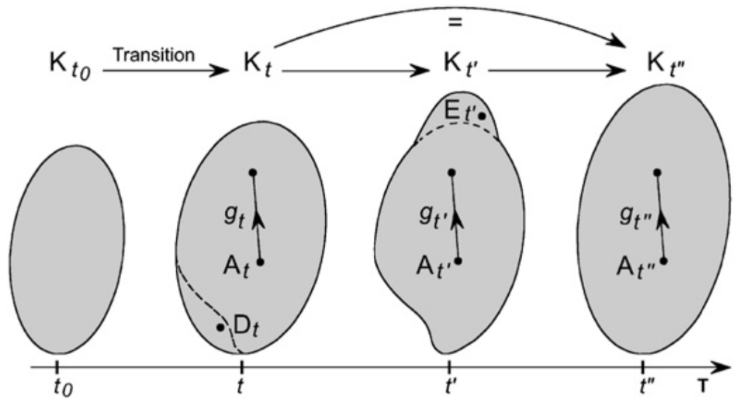


Figure 4.3 An evolutive system: Figure showing several successive configurations of an evolutive system $\mathcal{C}$, beginning from an initial configuration at time t_0 (from E&V, p. 154).

The properties of the evolutive system are inherited from those of their configuration Categories. If its configuration Categories are hierarchical, the ES is called a *hierarchical* ES (or HES). The evolutive system satisfies the multiplicity principle if its configuration Categories satisfy it too.

ESs therefore allow us to understand the complexification processes in §4.1.4.1 as specifying dynamic changes between two configuration Categories, with the functor specifying the actual development of the same system in time. As an illustration, a developmental shift from a system with only EPs to one with CFCs can be sketched diagrammatically as follows:

$$\mathcal{C}_{t_{EP}} \longrightarrow \mathcal{C}_{t_{Ph}} \longrightarrow \mathcal{C}_{t_{CFC}} \tag{4.10}$$

$$V_{EP} \qquad \begin{tikzcd} & V_{EP} & \\ v_{CFC} \arrow[ur] \arrow[r] & C_{CFC} \arrow[u] \end{tikzcd} \qquad \begin{tikzcd} & & V_{EP} & & \\ & v_{Ph} \arrow[ur] \arrow[rr] & & C_{Ph} \arrow[ul] & \\ Asp_{CFC} \arrow[ur] \arrow[r] & v_{CFC} \arrow[u] \arrow[r] & T_{CFC} \arrow[ur] \arrow[r] & C_{CFC} \arrow[u] & \end{tikzcd}$$

4.1.5.2 The role of the co-regulators: a hybrid model with Category Theory and Dynamical Systems Theory

Although the evolution of the system (viewed both 'statically' and dynamically) has been addressed, an open question remains as to how the options specifying the changes to be made in the complexification are chosen, and indeed what (or who) chooses them. Since there is no central regulating mechanism in the HES specified so far, the system should have some regulatory 'organs' – we will term these the *co-regulators*.

A co-regulator, as described in E&V, is simply a sub-system whose components (its agents) act collectively to carry out a specific function[11]. Complex systems are thus self-organised in the sense that their changes are to some extent self-regulated by a network within the broader system, which coordinates development. These co-regulators operate under the system's constraints, whether internal (coming from within the system's structure) or external (from the interaction of the system with environment or, in this context, the PLD). In a neurological system (e.g., nervous systems), an instance of a co-regulator could be a specialised brain area.

In the current context, this 'function' of the co-regulator will be primarily representation and grammar building. At each step, the co-regulator in this acquisitional system performs two main actions (adapted from E&V, p. 186):

i) **Internal observation**: the co-regulator forms a temporary 'model' of the current configuration of the HES, from its point of view (i.e., the co-regulators are modular, not able to access the HES and its internal components directly).

ii) **Regulation**: it selects a set of options to be effected in the complexification in HES, taking into account the results of (i) and the information received from the PLD. More plainly, it selects the new linguistic units to be absorbed into the representational system.

The crucial component of the co-regulators described in E&V is that they result in the formation of a hybrid model of cognitive systems which incorporates the familiar Category Theory alongside Dynamical Systems Theory (a theory thus far left aside in this chapter). There is a division of labour between these two frameworks, which will also be borrowed here: Category Theory (incorporating time) models the organisation of the representational system and its changes over time, while DST primarily studies the local dynamics of the

[11] The 'agentive' nature of the vocabulary used in E&V for co-regulators should not imply that they are endowed with some sort of intentionality. The operations they perform are just a consequence of the system's organisation and its dynamics.

co-regulators (as noted in Bosch, 2022, dynamical systems alone, without any additional frameworks like symbolic dynamics or E&V's evolutive systems, do not lend themselves easily to modelling discrete symbolic systems associated with the acquisitional dynamics).

Relatively little has been said about what these 'co-regulators' ultimately correspond to. I suggest that at least one co-regulator involved in regulating the evolutive system described here for language acquisition is the type of Complex Adaptive System described in Bosch (2022), which incorporates Biberauer's (2019) MMM approach. This entails that the complexifications to be made in successive stages are selected in the manner that MMM and neo-emergentism would predict (I leave open the possibility that there may be more co-regulators relevant to grammar construction than just this one, however). As a consequence, we expect feature/category postulation to abide by the constraints imposed by an MMM and DST system, including (but not limited to) observing linguistic principles such as Feature Economy and Feature Generalisation, structural homology and the Contrastivity requirement (§2.4). More simply, it provides an initial means with which to integrate the category-theoretic discussion so far (and the one in Song, 2019) with the outline of the DST framework in Bosch (2022), such that both systems (representational and dynamic) develop in the manner anticipated by neo-emergentism.

4.1.6 Interim summary: key results

To summarise, this chapter has introduced a novel way to conceptualise Biberauer and Roberts's (2015) hierarchy of emergent categories. We have argued that Ehresmann and Vanbremeersch (2007, 2019) provide a useful framework in which to understand not just B&R's hierarchy and the empirical patterns at stake in this thesis, but also neo-emergentist grammar construction more generally. Although the approach remains preliminary and, for instance, abstracts away from additional kinds of relations such as cross-categorial parallelisms (see Song, 2019), a proof of concept has nonetheless been sketched.

In this context, two theorems have stood out as being particularly fruitful for neo-emergentism. We have established that language is another of the many natural and cognitive systems that satisfies the multiplicity principle of E&V; that is, it is fundamentally a structurally-homologous system (**Definition 10**). This property has been shown to be a catalyst of the emergence process and of complexification steps (**Theorem 1**) and, more broadly, represents a key system-structuring principle in neo-emergent grammars.

Secondly, the conceptual intuition arising from MMM that 'interim' developmental stages in B&R are a key feature of the learning path can now also be understood as another manifestation of more general emergent processes, which proceed in this piece-meal manner as specified by the Iterated Complexification Theorem (**Theorem 3**). Category

Theory and DST thus concur with the empirical paradigm in this thesis, whereby 'mid-way' developmental stages are not simply *predicted*, but are also *sine qua non* conditions for the learning path to proceed successfully. Both the multiplicity principle and the Iterated Complexification Theorem reaffirm the preliminary biolinguistic potential of the hierarchically- and featurally-related categories that appear to typify neo-emergent systems.

The overall system is conceived of as an example of a Hierarchical Evolutive System — a mathematical and complex system describing the successive configurations of the system coordinated by a network of co-regulators (themselves understood as dynamical systems). Here, we have reintroduced the role of DST and the MMM-associated CAS in Bosch (2022), which has been argued to be the main agent specifying the options and complexifications to be realised in the system. The resulting developmental picture is fundamentally in line with the tenets of neo-emergentism, and arguably ratifies the approach further.

Ultimately, the primary feature of this developmental perspective on acquisition is one of *unification*, by harnessing the tools independently needed for other kinds of biological systems and characterising developing grammars in a similar light. This reflects one of the guiding goals of biolinguistics, namely that of finding principled explanations for language based on established properties of natural and biological systems.

5 Concluding remarks

Serratrice (1996, p. 172) recapitulates some of the questions we have been posing, many of which date back to the 1980s, but which have remained to the present day:

> "Although a maturational approach is appealing in that it accounts for successive developmental phases, it doesn't do so in a principled way. Proponents of the Maturation Hypothesis (MH) have argued that invoking maturation is the only way out of the paradox created by the Continuity Hypothesis (CH). If principles of UG are available from the onset of acquisition and the input is responsible for acquisition, why do certain constructions emerge before others? Why doesn't the input trigger these constructions earlier? [...] Explaining *stage transition* becomes increasingly difficult. [...] If the stage transition problem can be dealt with without postulating maturation of UG principles, then there is no need to do so. [...] While in theory any principle can mature, explaining stage transition by resorting to maturation weakens the explanatory power of the theory as it makes a claim which is virtually impossible to falsify."

This thesis has presented a new corpus study that qualifies the requirements on a theory of acquisition addressing Serratrice's (1996) *stage transition* and it has also reconsidered the value of 'syntactic maturation' as a notion with explanatory weight in a Three Factors context.

I have argued, firstly, that, rather than insisting on populating the inventory of Factor 1 with categorial priors and a hard-wired biological program pre-specifying their development, we should abandon both commitments, taking Serratrice's (1996) argument one step further in line with Chomsky (2005) and subsequent work. Doing so in turn fosters deeper, more nuanced explanations. This conceptual argument was then motivated by the empirical outcomes of the corpus study. The results indicate early development of the CP-domain (simultaneously with, or earlier than, TP), but reveal a novel generalisation, worthy of further investigation: production evidence for cartographic-type organisation in the CP-domain systematically emerges at a later stage. This has been interpreted to be at odds with nativist approaches to categories, both Maturation and Continuity, which by design cannot replicate 'shifts' in mental granularity.

The assumption of syntactic maturation and/or innate syntactic categories is what leads to this deadlock. Emergent categories allow the crosslinguistic developmental patterns to be rationalised. We can, with inward-growing (and other) approaches, mantain that the C-domain emerges early, but *also* concede that it is not cartographic in its first instantiation, this being a subsequent development. The findings and explananda that are theoretically most interesting, then, are those corresponding to **Generalisation 3**. 'Maturationalists' cannot incorporate these results. They have to dismiss them, in turn rejecting the possible categorial relations made by emergent syntactic systems. Biberauer and Roberts (2015), therefore, is the only alternative considered that can capitalise on the significance of these patterns. This neo-emergentist endeavor converged with Category Theory and Dynamical Systems Theory (as merged in Ehresmann and Vanbremeersch, 2007, 2019) in the last chapter, which further made the case for softly-assembled development and hierarchical Categories, satisfying the multiplicity principle.

Ultimately, the balance here tips in favour of neo-emergentist approaches to acquisition and variation: the empirical data in the corpus study asks us for flexibility and that is what a maximally minimalist faculty of language grants developing grammars.

References

Abler, W. L. (1989). On the particulate principle of self-diversifying systems. *Journal of Social and Biological Structures*, 12(1):1–13.

Adger, D. (2003). *Core syntax: A minimalist approach*. Oxford University Press, Oxford.

Alexiadou, A. and Anagnostopoulou, E. (1998). Parametrizing AGR: Word order, V-movement and EPP-checking. *Natural Language & Linguistic Theory*, 16(3):491–539.

Antinucci, F. and Miller, R. (1976). How children talk about what happened. *Journal Child Language*, 3:169–189.

Armon-Lotem, S. (1997). *The Minimalist Child: Parameters and Functional Heads in the Acquisition of Hebrew*. PhD thesis, Tel-Aviv University.

Awodey, S. (2006). *Category Theory*. Cambridge University Press, Cambridge.

Baas, N. A. (1992). Hyperstructures — a framework for emergence, hierarchies and complexity. In Bernard, A., Ducret, J.-J., and Grumbach, A., editors, *Proceedings Congrès sur l'émergence dans les modèles de la cognition*, pages 67–93. Télécom, Paris.

Babyonyshev, M. and Marin, S. (2006). Acquisition of Romanian pronominal clitics. *Catalan Journal of Linguistics*, 5:17–44.

Baker, M. (2008). The macroparameter in a microparametric world. In Biberauer, T., editor, *The Limits of Syntactic Variation*, pages 351–373. John Benjamins, Amsterdam.

Bambini, V. and Torregrossa, J. (2010). Cognitive categories behind early Topic/Comment structures. In Chini, M., editor, *Topic, strutture dell'informazione e acquisizione linguistica*, pages 35–58. Franco Angeli, Milan.

Barbosa, P. (1995). *Null subjects*. PhD thesis, Massachusetts Institute of Technology.

Bel, A. (2003). The syntax of subjects in the acquisition of Spanish and Catalan. *Probus: International Journal of Romance Linguistics*, 15(1):1–26.

Biberauer, T. (2010). Semi null-subject languages, expletives and expletive *pro* reconsidered. In Biberauer, T., Holmberg, A., Roberts, I., and Sheehan, M., editors, *Parametric Variation: Null Subjects in Minimalist Theory*, pages 153–99. Cambridge University Press, Cambridge.

Biberauer, T. (2011). In defence of lexico-centric parametric variation: two 3rd factor-constrained case studies. Paper presented at the *Workshop on Formal Grammar and Syntactic Variation: Rethinking Parameters* (Madrid).

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 (Special Issue)*, pages 45–88.

Biberauer, T. (2020). Emergent variation from a minimalist perspective: on the significance of imperatives. Talk presented at *Abralin ao Vivo - Linguists Online* (online), 22 July.

Biberauer, T. and Bosch, N. (2021). The case for Maximise Minimal Means: linguistic and general-cognitive perspectives. Talk presented at *SyntaxLab* (University of Cambridge), 16 November.

Biberauer, T. and Roberts, I. (2015). Rethinking formal hierarchies: A proposed unification. *Cambridge Occasional Papers in Linguistics*, 7:1–31.

Bjorkman, B. A. M. (2011). *BE-ing default: The morphosyntax of auxiliaries*. PhD thesis, Massachusetts Institute of Technology.

Bloom, L., Lifter, K., and Hafitz, J. (1980). Semantics of verbs and the development of verb inflection in language. *Language*, 56(2):386–412.

Bogaards, M., Boogaart, and Barbiers, S. (2022). The syntax of progressive and ingressive *aanhet*-constructions in Dutch. *Linguistics in the Netherlands*, 39:2–20.

Bol, G. W. (1995). Implicational scaling in child language acquisition: the order of production of Dutch verb constructions. In Verrips, M. and Wijnen, F., editors, *Papers from The Dutch-German Colloquium on Language Acquisition, Amsterdam Series in Child Language Development*, page 3. Institute for General Linguistics, Amsterdam.

Borer, H. and Wexler, K. (1987). The maturation of syntax. In Roeper, T. and Williams, E., editors, *Parameter Setting*, pages 123–172. Reidel, Dordrecht.

Bosch, N. (2022). Emergence, Complexity and Developing Grammars: a reinterpretation from a Dynamical Systems perspective. *Cambridge Occasional Papers in Linguistics*, 14(1):1–33.

Bosch, N. (2023). A note on Generalisation 3 (‘Cartography is Emergent’) in Bosch (2023): Alternative explanations for the patterns at Stage 3. Unpublished Ms., University of Cambridge. https://drive.google.com/file/d/1sQcdU6Czns9hnq9vbhWS_IOKJBlJD3Hj/view?usp=sharing.

Bosch, N. (accepted). Not all complementisers are late: a first look at the acquisition of illocutionary complementisers in Catalan and Spanish. Published as a working paper in *Cambridge Occasional Papers in Linguistics*, 15:1-38.

Brown, R. (2009). Memory Evolutive Systems. *Axiomathes*, 19(3):271–280.

Bruening, B. (2019). Disjunctive Selection is Necessary for “Hierarchy of Projections” and it Accounts for Transitivity Failures. Ms., University of Delaware.

Caprin, C. and Guasti, M. T. (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 Martin, R., Michaels, D., and Uriagereka, J., editors, *Step by step: essays on minimalist syntax in honor of Howard Lasnik*, pages 85–156. MIT Press, Cambridge, MA.

Chomsky, N. (2001). Derivation by Phase. In Kenstowicz, M., editor, *Ken Hale: A Life in Language*, pages 1–50. MIT Press, Cambridge, MA.

Chomsky, N. (2004). Beyond explanatory adequacy. In Belletti, A., editor, *Structures and beyond*, pages 104–31. Oxford University Press, Oxford.

Chomsky, N. (2005). Three factors in language design. *Linguistic Inquiry*, 36(1):1–22.

Chomsky, N. (2008). On Phases. In Freidin, R., Otero, C. P., and Zubizarreta, M. L., editors, *Foundational issues in linguistic theory*, pages 133–166. MIT Press, Cambridge, MA.

Cinque, G. (1999). *Adverbs and Functional Heads: A Cross-Linguistic Perspective.* Oxford University Press, Oxford.

Cipriani, P., Pfanner, P., Chilosi, A., Cittadoni, L., Ciuti, A., Maccari, A., Pantano, N., Pfanner, L., Poli, P., Sarno, S., Bottari, P., Cappelli, G., Colombo, C., , and Veneziano, E. (1989). *Protocolli diagnostici e terapeutici nello sviluppo e nella patologia del linguaggio.* Stella Maris Foundation, Italian Ministry of Health.

Clahsen, H. (1986). Verb inflections in German child language: Acquisition of agreement markings and the functions they encode. *Linguistics*, 24:79–121.

Clahsen, H. (1990). Constraints on parameter setting: A grammatical analysis of some acquisition stages in German child language. *Language Acquisition*, 1(4):361–391.

Clahsen, H. (1991). *Child Language and Developmental Dysphasia.* Studies in Speech Pathology and Clinical Linguistics. John Benjamins, Amsterdam.

Clahsen, H., Eisenbeiss, S., and Vainikka, A. (1994). The Seeds of Structure: A Syntactic Analysis of the Acquisition of Case Marking. In Hoekstra, T. and Schwartz, B., editors, *Language Acquisition Studies in Generative Grammar*, pages 85–118. John Benjamins, Amsterdam.

Clahsen, H., Kursawe, C., and Penke, M. (1996). Introducing CP: Wh-Questions and Subordinate Clauses in German Child Language. In Koster, C. and Wijnen, F., editors, *Proceedings of the Groningen Assembly on Language Acquisition*, pages 5–22. Center for Language and Cognition, Groningen.

Clahsen, H. and Penke, M. (1992). The acquisition of agremeent morphology and its syntactic consequences: New evidence on German child language from the Simone-corpus. In Meisel, J., editor, *The Acquisition of Verb Placement: Functional Categories and V2 Phenomena in Language Acquisition*, pages 181–224. Kluwer, Dordrecht.

Clahsen, H., Penke, M., and Parodi, T. (1993). Functional Categories in Early Child German. *Language Acquisition*, 3(4):395–429.

Cormany, E. (2015). Changes in Friulano subject clitics: Conflation and interactions with the left periphery. In Biberauer, T. and Walkden, G., editors, *Syntax over time: Lexical, morphological, and information–structural interactions*, pages 249–264. Oxford University Press, Oxford.

Corr, A. (2016). *Illocutionary complementisers and utterance syntax.* PhD thesis, University of Cambridge.

Corr, A. (2022). *The Grammar of the Utterance: How to Do Things with Ibero-Romance.* Oxford University Press, Oxford.

Cournane, A. and Klævik-Pettersen, E. (accepted). The role of the conservative learner in the rise and fall of V2. *Journal of Historical Syntax.*

Cowper, E. and Hall, D. C. (2014). Reductio ad discrimen: Where features come from. *Nordlyd*, 41(2):145–164.

Cowper, E. and Hall, D. C. (2022). Morphosemantic features in Universal Grammar: What we can learn from Marshallese pronouns and demonstratives. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 67(3):242–266.

Cruschina, S. (2012). *Discourse-Related Features and Functional Projections.* Oxford University Press, Oxford.

Cruschina, S. and Remberger, E.-M. (2018). Speaker-oriented syntax and root clause complementizers. *Linguistic Variation*, 18(2):336–358.

D'Alessandro, R. and Roberts, I. (2010). Past participle agreement in Abruzzese: split auxiliary selection and the null-subject parameter. *Natural Language & Linguistic Theory*, 28(1):41–72.

de Cat, C. (2007). *French Dislocation: Interpretation, Syntax, Acquisition.* Oxford University Press, Oxford.

De Clercq, K., Haegeman, L., Lohndal, T., and Meklenborg, C., editors (2023). *Adverbial Resumption in Verb Second Languages.* Oxford Studies in Comparative Syntax. Oxford University Press, Oxford.

de Haan, G. J. (1987). A theory-bound approach to the acquisition of verb placement in Dutch. In de Haan, G. J. and Zonneveld, W., editors, *Formal Parameters of Generative Grammar III*, pages 15–30. Foris, Dordrecht.

De Lisser, T. N., Durrleman, S., Shlonsky, U., and Rizzi, L. (2017). The Acquisition of Tense, Modal and Aspect markers in Jamaican Creole. *Journal of Child Language Acquisition and Development*, 5(4):219.

de Villiers, J. (1991). Why questions? In Maxfield, T. and Plunkett, B., editors, *Papers in the acquisition of WH*, pages 155–173. UMOP, University of Massachussets, Amherst.

den Besten, H. (1983). On the interaction of root transformations and lexical deletive rules. In Abraham, W., editor, *On the formal syntax of West-Germania*, pages 47–131. John Benjamins, Amsterdam.

Diercks, M. and Bossi, M. (2021). Minimalist syntax is psychologically real: Lessons from (counter)cyclicity. Ms. https://lingbuzz.net/lingbuzz/005671.

Diercks, M., Johns, K., Bar-Server, and Bossi, M. (2023). Developmental Minimalist Syntax. Ms. https://lingbuzz.net/lingbuzz/007134.

Diesing, M. (1992). *Indefinites.* MIT Press, Cambridge, MA.

Don, J., Fenger, P., and Koeneman, O. (2013). Micro-variation as a tool for linguistic analysis. Ms. https://lingbuzz.net/lingbuzz/001954.

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 University Press.

Dresher, B. E. (2014). The arch not the stones: universal feature theory without universal features. *Nordlyd*, 41(2):165–181.

Dumitrescu, D. (2015). Pragmatic and Discursive Aspects of the U.S. Spanish. *Informes del Observatorio*, 15:1–25.

Edelman, G. M. (1989). *The Remembered Present.* Basic Books, New York.

Edelman, G. M. and Gally, J. A. (2001). Degeneracy and complexity in biological systems. *PNAS*, 98(24):13763–13768.

Ehresmann, A. and Vanbremeersch, J.-P. (2007). *Memory Evolutive Systems: Hierarchy, Emergence, Cognition (Studies in Multidisciplinarity).* Elsevier, Amsterdam.

Ehresmann, A. and Vanbremeersch, J.-P. (2019). MES: A Mathematical Model for the Revival of Natural Philosophy. *Philosophies*, 4(1):9.

Eilenberg, S. and Mac Lane, S. (1945). General theory of natural equivalences. *Transactions of the American Mathematical Society*, 58(2):231–294.

Evers, A. and van Kampen, J. (2001). E-language, I-language and the order of parameter setting. *UiL OTS Working Papers*, 00105-S-S.

Ezeizabarrena, M.-J. and Garcia Fernandez, I. (2017). Length of Utterance, in Morphemes or in Words?: MLU3-w, a Reliable Measure of Language Development in Early Basque. *Frontiers in Psychology*, 8:2265.

Fernald, T. (2000). *Predicates and Temporal Arguments.* Oxford University Press, Oxford.

Folli, R. and Harley, H. (2005). Flavors of v. In Kempchinsky, P. and Slabakova, R., editors, *Aspectual Inquiries*, pages 95–120. Springer, Dordrecht.

Frascarelli, M. and Hinterhölzl, R. (2007). Types of topics in German and Italian. In Schwabe, K. and Winkler, S., editors, *On information structure, meaning and form: Generalizations across languages*, pages 87–116. John Benjamins, Amsterdam.

Frey, W. (2004). Notes on the syntax and the pragmatics of German left dislocation. In Lohnstein, H. and Trissler, S., editors, *The syntax and semantics of the left sentence periphery*, pages 203–233. De Gruyter Mouton, Berlin.

Freywald, U., Cornips, L., Ganuza, N., Nistov, I., and Opsahl, T. (2015). Beyond verb second – a matter of novel information-structural effects? Evidence from Norwegian, Swedish, German and Dutch. In Nortier, J. and Svendsen, B. A., editors, *Language, Youth and Identity in the 21st Century: Linguistic Practices across Urban Spaces*, pages 73–92. Cambridge University Press, Cambridge.

Friedmann, N., Belletti, A., and Rizzi, L. (2021). Growing Trees: The acquisition of the left periphery. *Glossa: a journal of general linguistics*, 6(1):131.

Friedmann, N. and Reznick, J. (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*. IUCL Publications, Indiana University.

Gavarró, A., Torrens, V., and Wexler, K. (2010). Object clitic omission: Two language types. *Language Acquisition*, 17(4):192–219.

Gibson, E. and Wexler, K. (1994). Triggers. *Linguistic Inquiry*, 25(3):407–454.

Giorgi, A. (2010). *About the Speaker: Towards a Syntax of Indexicality*. Oxford University Press, Oxford.

Giorgi, A. and Pianesi, F. (1997). *Tense and aspect: From semantics to morphosyntax*. Oxford University Press, Oxford.

Giurgea, I. and Remberger, E.-M. (2016). Illocutionary force. In Ledgeway, A. and Maiden, M., editors, *The Oxford Guide to the Romance Languages*, pages 863–878. Oxford University Press, Oxford.

Grimshaw, J. (1991). Extended projection. Unpublished Ms., Brandeis University.

Grinstead, J. (2000). Case, inflection and subject licensing in child Catalan and Spanish. *Journal Child Language*, 27(1):119–155.

Grinstead, J. (2004). Subjects and Interface Delay in Child Spanish and Catalan. *Language*, 80(1):40–72.

Guasti, M. T. (1993). Verb Syntax in Italian Child Grammar: Finite and Nonfinite Verbs. *Language Acquisition*, 3(1):1–40.

Haegeman, L. and Greco, C. (2016). Frame setters and the macro variation of subject-initial V2. Ms. https://ling.auf.net/lingbuzz/003226.

Haegeman, L. and Greco, C. (2018). West Flemish V3 and the interaction of syntax and discourse. *The Journal of Comparative Germanic Linguistics*, 21(1):1–56.

Haegeman, L. and Greco, C. (2020). Frame setters and microvariation of subject-initial Verb Second. In Woods, R. and Wolfe, S., editors, *Rethinking Verb Second*, pages 61–89. Oxford University Press, Oxford.

Haegeman, L. and Hill, V. (2013). The syntacticization of discourse. In Folli, R., Sevdali, C., and Truswell, R., editors, *Syntax and its Limits*, pages 370–390. Oxford University Press, Oxford.

Hale, K. (1986). Notes on World View and Semantic Categories: Some Warlpiri Examples. In Muysken, P. and van Riemsdijk, H., editors, *Features and projections*, pages 233–256. De Gruyter Mouton, Berlin.

Hall, D. C. (2007). *The Role and Representation of Contrast in Phonological Theory.* PhD thesis, University of Toronto.

Harley, H. and Ritter, E. (2002). Person and number in pronouns: A feature-geometric analysis. *Language*, 78(3):482–526.

Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science*, 298(5598):1569–1579.

Heim, J. and Wiltschko, M. (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 *LAGB Annual Meeting 2021* (online), 8 September.

Hinzen, W. and Wiltschko, M. (2023). Modelling non-specific linguistic variation in cognitive disorders. *Journal of Linguistics*, 59(1):61–87.

Hoekstra, T. and Jordens, P. (1996). From adjunct to head. In Hoekstra, T. and Schwartz, B., editors, *Language Acquisition Studies in Generative Grammar*, pages 119–149. John Benjamins.

Holmberg, A. and Platzack, C. (1995). *The Role of Inflection in Scandinavian Syntax.* Oxford University Press, Oxford.

Hsu, B. (2017). Verb second and its deviations: An argument for feature scattering in the left periphery. *Glossa: a journal of general linguistics*, 2(1):35.

Hsu, B. (2021). Coalescence: A Unification of Bundling Operations in Syntax. *Linguistic Inquiry*, 52(1):39–87.

Hyams, N. (1986). *Language Acquisition and the Theory of Parameters.* Reidel, Dordrecht.

Hyams, N. (1992). Morphosyntactic development in Italian and its relevance to parameter-setting models: Comments on the paper by Pizzuto and Casselli. *Journal of Child Language*, 19(3):695–709.

Hyams, N. (1994). Commentary: Null subjects in child language and the implications of cross-linguistic variation, Volume 2. In B. Lust, M. S. and Whitman, J., editors, *Syntactic theory and first language acquisition: Cross-linguistic perspectives*, pages 287–300. Erlbaum, Hillsdale, NJ.

Hyams, N. (1996). The Underspecification of Functional Categories in Early Grammar. In Clahsen, H., editor, *Generative Perspectives on Language Acquisition: Empirical findings, theoretical considerations and crosslinguistic comparisons*, pages 91–127. John Benjamins, Amsterdam.

Jakobson, R. (1941). *Kindersprache, Aphasie, und allgemeine Lautgesetze*. Uppsala Universitets Årsskrift, Uppsala.

Jakobson, R. and Halle, M. (1956). *Fundamentals of language*. De Gruyter Mouton, The Hague.

Jordens, P. (1990). The acquisition of verb placement in Dutch and German. *Linguistics*, 28(6):1407–1448.

Kayne, R. S. (1991). Romance clitics, verb movement and PRO. *Linguistic Inquiry*, 22(4):647–686.

Kratzer, A. and Selkirk, E. (2007). Phase theory and prosodic spellout: The case of verbs. *The Linguistic Review*, 24(2-3):93–135.

Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, 55(3-4):243–276.

Laka, M. I. (1990). *Negation in syntax: on the nature of functional categories and projections*. PhD thesis, Massachusetts Institute of Technology.

Larson, R. K. (2021). Rethinking cartography. *Language*, 97(2):245–268.

Lebeaux, D. (1988). *Language acquisition and the form of the grammar*. PhD thesis, University of Massachussets.

Lee, N. and Cournane, A. (2019). The journey, not the endstate: finding innovation in the dynamics of L1A. Talk presented at *Diachronic Generative Syntax (DiGS) 2019* (Arizona State University).

Leivada, E. and Murphy, E. (2022). A demonstration of the uncomputability of parametric models of language acquisition and a biologically plausible alternative. *Language Development Research*, 2(1):105–138.

Leivada, E. and Westergaard, M. (2019). Universal linguistic hierarchies are not innately wired. evidence from multiple adjectives. *PeerJ*, 7:e7438.

Lewontin, R. C. (2000). *The triple helix: Gene, organism, and environment*. Harvard University Press, Cambridge, MA.

Lidz, J. (2022). Parser-Grammar Transparency and the Development of Syntactic Dependencies. *Language Acquisition*, pages 1–12.

Llinàs-Grau, M. and Ojea, A. I. (2000). PhonBank Spanish Llinàs/Ojea Corpus. https://doi.org/10.21415/t5k01q.

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk*. Erlbaum, Mahwah, NJ.

Manzini, M. R. (2023). Romance pronominal clitics as pure heads. *Journal of Linguistics*, 59(1):89–119.

Marušič, F., Mišmaš, P., and Zaucer, R. (2019). Looking for Cognitive Foundations of Functional Sequences. *Studies in Polish Linguistics*, 1:53–75. Special Volume.

Meelen, M., Mourigh, K., and Cheng, L. L.-S. (2020). V3 in urban youth varieties of Dutch. In Bárány, A., Biberauer, T., Douglas, J., and Vikner, S., editors, *Syntactic architecture and its consequences II: Between syntax and morphology*, pages 327–355. Language Science Press, Berlin.

Meira, M. and Grolla, E. (2022). The 'Growing Trees' approach and the Development of the Left Periphery of the Sentence in Brazilian Portuguese. Talk presented at *Generative Approaches to Language Acquisition (GALA) 15* (Goethe Universität Frankfurt), September 22-24.

Meklenborg, C. (2020). Adverbial resumptive particles and Verb Second. In Woods, R. and Wolfe, S., editors, *Rethinking Verb Second*, pages 90–125. Oxford University Press, Oxford.

Mielke, J. (2008). *The emergence of distinctive features.* Oxford University Press, Oxford.

Miller, M. (1979). *The Logic of Language Development in Early Childhood.* Springer-Verlag, Berlin.

Mišmaš, P., Marušič, F., and Žaucer, R. (2018). Looking for an extralinguistic source of the strict order of adjectives. Paper presented at the *A Workshop* (Bled, Slovenia).

Mitrofanova, N. (2018). Early Underspecification of Functional Categories: Evidence from the Acquisition of Locative PPs in Russian. *Language Acquisition*, 25(4):341–365.

Miyagawa, S. (2022). *Syntax in the Treetops.* MIT Press, Cambridge, MA.

Mohr, S. (2005). *Clausal architecture and subject positions.* John Benjamins, Amsterdam.

Montes, R. (1987). *Secuencias de clarificación en conversaciones con niños.* Universidad Autónoma de Puebla.

Moscati, V. and Rizzi, L. (2021). The Layered Syntactic Structure of the Complementizer System: Functional Heads and Multiple Movements in the Early Left-Periphery. A Corpus Study on Italian. *Frontiers in Psychology*, 12.

Nilsen, Ø. (2003). *Eliminating Positions: Syntax and semantics of sentence modification.* PhD thesis, Utrecht University.

Ordóñez, F. (1997). *Word order and clause structure in Spanish and other Romance languages.* PhD thesis, CUNY, New York.

Paradis, J. and Genesee, F. (1997). On Continuity and the Emergence of Functional Categories in Bilingual First-Language Acquisition. *Language Acquisition*, 6(2):91–124.

Perkins, L. and Lidz, J. (2021). Eighteen-month-old infants represent nonlocal syntactic dependencies. *PNAS*, 118(41):e2026469118.

Perkins, L., Ying, Y., Williams, A., and Lidz, J. (2021). Object wh-questions with unknown verbs are transitive for 20-month-olds. Talk presented at *Boston University Conference on Language Development 46* (Boston).

Pinker, S. (1984). *Language learnability and language development.* Cambridge University Press, Cambridge.

Platzack, C. (1986). Comp, Infl, and Germanic word order. In Hellan, L. and Koch, K., editors, *Topics in Scandinavian syntax*, pages 185–234. Reidel, Dordrecht.

Platzack, C. (1992). Functional Categories and Early Swedish. In Meisel, J., editor, *The Acquisition of Verb Placement: Functional Categories and V2 Phenomena in Language Acquisition.* Kluwer, Dordrecht.

Poeppel, D. and Wexler, K. (1993). The Full Competence Hypothesis of Clause Structure in Early German. *Language*, 69(1):1–33.

Poletto, C. (2002). The left-periphery of V2-Rhaetoromance dialects: A new view on V2 and V3. In Barbiers, S., Cornips, L., and van der Kleij, S., editors, *Syntactic Microvariation*, pages 214–242. Meertens Institute, Amsterdam.

Pollock, J.-Y. (1989). Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry*, 20(3):365–424.

Pratt, J. W. (1959). Remarks on Zeros and Ties in the Wilcoxon Signed Rank Procedures. *Journal of the American Statistical Association*, 54(287):655–667.

R Core Team (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Radford, A. (1988). Small children's small clauses. *Transactions of the Philological Society*, 86(1):1–43.

Radford, A. (1990). *Syntactic theory and the acquisition of English syntax: The nature of early child grammars of English.* Wiley Blackwell, Oxford.

Rakhlin, N. and Progovac, L. (2021). Hierarchical clause structure as a tool for cognitive advances in early childhood. *Language Sciences*, 83:101316.

Ramchand, G. (2023). The Symbolic Domain. Talk presented at *V-NYI 6 Winter School*, 23 January.

Ramchand, G. and Svenonius, P. (2014). Deriving the functional hierarchy. *Language sciences*, 46:152–174.

Ramchand, G. C. (2008). *Verb Meaning and the Lexicon: A First Phase Syntax.* Cambridge University Press, Cambridge.

Raposo, E. P. and Uriagereka, J. (2008). Clitic Placement in Western Iberian: A Minimalist View. In Cinque, G. and Kayne, R. S., editors, *The Oxford Handbook of Comparative Syntax*, pages 639–697. Oxford University Press, Oxford.

Rice, K. and Avery, P. (1995). Variability in a deterministic model of language acquisition: A theory of segmental elaboration. In Archibald, J., editor, *Phonological acquisition and phonological theory*, pages 23–42. Psychology Press, New York.

Ritter, E. and Wiltschko, M. (2014). The composition of INFL. *Natural Language & Linguistic Theory*, 32(4):1331–1386.

Rizzi, L. (1993). 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 Haegeman, L., editor, *Elements of grammar*, pages 281–337. Kluwer, Dordrecht.

Roberts, I. (2019). *Parameter Hierarchies and Universal Grammar.* Oxford University Press, Oxford.

Roberts, I. (2021). *Diachronic Syntax.* Oxford University Press, Oxford, 2nd edition.

Roberts, I. and Roussou, A. (2003). *Syntactic change: A minimalist approach to grammaticalization.* Cambridge University Press, Cambridge.

Roeper, T. (1992). From the initial state to V2: Acquisition principles in action. In Meisel, J., editor, *The Acquisition of Verb Placement: Functional Categories and V2 Phenomena in Language Acquisition*, pages 333–370. Kluwer, Dordrecht.

Roeper, T. and Rohrbacher, B. (1994). Null Subjects in Early Child English and the Theory of Economy of Projection. *IRCS Technical Reports Series 94-16*, page 162.

Ross, J. R. (1970). On Declarative Sentences. In Jacobs, R. and Rosenbaum, P., editors, *Readings in English Transformational Grammar*, pages 166–183. Blaisdell, Waltham, MA.

Saikkonen, T. (2018). *The Development of Finnish Syntax between the Ages of 1 Year, 8 Months and 3 Years.* PhD thesis, University of Helsinki.

Salustri, M. and Hyams, N. (2006). Looking for the universal core of the RI stage. In Torrens, V. and Escobar, L., editors, *The Acquisition of Syntax in Romance Languages*, pages 159–182. John Benjamins, Amsterdam.

Samuels, B. D., Andersson, S., Sayeed, O., and Vaux, B. (2022). Getting ready for primetime: Paths to acquiring substance-free phonology. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 67(4):552–580.

Sanfelici, E. and Gallina, C. (2022). The timing of production: on the acquisition of Italian prepositions. *Isogloss*, 8(2):1–22.

Santelmann, L. (2003). The acquisition of Swedish wh-questions. In Josefsson, G., Platzack, C., and Håkansson, G., editors, *The Acquisition of Swedish Grammar*, pages 261–307. John Benjamins, Amsterdam.

Schaeffer, J. (2000). *The Acquisition of Direct Object Scrambling and Clitic Placement.* John Benjamins, Amsterdam.

Scheer, T. (2022). 3 x Phonology. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 67(4):444–499.

Schelletter, C., Sinka, I., and Garman, M. (1997). Latvian/English and German/English bilingual acquisition: new light on Universal Grammar. In Cabeza Pereiro, M. d. C., Lorenzo Suárez, A. M., and Rodríguez Yáñez, X. P., editors, *Comunidades e individuos bilingües: actas do I Simposio Internacional sobre o Bilingüismo*, pages 442–452. Universidade de Vigo, Galicia.

Schmerse, D., Lieven, E., and Tomasello, M. (2013). Error patterns in young German children's wh-questions. *Journal of Child Language*, 40(3):656–671.

Schröder, B. (2016). *Ordered sets: An introduction with connections from combinatorics to topology*. Birkhäuser, Cham, 2nd edition.

Schwartz, B. and Vikner, S. (1989). All verb-second clauses are CPs. *Working Papers in Scandinavian Syntax*, 43:27–49.

Schwartz, B. and Vikner, S. (1996). The verb always leaves IP in V2 clauses. In Belletti, A. and Rizzi, L., editors, *Parameters and functionals heads*, pages 11–62. Oxford University Press, Oxford.

Scontras, G., Degen, J., and Goodman, N. D. (2017). Subjectivity Predicts Adjective Ordering Preferences. *Open Mind*, 1(1):53–66.

Seidl, A., Hollich, G., and Jusczyk, P. W. (2003). Early Understanding of Subject and Object Wh-Questions. *Infancy*, 4(3):423–436.

Serra, M. and Solé, R. (1989). CHILDES Catalan Serra/Solé Corpus. https://doi.org/10.21415/t58w3n.

Serratrice, L. (1996). The emergence of functional categories in bilingual first language acquisition. In *Proceedings of the Edinburgh Linguistics Department Conference 96*, pages 170–181.

Serratrice, L. (2000). *Emergence of functional categories in bilingual first language acquisition*. PhD thesis, University of Edinburgh.

Shirai, J., Shirai, H., and Furuta, Y. (2000). Acquisition of Sentence-Final Particles in Japanese. In Perkins, M. and Howard, S., editors, *New Directions in Language Development and Disorders*, pages 243–250. Springer, Dordrecht.

Soares, C. (2003). The C-domain and the acquisition of European Portuguese: The case of wh-questions. *Probus: International Journal of Romance Linguistics*, 15:147–176.

Song, C. (2019). *On the Formal Flexibility of Syntactic Categories*. PhD thesis, University of Cambridge.

Song, C. (2022). Weakening cartography: On the formal foundation of functional hierarchies. Poster presented at the *24th Seoul International Conference on Generative Grammar* (SICOGG24), August 12-14.

Speas, M. (1994). Null Arguments in A Theory of Economy of Projection. *University of Massachusetts Occasional Papers in Linguistics*, 20:179–208.

Speas, P. and Tenny, C. (2003). Configurational properties of point of view roles. In Di Sciullo, A. M., editor, *Asymmetry in Grammar*, pages 315–344. John Benjamins, Amsterdam.

Svenonius, P. (2006). The emergence of axial parts. *Nordlyd*, 33:49–77.

Svenonius, P. (2008). Projections of P. In Asbury, A., Dotlacil, J., Gehrke, G., and Nouwen, R., editors, *Syntax and Semantics of Spatial P*, pages 63–84. John Benjamins, Amsterdam.

Svenonius, P. (2016). Emergent extended projections. Paper presented at MIT, 4 November.

Tedeschi, R. (2009). *Acquisition at the Interfaces: A Case Study on Object Clitics in Early Italian*. PhD thesis, Universiteit Utrecht.

Thom, R. (1988). *Esquisse d'une Sémiophysique*. InterEditions, Paris.

Tsimpli, I. M. (1992). *Functional categories and maturation: The prefunctional stage of language acquisition*. PhD thesis, University College London.

Tsimpli, I. M. (2005). Peripheral positions in early Greek. In Stavrou, M. and Terzi, A., editors, *Advances in Greek Generative Syntax: In honor of Dimitra Theophanopoulou-Kontou*, pages 179–216. John Benjamins, Amsterdam.

Tucker, M. and Hirsh-Pasek, K. (1993). Systems and language: Implications for acquisition. In Smith, L. B. and Thelen, E., editors, *A dynamic systems approach to development: Applications*, pages 359–384. MIT Press, Cambridge, MA.

Uriagereka, J. (1995). Aspects of the Syntax of Clitic Placement in Western Romance. *Linguistic Inquiry*, 26(1):79–123.

Vainikka, A. and Young-Scholten, M. (2011). *The Acquisition of German: Introducing Organic Grammar*. De Gruyter Mouton, Berlin.

van Craenenbroeck, J. (2006). Transitivity failures in the left periphery and foot-driven movement operations. In van de Weijer, J. and Los, B., editors, *Linguistics in the Netherlands 2006*, pages 52–64. John Benjamins, Amsterdam.

van Craenenbroeck, J. and Haegeman, L. (2007). The Derivation of Subject-Initial V2. *Linguistic Inquiry*, 38(1):167–178.

van Kampen, J. (2009). The Non-Biological Evolution of Grammar: Wh-Question Formation in Germanic. *Biolinguistics*, 3(2-3):154–185.

van Kampen, J. (2010). Typological guidance in the acquisition of V2 Dutch. *Lingua*, 120(2):264–283.

Vendler, Z. (1957). Verbs and times. *The Philosophical Review*, 66(2):143–160.

Verrips, M. and Weissenborn, J. (1992). Routes to verb placement in early German and French. The independence of finiteness and agreement. In Meisel, J., editor, *The Acquisition of Verb Placement: Functional Categories and V2 Phenomena in Language Acquisition*, pages 283–331. Kluwer, Dordrecht.

Vikner, S. (1995). *Verb movement and expletive subjects in the Germanic languages.* Oxford University Press, Oxford.

Vikner, S. (2001). Verb movement variation in Germanic and Optimality Theory. Habilitationsschrift, University of Tübingen.

Vikner, S. (2005). Immobile Complex Verbs in Germanic. *Journal of Comparative Germanic Linguistics*, 8:83–115.

Villa-García, J. (2012). Recomplementation and locality of movement in Spanish. *Probus: International Journal of Romance Linguistics*, 24(2):257–314.

Villa-García, J. and González Rodríguez, R. (2020). On the contrasts between *sí* 'yes' and *sí que* 'yes that' in Spanish and the structure of the Complementiser Phrase domain. *Studies in Hispanic and Lusophone Linguistics*, 13(2):451–484.

Villa-García, J. (2014). Spanish subjects can be subjects: Acquisitional and empirical evidence. *IBERIA: An International Journal of Theoretical Linguistics*, 4(2):124–169.

Villa-García, J. and Ott, D. (2022). Recomplementation as a paratactic phenomenon: Evidence from Spanish and English. *Journal of Linguistics*, page 1–32. First View.

Wagner, L. (2001). Aspectual influences on early tense comprehension. *Child Language*, 28(3):661–681.

Walkden, G. (2017). Language contact and V3 in Germanic varieties new and old. *The Journal of Comparative Germanic Linguistics*, 20:49–81.

Westergaard, M. (2009). *The Acquisition of Word Order.* John Benjamins, Amsterdam.

Wexler, K., Gavarró, A., and Torrens, V. (2004). Feature checking and object clitic omission in child Catalan and Spanish. In Bok-Bennema, R., Hollebrandse, B., Kampers-Manhe, B., and Sleeman, P., editors, *Romance Languages and Linguistic Theory 2002*, pages 253–270. John Benjamins, Amsterdam.

Wiltschko, M. (2014). *The Universal Structure of Categories: Towards a Formal Typology.* Cambridge University Press, Cambridge.

Wiltschko, M. (2021). *The Grammar of Interactional Language.* Cambridge University Press, Cambridge.

Wolfe, S. (2015). The nature of Old Spanish verb second reconsidered. *Lingua*, 164:132–155.

Zeijlstra, H. (2004). *Sentential Negation and Negative Concord.* PhD thesis, University of Amsterdam.

Zwart, J.-W. (1993). *Dutch syntax: a Minimalist approach.* PhD thesis, University of Groningen.

Zwart, J.-W. (1997). *Morphosyntax of verb movement: a Minimalist approach to the syntax of Dutch.* Springer, Dordrecht.

Zwart, J.-W. (2009). Uncharted territory? Towards a non-cartographic account of Germanic syntax. In Alexiadou, A., Hankamer, J., McFadden, T., Nuger, J., and Schäfer, F., editors, *Advances in comparative Germanic syntax*, pages 59–83. John Benjamins, Amsterdam.

A Full tables

This Appendix contains the complete set of tables, which record the structures attested in every recording for each child. The criteria for colour shading follow those outlined in §3.1: namely, a given domain is assumed to have emerged (and the cells corresponding to that domain where a structure is attested are shaded in a given colour), once a structure in that domain is first attested and is followed soon after (generally, within the next ∼ 2 files) by other structures diagnosing that same domain. TP-structures are shaded in orange, CP-structures in blue and, finally, cartographic-type structures receive the colour green. This criterion was relaxed in two instances: in the case of Irene, whose recordings are extremely brief at the earliest stages (often having < 50 words by the child), meaning they often contained very few to no verbs and, therefore, none of the structures analysed. It was also weakened in the case of the Split CP-diagnostics, which, by their very nature, are less common than the CP-structures which they combine.

The apparent counterexamples discussed in §3.2.3.3 are noted as '?' in the cells where they are found, to reflect their ambiguous nature. In Germanic, if no wh-questions are attested in a given file, but wh-less questions *are* attested, 'wh-less' is noted in the cell instead of '✓', to signal the absence of wh-words. For completeness, the ages of the children at each recording are also noted next to the MLUs with which they are associated. In future appendices, however, these will be omitted, given the focus on MLU (see §3.1.2).

As noted in §3, structures that appeared as an identical repetition of an utterance by an adult were excluded. We also did not count utterances in which the child was singing or reciting lyrics, as well as unintelligible utterances (either due to pronunciation or transcription). Cases of potential right-dislocations (e.g., instantiating A-bar movement) were also not counted. This is because their position in the structure cannot be reliably disambiguated when the relevant elements are in (close to) sentence-final position. They are, in principle, ambiguous between genuine right-dislocation structures, base-generation ones or, in some instances, after-thoughts. For several children, importantly, recordings are unavailable to help disambiguate between these possibilities. As a result, right-dislocation structures were excluded for all ten children.

A.1 Romance

Catalan

Table A.1 Production of structures by Laura

Age	MLU	S-Neg-V	S-Adv-V	S-Cl-V	Aux	Wh-Q	Top/Foc	Illoc	Embed	Split CP
1;07.20	1.03									
1;09.07	1.09									
1;10.22	1.15							✓		
1;11.12	1.15							✓		
2;02.05	1.35							✓		
2;02.13	1.3					✓				
2;04.11	1.44				✓	✓				
2;05.08	1.64									
2;06.25	1.76				✓	✓				
2;07.20	1.78	✓		✓	✓	✓		✓		
2;08.30	1.88			✓	✓	✓	✓	✓		
2;11.17	1.98	✓	✓		✓	✓	✓	✓		?
3;00.02	2.42	✓		✓	✓	✓	✓	✓	✓	
3;03.21	3.47	✓		✓	✓	✓	✓	✓	✓	✓
3;05.13	2.54	✓		✓	✓	✓	✓	✓	✓	✓
3;10.00	2.97	✓		✓	✓	✓	✓	✓	✓	✓
3;10.01	2.91	✓		✓	✓	✓	✓	✓	✓	✓
3;11.12	3.0	✓		✓	✓	✓	✓	✓	✓	✓
4;00.10	3.18	✓	✓		✓	✓	✓	✓	✓	✓

Table A.2 Production of structures by Gisela

Age	MLU	S-Neg-V	S-Adv-V	S-Cl-V	Aux	Wh-Q	Top/Foc	Illoc	Embed	Split CP
1;07.14	1.04									
1;08.03	1.02									
1;08.24	1.13							✓		
1;09.00	1.16									
1;10.07	1.14									
1;11.11	1.09									
2;01.23	1.53									
2;02.06	1.5									
2;04.25	1.58		✓		✓			✓		
2;06.23	2.32	✓		✓		✓		✓		

2;08.00	2.61	✓		✓	✓	✓	✓	✓	✓	✓
2;09.16	2.68	✓		✓	✓	✓	✓	✓	✓	
2;11.00	2.6	✓			✓		✓	✓	✓	
3;00.29	2.63	✓		✓	✓		✓	✓		
3;05.15	2.66			✓	✓	✓	✓	✓	✓	✓
3;06.28	3.51	✓	✓	✓	✓	✓	✓	✓	✓	
3;10.02	2.95	✓	✓	✓	✓	✓	✓	✓	✓	✓
3;11.14	2.71	✓	✓	✓	✓	✓	✓	✓	✓	✓
4;00.24	2.22	✓		✓	✓		✓			
4;02.03	3.41	✓			✓	✓	✓	✓	✓	✓

Italian

Table A.3 Production of structures by Martina

Age	MLU	S-Neg-V	S-Adv-V	S-Cl-V	Aux	Wh-Q	Top/Foc	Illoc	Embed	Split CP
1;07.18	1.26				✓					
1;08.02	1.9					✓		✓		
1;08.17	1.57					✓				
1;09.01	1.59									
1;10.29	1.66				✓	✓	✓			
1;11.02	1.99				✓	✓		✓		
1;11.20	1.99	✓			✓	✓	✓		✓	
2;01.12	1.86				✓	✓	✓	✓		
2;03.01	2.55	✓				✓		✓	✓	
2;03.22	2.64	✓			✓	✓			✓	
2;04.13	2.69	✓			✓	✓	✓	✓		✓
2;05.21	2.37				✓	✓	✓	✓	✓	✓
2;07.15	2.55	✓		✓	✓	✓	✓		✓	✓

Table A.4 Production of structures by Rosa

Age	MLU	S-Neg-V	S-Adv-V	S-Cl-V	Aux	Wh-Q	Top/Foc	Illoc	Embed	Split CP
1;07.13	1.27					✓				
1;09.11	1.5					✓				
1;10.08	1.44					✓				
1;11.24	1.39				✓	✓				
2;01.14	1.41					✓				
2;01.29	1.5				✓[1]	✓				?
2;02.11	1.54					✓				
2;04.23	1.75					✓	✓			
2;04.29	1.78				✓	✓	✓			
2;05.25	2.14	✓		✓	✓	✓	✓			
2;06.29	2.6				✓	✓			✓	
2;07.00	1.34					✓				
2;07.26	2.78	✓			✓	✓	✓		✓	
2;09.04	2.87	✓		✓	✓	✓	✓	✓	✓	
2;09.24	2.54				✓	✓	✓	✓	✓	
2;10.14	2.5				✓	✓	✓		✓	✓
2;11.12	3.03				✓	✓	✓		✓	
2;11.30	2.6	✓			✓	✓	✓		✓	✓
3;00.24	3.07	✓			✓	✓	✓	✓	✓	✓

3;01.29	2.89	✓			✓	✓	✓		✓	✓
3;03.23	3.24	✓		✓	✓	✓	✓	✓	✓	✓

[1]This file in principle already meets the emergence criterion outlined in §3 and above. However, both auxiliaries at MLU 1.39 and 1.5 instantiate the same utterance *s'è rotto* 'it has broken'. Given that these could be unproductive forms, as noted in §3.2.3.2, I did not count MLU 1.39 or 1.5 as the point of emergence of TP, favouring instead the next recording with TP-structures (MLU 1.78). This is also in accordance with the empirical conclusions reached in §3.2.3.2, which established that these auxiliary constructions are not clear evidence for TP.

Spanish

Table A.5 Production of structures by Irene

Age	MLU	S-Neg-V	S-Adv-V	S-Cl-V	Aux	Wh-Q	Top/Foc	Illoc	Embed	Split CP
0;11.01	1.42									
1;01.25	1.08									
1;01.28	1.05									
1;02.05	1.27									
1;04.16	1.32					✓				
1;04.17	1.0									
1;05.01	1.59									
1;05.15	1.53									
1;05.27	1.43									
1;06.01	1.57									
1;06.16a	1.61	✓								
1;06.16b	2.15									
1;07.05a	1.84									
1;07.05b	1.69					✓				
1;07.05c	1.8									
1;07.22a	1.94	✓								
1;07.22b	1.94	✓								
1;07.22c	3.0									
1;08.09a	1.88							✓		
1;08.09b	2.24						✓			
1;08.26	2.28					✓				
1;09.10	3.28				✓	✓			✓	
1;09.28a	2.53					✓		✓	✓	
1;09.28b	3.09						✓	✓		
1;10.16	2.62				✓	✓			✓	
1;10.29a	2.36					✓				
1;10.29b	2.22					✓				
1;11.01	3.09					✓	✓	✓	✓	
1;11.13	2.95				✓	✓	✓	✓	✓	✓
1;11.30	3.22			✓	✓	✓	✓	✓	✓	✓
2;00.13	3.09			✓	✓	✓		✓	✓	✓
2;00.28	3.11				✓	✓		✓	✓	
2;01.18	4.94	✓		✓	✓	✓	✓	✓	✓	✓
2;01.29	3.24				✓	✓	✓	✓	✓	
2;02.14	3.96			✓		✓	✓	✓	✓	
2;02.29	3.08				✓	✓		✓	✓	
2;03.13	3.44				✓	✓	✓	✓	✓	✓
2;03.28	3.44				✓	✓		✓	✓	

2;04.13a	4.2	✓			✓	✓			✓	
2;04.13b	2.7			✓		✓	✓		✓	
2;04.28	2.99	✓			✓	✓	✓	✓	✓	✓
2;05.13a	3.0				✓	✓		✓	✓	✓
2;05.13b	2.83			✓		✓	✓	✓	✓	
2;05.27	3.99			✓	✓	✓		✓	✓	
2;06.12	5.12	✓		✓		✓		✓	✓	
2;06.26	4.12			✓	✓	✓	✓	✓	✓	✓
2;07.09	4.12	✓			✓	✓	✓	✓	✓	✓
2;07.28	4.41			✓		✓	✓	✓	✓	✓
2;08.14	2.59				✓	✓		✓	✓	✓
2;08.27	3.45				✓	✓	✓	✓	✓	✓
2;09.11	3.65		✓	✓	✓	✓	✓	✓	✓	✓
2;09.26	3.23	✓		✓	✓	✓	✓	✓	✓	✓
2;10.13	4.74				✓	✓	✓	✓	✓	✓
2;10.28	5.13	✓		✓	✓	✓	✓	✓	✓	✓
2;11.11	3.19	✓		✓	✓	✓	✓	✓	✓	✓
2;11.27	4.01		✓	✓	✓	✓	✓	✓	✓	✓
3;00.24	3.52					✓	✓	✓	✓	✓
3;01.22	3.11		✓	✓		✓	✓	✓	✓	✓
3;02.19	3.38	✓		✓	✓	✓	✓	✓	✓	✓

Table A.6 Production of structures by Koki

Age	MLU	S-Neg-V	S-Adv-V	S-Cl-V	Aux	Wh-Q	Top/Foc	Illoc	Embed	Split CP
1;07.20	1.96						✓			
1;09.18	2.54	✓		✓	✓					
1;11.25	2.47	✓			✓	✓	✓		✓	
2;01.29	2.51	✓			✓	✓	✓			
2;02.27	2.47				✓	✓	✓	✓	✓	
2;03.21	2.07				✓	✓	✓	✓		
2;04.18	2.69	✓		✓	✓	✓	✓		✓	✓
2;05.24	3.08	✓	✓	✓	✓	✓	✓		✓	✓
2;06.10	2.71	✓	✓	✓	✓	✓	✓	✓	✓	✓
2;07.10	3.61	✓	✓	✓	✓	✓	✓	✓	✓	✓
2;08.09	2.75	✓	✓	✓		✓	✓	✓	✓	✓
2;09.14	2.93			✓	✓	✓	✓		✓	
2;11.14	3.38	✓	✓	✓	✓	✓	✓	✓	✓	✓

A.2 Germanic

German

Table A.7 Production of structures by Kerstin

Age	MLU	S-Neg-V	S-Adv-V	Aux	V2	Wh-Q	Y/N-Q	Top/Foc	Embed	Split CP
1;03.22	1.09									
1;04.13	1.31									
1;05.03	1.38									
1;05.06	1.38									
1;05.17	1.58					Wh-less				
1;05.24	1.36					✓				
1;06.06	1.45									
1;06.13	1.46									
1;06.20	1.37									
1;07.09	1.58					✓				
1;07.10	1.39					✓				
1;07.24	1.34									
1;08.07	1.2									
1;08.22	1.4									
1;08.26	1.12									
1;10.03	1.28					✓				
1;10.05	1.38				✓	✓				
1;11.20	1.53				✓					
2;00.05	1.76				✓			✓		
2;00.10	1.68				✓		✓	✓	[2]	
2;01.01	1.59			✓	✓	✓		✓		
2;01.02	1.58		✓		✓	✓	✓			
2;02.20	1.72		✓		✓	✓	✓			
2;02.21	1.65	✓	✓		✓	Wh-less				
2;03.01	1.81				✓	✓				
2;03.02	1.86		✓		✓	✓	✓	✓		?
2;04.14	1.98		✓		✓			✓		
2;04.16	1.67		✓	✓	✓	Wh-less	✓	✓		?
2;05.12	1.8	✓	✓	✓	✓	✓	✓	✓		
2;05.14	1.78	✓			✓			✓		?
2;06.02	2.25	✓	✓		✓	✓	✓	✓		
2;06.03	1.62		✓		✓					
2;07.23	2.13	✓	✓	✓	✓	✓	✓	✓	✓	
2;09.11	2.32	✓	✓	✓	✓	✓	✓	✓		✓
2;10.27	2.38		✓	✓	✓	✓	✓	✓	✓	

3;02.08	2.68	✓	✓	✓	✓	✓	✓	✓	✓	✓
3;04.03	2.89	✓	✓	✓	✓	✓	✓	✓	✓	✓

Table A.8 Production of structures by Simone

Age	MLU	S-Neg-V	S-Adv-V	Aux	V2	Wh-Q	Y/N-Q	Top/Foc	Embed	Split CP
1;09.11	1.54					✓				
1;10.20	1.62				✓	✓		✓		
1;10.21	1.69				✓	✓		✓		
1;10.22	1.71				✓	✓		✓		
1;10.27	1.52									
1;10.28	1.94		✓		✓	✓		✓		
1;11.13	1.53		✓		✓	✓				
1;11.14	1.88		✓		✓	✓				
1;11.23	2.21		✓	✓	✓	✓		✓		
2;00.01	2.27		✓	✓	✓	✓				
2;00.03	2.28		✓		✓	✓		✓		
2;00.05	2.31		✓		✓	✓		✓		
2;00.23	2.31	✓	✓		✓	✓	✓	✓		
2;00.26	1.98	✓	✓		✓	✓		✓		
2;01.12	1.93	✓	✓	✓	✓	✓	✓	✓		?
2;01.16	1.63			✓	✓	✓		✓		
2;01.18	1.72			✓	✓	✓		✓		
2;01.19	1.78		✓	✓	✓	✓		✓		
2;01.20	2.03				✓	✓				
2;01.21	1.79	✓			✓	✓		✓		
2;02.03	1.71			✓	✓	✓		✓		
2;02.04	1.94		✓	✓	✓	✓		✓		
2;02.07	1.66	✓	✓	✓	✓	✓		✓		
2;02.18	2.22			✓	✓	✓				
2;02.19	2.0					✓				
2;02.20	2.09	✓	✓	✓	✓	✓	✓	✓		
2;02.21	1.99	✓	✓	✓	✓	✓		✓		
2;04.17	1.82	✓			✓	✓		✓		?
2;04.19	1.89			✓	✓	✓				
2;04.20	1.96			✓	✓	✓		✓	✓	
2;04.21	1.92		✓	✓	✓	✓		✓		

[2]This file contains an isolated conditional clause headed by *ob* 'if', 13 files before other embedding markers appear in Kerstin. The conversational context suggests this is likely a transcription error and the utterance with the conditional clause was made by the parent. Given its ambiguous nature, I will not count this as a instance of embedding by Kerstin.

2;05.13	2.52		✓	✓	✓	✓		✓	✓	
2;05.16	2.35				✓	✓		✓		
2;05.19	2.62	✓		✓	✓	✓		✓	✓	
2;05.22	2.67			✓	✓				✓	
2;06.10	3.35			✓	✓	✓		✓	✓	
2;06.16	4.04			✓	✓	✓		✓	✓	
2;06.23	2.78		✓	✓	✓	✓	✓	✓	✓	✓
2;06.24	2.27			✓	✓					
2;06.26	2.88	✓	✓	✓	✓	✓		✓	✓	
2;06.28	3.43				✓			✓		
2;07.04	4.89	✓	✓	✓	✓	✓	✓	✓	✓	✓
2;07.19	4.0	✓	✓	✓	✓	✓	✓	✓	✓	✓
2;07.23	2.67	✓		✓	✓	✓	✓	✓	✓	✓
2;08.08	2.97		✓		✓	✓	✓	✓	✓	
2;08.09	2.9		✓		✓	✓	✓	✓	✓	
2;08.15	2.5	✓	✓	✓	✓	✓		✓	✓	
2;08.16	2.0				✓		✓			
2;09.10	3.47		✓		✓	✓		✓		
2;09.26	2.85	✓	✓	✓	✓	✓	✓	✓	✓	✓
2;09.28	3.46		✓	✓	✓	✓		✓	✓	

Dutch

Table A.9 Production of structures by Josse

Age	MLU	S-Neg-V	S-Adv-V	Aux	V2	Wh-Q	Y/N-Q	Top/Foc	Embed	Split CP
2;00.07	1.2				✓	✓				
2;00.21	1.44				✓					
2;01.12	1.55				✓					
2;01.26	1.59				✓	✓				
2;02.08	1.69		✓		✓	✓				
2;02.22	1.74				✓	✓				
2;03.28	1.94	✓			✓	✓	✓	✓		
2;04.11	2.14		✓		✓	✓	✓	✓		?
2;04.25	1.9				✓	✓		✓		
2;05.11	1.59				✓		✓	✓		
2;06.01	2.17			✓	✓	✓		✓		
2;06.22	2.11	✓	✓		✓	✓	✓	✓		
2;07.06	2.19		✓	✓	✓	✓	✓	✓		
2;07.20	2.46		✓	✓	✓	✓	✓	✓		?
2;08.04	2.25		✓	✓	✓	✓	✓	✓		
2;08.18	2.55		✓	✓	✓	✓	✓	✓		
2;09.02	2.42	✓	✓	✓	✓	✓	✓	✓	✓	
2;09.16	2.8		✓	✓	✓	✓	✓	✓		
2;11.09	3.57	✓	✓	✓	✓	✓	✓	✓	✓	✓
2;11.23	2.98	✓	✓	✓	✓		✓	✓	✓	
3;00.06	3.06	✓	✓	✓	✓	✓	✓	✓	✓	✓
3;00.20	4.01	✓	✓	✓	✓	✓	✓	✓	✓	
3;01.10	3.91	✓	✓	✓	✓	✓	✓	✓	✓	✓
3;01.24	3.78	✓	✓	✓	✓	✓	✓	✓	✓	✓
3;02.15	3.95	✓	✓	✓	✓	✓	✓	✓	✓	✓
3;02.29	3.19	✓	✓	✓	✓	✓	✓	✓	✓	
3;03.27	3.39	✓	✓	✓	✓	✓	✓	✓	✓	✓
3;04.17	3.2	✓	✓	✓	✓	✓	✓	✓	✓	✓

Table A.10 Production of structures by Sarah

Age	MLU	S-Neg-V	S-Adv-V	Aux	V2	Wh-Q	Y/N-Q	Top/Foc	Embed	Split CP
1;06.16	1.12									
1;07.21	1.17					Wh-less				
1;08.28	1.07									
1;09.10	1.17					Wh-less				

1;10.05	1.09				✓					
1;10.13	1.17				✓					
1;11.01	1.25				✓					
1;11.15	1.37				✓	Wh-less				
2;00.17	1.68	✓			✓		✓	✓		
2;01.10	1.88		✓	✓	✓		✓	✓		
2;02.18	2.11	✓	✓		✓	✓	✓			
2;03.16	2.05		✓	✓	✓	✓	✓			
2;04.02	2.53	✓	✓	✓	✓	✓	✓	✓		
2;04.09	2.34			✓	✓	✓	✓			
2;04.27	2.46	✓	✓	✓	✓	✓	✓	✓		?
2;05.09	2.47	✓	✓	✓	✓	✓	✓	✓		?
2;05.22	2.59		✓	✓	✓	✓	✓	✓		?
2;06.04	2.74		✓	✓	✓	✓	✓	✓		
2;06.11	2.45	✓	✓	✓	✓		✓	✓		
2;06.18	2.8		✓	✓	✓	Wh-less	✓	✓		
2;07.16	2.51	✓	✓	✓	✓	✓	✓	✓		
2;08.06	2.66		✓	✓	✓	✓	✓	✓		
2;08.19	2.97	✓	✓	✓	✓	✓	✓	✓		
2;09.02	2.59	✓	✓	✓	✓	✓	✓	✓		
2;09.07	3.15	✓	✓	✓	✓	✓	✓	✓		
2;10.18	2.88	✓	✓	✓	✓	✓	✓	✓		
2;11.03	2.87	✓	✓	✓	✓	✓	✓	✓		
2;11.27	3.64	✓	✓	✓	✓	✓	✓	✓		
3;00.19	3.52	✓	✓	✓	✓	✓	✓	✓	✓	✓
3;01.17	3.06		✓	✓	✓	✓	✓	✓	✓	✓
3;02.13	3.82	✓	✓	✓		✓	✓	✓	✓	✓
3;03.21	3.05	✓	✓		✓	✓	✓	✓	✓	✓
3;04.13	3.15	✓	✓	✓	✓	✓	✓	✓	✓	✓
3;05.30	2.89	✓	✓	✓	✓	✓	✓	✓	✓	✓
3;07.25	3.24	✓	✓		✓	✓	✓	✓	✓	
3;10.07	3.71	✓	✓	✓	✓	✓	✓	✓	✓	✓
3;11.04	4.07	✓	✓	✓	✓	✓	✓	✓	✓	✓
4;00.11	3.81	✓	✓	✓	✓	✓	✓	✓	✓	✓
4;00.30	4.08	✓	✓	✓	✓	✓	✓	✓	✓	
4;01.11	4.66	✓	✓	✓	✓	✓	✓	✓	✓	✓
4;03.04	5.37	✓	✓	✓	✓	✓	✓	✓	✓	✓
4;04.28	4.28	✓	✓	✓	✓	✓	✓	✓	✓	
4;05.29	4.7	✓	✓	✓	✓	✓	✓	✓	✓	✓
4;06.12	5.06	✓	✓	✓	✓	✓	✓	✓	✓	✓
4;07.25	4.62	✓	✓	✓	✓	✓	✓	✓	✓	✓
4;08.03	5.03	✓	✓	✓	✓	✓	✓	✓	✓	✓
4;09.13	6.07	✓	✓	✓	✓	✓	✓	✓	✓	✓

4;09.29	5.2	✓	✓	✓	✓	✓	✓	✓	✓	✓
4;11.15	4.01	✓	✓	✓	✓	✓	✓	✓	✓	✓
5;02.13	4.92	✓	✓	✓	✓	✓	✓	✓	✓	✓

B Quantitative analyses of the CP domain

Appendix B compiles the results of the secondary analyses on the CP-domain during Stages 1 and 2 (including also the recordings for the stage where no structures are yet attested). This analysis counted the frequency of occurrences of all CP-structures analysed, including V2. For wh-questions, their variety was also analysed, in particular the types of wh-words and the types of verbs used in each wh-question. The quantitative data collected for wh-questions is presented in **Section B.1**. The study on early development of knowledge of V2 in Germanic is reported in **Section B.2**. All CP-structures attested until Stages 1 and 2 are detailed in **Section B.3**, including wh-questions and V2.

For all studies, any identical repetitions of an adult's utterance or of the child's previous utterance were not counted. Instances where the child was singing or reciting lyrics were excluded. Unintelligible utterances, as a result of pronunciation or transcription unclarity, were also discarded.

B.1 Wh-questions

The first supplementary analysis quantified the use of wh-questions during Stages 1 and 2. The tables below report the types and tokens of wh-words used per child and language. Additionally, we also note the proportion of copular wh-questions out of the total wh-questions for each recording. For Germanic languages, wh-less questions are also quantified and reported next to the number of copular wh-questions. The few cases of wh-questions containing non-finite verbal forms are noted via a footnote next to the total count of the relevant recording.

B.1.1 Romance

Catalan

Gloss: on ‘where’, *què* ‘what’, *qui* ‘who’, *perquè* ‘why’, *com* ‘how’.

Table B.1 Production of wh-questions by Laura

MLU	On	Què	Qui	Perquè	Com	Other	Copular	**Total**
1.03								
1.09								
1.15								
1.15								
1.35								
1.3		1					1	1
1.44		1					0	1
1.64								
1.76		1	2				3	3
1.78	1						1	1
1.88	1	1					1	2
1.98		3					1	3
2.42	1	1	1			1	2	4

Table B.2 Production of wh-questions by Gisela

MLU	On	Què	Qui	Perquè	Com	Other	Copular	**Total**
1.04								
1.02								
1.13								
1.16								
1.14								
1.09								
1.53								
1.5								

1.58			
2.32	1	0	1

Italian

Gloss: dove ‘where’, *cosa* ‘what’, *che* ‘what’, *chi* ‘who’, *perché* ‘why’, *come* ‘how’.

Table B.3 Production of wh-questions by Martina

MLU	Dove	Cosa	Che	Chi	Perché	Come	Other	Copular	**Total**
1.26									
1.9		1						1	1
1.57	5							2	5[1]
1.59									
1.66		1						1	1
1.99				1				1	1
1.99	4							4	4
1.86	1	2						3	3
2.55	1		1		1			1	3
2.64	3							3	3

Table B.4 Production of wh-questions by Rosa

MLU	Dove	Cosa	Che	Chi	Perché	Come	Other	Copular	**Total**
1.27			1	1				2	2
1.5				1				1	1
1.44				2				2	2
1.39	7		1	2				10	10
1.41	6							6	6
1.5	2	1						3	3
1.54	2		1	1				4	4
1.75	3	2	10	12				27	27

[1] 2 wh-questions are non-finite.

1.78	3	4	13	13			32	33
2.14	1	3				1	4	5
2.6	14	1	1	2			18	18
1.34	3			2			5	5
2.78	3			1			4	4
2.87	9						9	9
2.54	3			1			4	4

Spanish

Gloss: dónde ‘where’, *qué* ‘what’, *quién* ‘who’, *por qué* ‘why’, *cómo* ‘how’.

Table B.5 Production of wh-questions by Irene

MLU	Dónde	Qué	Quién	Por qué	Cómo	Other	Copular	**Total**
1.42								
1.08								
1.05								
1.27								
1.32			1				1	1
1.0								
1.59								
1.53								
1.43								
1.57								
1.61								
2.15								
1.84								
1.69		1					0	1
1.8								
1.94								
1.94								
3.0								
1.88								
2.24								

2.28					1		0	1
3.28		1					0	1
2.53		1					0	1
3.09								
2.62	3	1	1				2	5
2.36	1						1	1
2.22		1					0	1
3.09	2	2			1		2	6

Table B.6 Production of wh-questions by Koki

MLU	Dónde	Qué	Quién	Por qué	Cómo	Other	Copular	**Total**
1.96								
2.54								
2.47	5	3					7	8
2.51	1						1	1
2.47	1	13					13	14
2.07		8			1		6	9

B.1.2 Germanic

German

Gloss: wo 'where', *was* 'what', *wer* 'who', *warum* 'why', *wie* 'how'.

Table B.7 Production of wh-questions by Kerstin

MLU	Wo	Was	Wer	Warum	Wie	Other	Wh-less	Copular	**Total**
1.09									
1.31									
1.38									
1.38									
1.58							1	1	1
1.36	1	1					0	2	2
1.45									
1.46									
1.37									
1.58	1						0	1	1
1.39	2	1					0	3	3
1.34									
1.2									
1.4									
1.12									
1.28	1						0	1	1
1.38	2						0	2	2
1.53									
1.76									
1.68									
1.59	1						0	1	1
1.58	1						0	1	1
1.72	1						0	1	1
1.65							1	1	1
1.81									
1.86		1					0	0	1
1.98									
1.67							1	1	1

1.8	1						0	1	1
1.78									
2.25	2	2					0	4	4
1.62									
2.13		1					0	0	1[2]

Table B.8 Production of wh-questions by Simone

MLU	Wo	Was	Wer	Warum	Wie	Other	Wh-less	Copular	**Total**
1.54	1						0	1	1
1.62	9						0	9	9
1.69	4						0	4	4
1.71	2						0	2	2
1.52									
1.94	1						0	1	1
1.53	2						0	2	2
1.88	10						0	10	10
2.21	2						0	2	2
2.27	15						0	15	15
2.28	3						0	3	3
2.31		1					1	2	2
2.31	1						0	1	1
1.98	5						0	5	5
1.93	1	1					0	2	2
1.63	6					2	0	6	8
1.72	3						0	3	3
1.78	5						0	5	5
2.03	1						0	1	1
1.79	4	1					0	5	5
1.71	9	1					0	10	10
1.94	14						0	14	14
1.66	1						1	2	2
2.22	2						0	2	2

[2] 1 wh-question is non-finite.

2.0			1				0	1	1
2.09	13	2	3				0	16	18
1.99	9		1				0	8	10
1.82	2						1	3	3
1.89	3						3	6	6
1.96		2					1	2	3
1.92	3						0	3	3
2.52	2		2				0	2	4
2.35	2	1					0	2	3
2.62	1	1					0	1	2
2.67									
3.35		2					0	1	2
4.04		2					0	2	2

Dutch

Gloss: waar ‘where’, *wat* ‘what’, *wie* ‘who’, *waarom* ‘why’, *hoe* ‘how’.

Table B.9 Production of wh-questions by Josse

MLU	Waar	Wat	Wie	Waarom	Hoe	Other	Wh-less	Copular	**Total**
1.2									
1.44									
1.55									
1.59		2					4	0	6
1.69		1					3	0	4
1.74	1						1	1	2
1.94		1					3	3	4
2.14		2					0	0	2[3]
1.9	1						3	4	4
1.59									
2.17	2						4	5	6
2.11	7	1					4	12	12
2.19	1						1	2	2
2.46		1					0	1	1

2.25	4	1					0	3	5
2.55		1					0	0	1
2.42	2	1					1	3	4
2.8	1	6					1	2	8

Table B.10 Production of wh-questions by Sarah

MLU	Waar	Wat	Wie	Hoe	Waarom	Other	Wh-less	Copular	**Total**
1.12									
1.17							1	1	1
1.07									
1.17							1	1	1
1.09									
1.17									
1.25									
1.37							1	1	1
1.68									
1.88									
2.11		1					38	39	39
2.05	2						0	2	2
2.53	2	5					0	1	7
2.34		2					0	2	2
2.46	3	9			1		1	12	14[4]
2.47	5	2					0	6	7
2.59	2						0	2	2
2.74		2					0	2	2
2.45									
2.8							1	0	1
2.51	1	4					0	1	5
2.66	1						0	1	1
2.97		4			1		0	2	5
2.59	5	2	1				1	7	9
3.15	1	7	1				3	9	12

[3] 1 wh-question is non-finite.

2.88	1			0	1	1
2.87	2	7	1	1	9	11
3.64	1	1	1	0	3	3

[4] 1 wh-question is non-finite.

B.2 V2 in Germanic

This quantitative analysis determined the proportion of verbs that fall in V2/V1 position *vs* in V-final position vis-à-vis their morphological specifications ([±FINITE]). Included in the analysis are multi-word utterances that are spontaneous (i.e., not imitated, repeated, sung, or recited). This, then, helps determine not just children's knowledge of the positional distinctions between finite and non-finite verbs, but also their consistency in deploying them correctly in production (see Poeppel and Wexler, 1993, for quantitative data from a later developmental stage).

The study excludes one-word utterances, where the verb's position cannot be established, repetitions of both an adult utterance and the child's previous utterance, and any utterances which are not interpretable because of transcript or pronunciation unclarity. The study also excludes any two-word or three-word utterances from which the verb's position cannot be ascertained, e.g., subject-initial intransitives, such as *die Puppe weint* 'the doll cries', where the verb is finite but is both in second position and sentence-final. Only those multiword utterances that provided sufficient structural cues as to the location of the verb were included (e.g., those containing adverbs, negation or objects, among others).

Lastly, it is worth remarking that Simone speaks a southern German dialect (this becomes relevant again in **Appendix C**). In her variety, importantly, the verb ending *-e* is ambiguous between 1SG and the infinitive (cf. the Standard German infinitival marker *-en*). Following Verrips and Weissenborn (1992), I did not exclude verbs ending in *-e* from the study, insofar as their exclusion could be dismissing important material regarding V2 competence. I assumed that, if the verb was placed in a clearly raised position (V2/V1-position), the form was likely a 1SG one. Conversely, if the verb was in V-final position, I inferred that the form was feasibly non-finite. Note that an inevitable limitation of this syncretism is that it is impossible to ascertain errors of placement with respect to verb finiteness when it comes this verb ending.

Due to time constraints and since the overall developmental trend is apparent from the first recordings, only the first ten files after the start of Stage 1 were analysed for each child.

German

Table B.11 Finiteness and verb position in Kerstin

	V2/V1		V-final	
MLU	Finite	Non-finite	Finite	Non-finite
1.28	6	0	0	0
1.38	5	0	0	2
1.53	11	0	0	2
1.76	20	0	2	2
1.68	11	0	1	3
1.59	13	1	0	8
1.58	13	0	0	3
1.72	15	0	1	7
1.65	18	0	0	9
1.81	17	0	0	4

$\chi^2(1) = 148.301$, $p < 0.001$

Table B.12 Finiteness and verb position in Simone

	V2/V1		V-final	
MLU	Finite	Non-finite	Finite	Non-finite
1.54	8	0	0	0
1.62	49	2	0	52
1.69	26	0	0	13
1.71	25	0	1	26
1.52	0	0	0	0
1.94	5	0	0	10
1.53	15	0	0	26
1.88	16	0	0	18
2.21	21	0	0	14
2.27	47	1	3	36

$\chi^2(1) = 386.437$, $p < 0.001$

Dutch

Table B.13 Finiteness and verb position in Josse

	V2/V1		V-final	
MLU	Finite	Non-finite	Finite	Non-finite
1.2	3	0	0	9
1.44	5	1	0	3
1.55	14	2	0	9
1.59	12	0	0	11
1.69	10	0	0	8
1.74	3	1	0	12
1.94	30	0	0	9
2.14	22	0	0	27
1.9	21	0	0	11
1.59	11	1	0	2

$\chi^2(1) = 217.519$, $p < 0.001$

Table B.14 Finiteness and verb position in Sarah

	V2/V1		V-final	
MLU	Finite	Non-finite	Finite	Non-finite
1.16	6	1	0	0
1.09	5	0	0	2
1.17	4	0	0	1
1.25	7	0	0	7
1.37	6	0	0	12
1.68	17	0	0	12
1.88	26	2	0	16
2.11	60	0	0	12
2.05	46	2	0	14
2.53	90	0	0	29

$\chi^2(1) = 353.248$, $p < 0.001$

B.3 All CP-structures until Stage 1 and 2

A summary of all of the CP-structures attested during Stages 1 and 2 is given below. This includes the structures already reported above (wh-questions and V2) and the ones yet to be reported (topicalisations/focalisations, yes/no questions, illocutionary complementisers, finite embedding markers). I compress them in a single table per language for ease of exposition and reference.

The same provisos that applied earlier regarding utterances excluded also hold here.

B.3.1 Romance

Catalan

Table B.15 CP-structures produced by Laura until Stages 1 and 2

	Wh-Q		**Top/Foc**	**Illoc**	**Embed**
MLU	Copular	Total			
1.03					
1.09					
1.15				2	
1.15				1	
1.35				1	
1.3	1	1			
1.44	0	1			
1.64					
1.76	3	3			
1.78	1	1		3	
1.88	1	2	2	7	
1.98	1	3	1	6	
2.42	2	4	1	22	4

Table B.16 CP-structures produced by Gisela until Stage 2

	Wh-Q		**Top/Foc**	**Illoc**	**Embed**
MLU	Copular	Total			
1.04					
1.02					
1.13				1	
1.16					
1.14					
1.09					
1.53					
1.5					
1.58				2	
2.32	0	1		4	

Italian

Table B.17 CP-structures produced by Martina until Stages 1 and 2

	Wh-Q		**Top/Foc**	**Illoc**	**Embed**
MLU	Copular	Total			
1.26					
1.9	1	1		1	
1.57	2	5			
1.59					
1.66	1	1	1		
1.99	1	1		2	
1.99	4	4	1		2
1.86	3	3	1	1	
2.55	1	3		3	4
2.64	3	3			2

Table B.18 CP-structures produced by Rosa until Stages 1 and 2

	Wh-Q		**Top/Foc**	**Illoc**	**Embed**
MLU	Copular	Total			
1.27	2	2			
1.5	1	1			
1.44	2	2			
1.39	10	10			
1.41	6	6			
1.5	3	3			
1.54	4	4			
1.75	27	27	3		
1.78	32	33	2		
2.14	4	5	2		
2.6	18	18			1
1.34	5	5			
2.78	4	4	2		2
2.87	9	9	2	1	4
2.54	4	4	1	2	1

Spanish

Table B.19 CP-structures produced by Irene until Stages 1 and 2

	Wh-Q		**Top/Foc**	**Illoc**	**Embed**
MLU	Copular	Total			
1.42					
1.08					
1.05					
1.27					
1.32	1	1			
1.0					
1.59					
1.53					
1.43					

1.57					
1.61					
2.15					
1.84					
1.69	0	1			
1.8					
1.94					
1.94					
3.0					
1.88				1	
2.24			1		
2.28	0	1			
3.28	0	1			1
2.53	0	1		3	1
3.09			1	4	
2.62	2	5			1
2.36	1	1			
2.22	0	1			
3.09	2	5	1	2	1

Table B.20 CP-structures produced by Koki at Stage 2

	Wh-Q		**Top/Foc**	**Illoc**	**Embed**
MLU	Copular	Total			
1.96			1		
2.54					
2.47	7	8	2		1
2.51	1	1	1		
2.47	13	14	2	1	3
2.07	6	9	1	1	

B.3.2 Germanic

German

Table B.21 CP-structures produced by Kerstin until Stages 1 and 2

	V2	Wh-Q			Y/N-Q	Top/Foc	Embed
MLU		Wh-less	Copular	Total			
1.09							
1.31							
1.38							
1.38							
1.58		1	1	1			
1.36		0	2	2			
1.45							
1.46							
1.37							
1.58		0	1	1			
1.39		0	3	3			
1.34							
1.2							
1.4							
1.12							
1.28		0	1	1			
1.38	✓	0	2	2			
1.53	✓						
1.76	✓					4	
1.68	✓				1	2	
1.59	✓	0	1	1		1	
1.58	✓	0	1	1	2		
1.72	✓	0	1	1	1		
1.65	✓	1	1	1			
1.81	✓	0	1	1			
1.86	✓	0	0	1	1	2	
1.98	✓					5	
1.67	✓	1	1	1	2	2	
1.8	✓	0	1	1	4	2	

1.78	✓					3	
2.25	✓	0	4	4	2	2	
1.62	✓						
2.13	✓	0	0	1	8	4	1

Table B.22 CP-structures produced by Simone until Stages 1 and 2

	V2	**Wh-Q**			**Y/N-Q**	**Top/Foc**	**Embed**
MLU		Wh-less	Copular	Total			
1.54		0	1	1			
1.62	✓	0	9	9		1	
1.69	✓	0	4	4		1	
1.71	✓	0	2	2		1	
1.52							
1.94	✓	0	1	1		1	
1.53	✓	0	2	2			
1.88	✓	0	10	10			
2.21	✓	0	2	2		1	
2.27	✓	0	15	15			
2.28	✓	0	3	3		1	
2.31	✓	1	2	2		1	
2.31	✓	0	1	1	1	1	
1.98	✓	0	5	5		1	
1.93	✓	0	2	2	1	3	
1.63	✓	0	6	8		4	
1.72	✓	0	3	3		6	
1.78	✓	0	5	5		6	
2.03	✓	0	1	1			
1.79	✓	0	5	5		2	
1.71	✓	0	10	10		3	
1.94	✓	0	14	14		6	
1.66	✓	1	2	2		2	
2.22	✓	0	2	2			
2.0		0	1	1			

2.09	✓	0	16	18	1	11	
1.99	✓	0	8	10		9	
1.82	✓	1	3	3		3	
1.89	✓	3	6	6			
1.96	✓	1	2	3		7	2
1.92	✓	0	3	3		7	
2.52	✓	0	2	3		10	7
2.35	✓	0	2	4		7	
2.62	✓	0	1	2		6	4
2.67	✓						4
3.35	✓	0	1	2		1	3
4.04	✓	0	2	2		3	4

Dutch

Table B.23 CP-structures produced by Josse until Stages 1 and 2

	V2	**Wh-Q**			**Y/N-Q**	**Top/Foc**	**Embed**
MLU		Wh-less	Copular	Total			
1.2	✓	0	0	1			
1.44	✓						
1.55	✓						
1.59	✓	4	0	6			
1.69	✓	3	0	4			
1.74	✓	1	1	2			
1.94	✓	3	3	4	1	1	
2.14	✓	0	0	2	2	1	
1.9	✓	3	4	4		6	
1.59	✓				1	2	
2.17	✓	4	5	6		5	
2.11	✓	4	12	12	2	5	
2.19	✓	1	2	2	6	1	
2.46	✓	0	1	1	8	8	
2.25	✓	0	3	5	3	9	

2.55	✓	0	0	1	4	10	
2.42	✓	1	3	4	9	10	1
2.8	✓	1	2	8	1	10	

Table B.24 CP-structures produced by Sarah until Stages 1 and 2

MLU	**V2**	**Wh-Q**			**Y/N-Q**	**Top/Foc**	**Embed**
		Wh-less	Copular	Total			
1.12							
1.17		1	1	1			
1.07							
1.17		1	1	1			
1.09	✓						
1.17	✓						
1.25	✓						
1.37	✓	1	1	1			
1.68	✓				1	2	
1.88	✓				1	1	
2.11	✓	38	39	39	2		
2.05	✓	0	2	2	2		
2.53	✓	0	1	7	10	6	
2.34	✓	0	2	2	3		
2.46	✓	1	12	14	5	4	
2.47	✓	0	6	7	9	3	
2.59	✓	0	2	2	8	9	
2.74	✓	0	2	2	1	7	
2.45	✓				4	4	
2.8	✓	1	0	1	3	5	
2.51	✓	0	1	5	5	5	
2.66	✓	0	1	1	10	10	
2.97	✓	0	2	5	17	8	
2.59	✓	1	7	9	6	10	
3.15	✓	3	9	12	6	16	
2.88	✓	0	1	1	4	7	

2.87	✓	1	9	11	3	10
3.64	✓	0	3	3	4	9

C Morphological analysis

The results of the morphological analysis are presented in this third and final Appendix. This study quantified the frequency of (i) person-number combinations as part of agreement morphemes, (ii) tense, aspectual and mood morphology (including auxiliaries) and (iii) imperatives produced in each file. For all children, the period studied spanned every recording until the end of Stage 2 (including, therefore, the entirety of Stages 1 and 2, and also the initial period, in some children, where none of the structures analysed are attested). This is the same developmental period that the studies in **Appendix B** examined.

The complete results of this study are reported in **Section C.2**. Additionally, a less detailed overview of the development of agreement morphology in the ten children is offered in **Section C.1**, for ease of reference. As noted in §3.1.5, a morpheme was considered productive once it has been used with at least three different verbs (excluding copulas and the Catalan/Spanish impersonal verbs *haver-hi/haber* 'there is/are', which only conjugate with 3rd person; following, partly, the criteria in Caprin and Guasti, 2009). Because of this, we also recorded the verbs with which children produced each of the person-number morphemes in every recording. In the interests of legibility, however, only the first five verbs in which each morpheme was produced are noted in the tables in **Section C.2**. Since the focus of this study lies in *inflected* forms, where auxiliaries are used, the infinitive of the inflected auxiliary is listed, instead of the infinitive of the participial or accompanying non-finite form[1]. These are distinguished from lexical verbs via a subscript, e.g., Italian *avere*$_{aux}$ 'have'. When these auxiliaries have a homophonous lexical verb in the language (as is the case for Italian *avere* 'to have' or *essere* 'to be'), they are counted as one for the purposes of the productivity metric if, and only if, the inflectional paradigm is identical for both the lexical verb and the auxiliary. They count as separate when the inflectional paradigms diverge: this is only the case for the Catalan go-past auxiliary *anar*, which has a

[1]This assumption is necessary in view of the productivity metric outlined above. Counting the participle or the non-finite verb as the infinitival form of an auxiliary or periphrastic construction would artificially and significantly increase the number of 'inflected' verbs that children would seem to use, when in fact these reduce to inflected variations of a limited set of auxiliaries (combined with a wider range of participial or other non-finite forms).

distinct paradigm for some person-number combinations compared to the lexical verb *anar* 'to go'.

Additionally, when negative imperatives in Ibero-Romance (Catalan and Spanish) are used with the expected subjunctive morphology, this is noted in parentheses under the total count of imperatives in the relevant columns in **Section C.2**, to reflect the proportion of negative imperatives (with subjunctive morphology) out of the total count of imperatives. Although the study examined other instances of mood morphology, no cases of subjunctive morphology, besides negative imperatives in Ibero-Romance, or other moods (e.g., conditional) were found. Since mood morphology is absent at Stages 1-2, it is omitted from the columns below.

Errors in the use of agreement morphemes (e.g., use of an ending of an incorrect person-number combination) were distinguished from *correct* productions and are also noted in all tables as *Err*. Additionally, inflected forms are distinguished from incorrect use of non-finite forms in finite contexts, namely *Root Infinitives*. The latter are noted as INF under the relevant Agreement columns/tables. Note also that 2PL is never attested in any child across the periods studied. This is possibly because the kinds of conversational contexts and interactions do not favour its use (as also noted in the study on Italian children in Caprin and Guasti, 2009, p. 30). Therefore, I omit 2PL in all relevant tables.

Finally, like in **Appendix B**, **Section B.2**, we also remark that Simone speaks a southern German dialect. This is relevant insofar as verbs are often inflected with *-e*, an agreement marker which is crucially ambiguous between 1SG and the infinitive (cf. the Standard German infinitival marker *-en*). I did not exclude forms with *-e* on the grounds of this ambiguity, following previous work (Verrips and Weissenborn, 1992). This approach would also be inappropriate for present purposes, in that it would discard all 1SG morphology. That decision would be undesirable, given one of the aims of the study was to ascertain the productivity of each person-number morpheme at various stages of development (see §3). I assumed, instead, that whenever a verb ending in *-e* was placed in transparent verb-final position, it signalled infinitival morphology, *not* 1SG agreement. Similarly, if a finite verb ending in *-e* had raised to the V2/V1 position (following the diagnostics outlined in §3.1.3.1, I took that to indicate likely 1SG morphology.

As per earlier Appendices, this study also excluded repetitions and imitations of an adult utterance or of the child's previous utterance. Utterances in which the child was singing or reciting lyrics were also discarded, as well as unintelligible utterances (either due to pronunciation or transcription). Furthermore, in the comparatively few cases where a given morpheme was ambiguous (because of syncretism) and this ambiguity was not resolvable based on context or accompanying subjects (if any), these verb forms were discarded. This applied primarily to Simone, given the syncretism between 1SG and INF raised above.

The glosses for the abbreviations used in the tables below can be found in the **List of abbreviations** (page xix). Note that, in Catalan, the 'periphrastic past' (abbreviated as PERIPHR PST), often also called a 'go past', refers to the tense formed with the auxiliary verb *anar* 'to go' plus the infinitive.

C.1 Summary of the development of agreement morphology

C.1.1 Romance

Catalan

Table C.1 Laura's development of agreement morphology

MLU	3SG	1SG	2SG	3PL	1PL	INF	*Err*
1.03							
1.09	3 (100%)						
1.15	8 (80%)		2 (20%)				
1.15	21 (95.5%)					1 (4.5%)	
1.35	4 (33.3%)	6 (50%)				2 (16.7%)	
1.3	4 (100%)						
1.44	6 (85.7%)		1 (14.3%)				
1.64	17 (63%)	6 (22.2%)	2 (7.4%)	2 (7.4%)			
1.76	10 (66.7%)	4 (26.7%)				1 (6.7%)	
1.78	20 (37.7%)	18 (34%)	7 (13.2%)	3 (5.7%)		5 (9.4%)	
1.88	41 (54.7%)	18 (24%)	12 (16%)	1 (1.3%)	1 (1.3%)	2 (2.7%)	
1.98	52 (62.7%)	24 (28.9%)	3 (3.6%)		1 (1.2%)	3 (3.6%)	
2.42	55 (38.7%)	49 (34.5%)	29 (20.4%)	3 (2.1%)	5 (3.5%)		1 (0.7%)

Table C.2 Gisela's development of agreement morphology

MLU	3sg	1sg	2sg	3pl	1pl	inf	*Err*
1.04							
1.02	1 (100%)						
1.13	7 (100%)						
1.16	3 (100%)						
1.14	12 (100%)						
1.09	4 (100%)						
1.53	4 (100%)						
1.5	12 (60%)	7 (35%)		1 (5%)			
1.58	15 (88.2%)	1 (5.9%)		1 (5.9%)			
2.32	4 (25%)	11 (68.8%)		1 (6.3%)			

Italian

Table C.3 Martina's development of agreement morphology

MLU	3sg	1sg	2sg	3pl	1pl	inf	*Err*
1.27	3 (75%)	1 (25%)					
1.9	14 (87.5%)	1 (6.3%)					1 (6.3%)
1.57	14 (73.7%)	3 (15.8%)	1 (5.3%)			1 (5.3%)	
1.59	3 (42.9%)	2 (28.6%)	1 (14.3%)			1 (14.3%)	
1.66	14 (48.3%)	9 (31%)	3 (10.3%)	1 (3.4%)		1 (3.4%)	1 (3.4%)
1.99	29 (61.7%)	8 (17%)	6 (12.8%)			3 (6.4%)	1 (2.2%)
1.99	34 (50.7%)	20 (29.9%)	2 (3%)			10 (14.9%)	1 (1.5%)
1.86	14 (63.6%)	4 (18.2%)	2 (9.1%)	1 (4.5%)		1 (4.5%)	
2.55	26 (51%)	18 (35.3%)	6 (11.8%)		1 (2%)		
2.64	49 (53.3%)	23 (25%)	16 (17.4%)	2 (2.2%)	1 (1.1%)		1 (1.1%)

Table C.4 Rosa's development of agreement morphology

MLU	3SG	1SG	2SG	3PL	1PL	INF	*Err*
1.27	4 (100%)						
1.5	1 (100%)						
1.44	3 (75%)					1 (25%)	
1.39	7 (87.5%)					1 (12.5%)	
1.41	15 (100%)						
1.5	11 (78.6%)		2 (14.3%)			1 (7.1%)	
1.54	8 (66.7%)	2 (16.7%)	2 (16.7%)				
1.75	45 (64.3%)	6 (8.6%)	12 (17.1%)	6 (8.6%)		1 (1.4%)	
1.78	41 (82%)	2 (4%)	4 (8%)	2 (4%)		1 (2%)	
2.14	18 (48.6%)	10 (27%)	8 (21.6%)				1 (2.7%)
2.6	43 (75.4%)	8 (14%)	4 (7%)	1 (1.8%)			1 (1.8%)
1.34	7 (63.6%)	1 (9.1%)	1 (9.1%)			2 (18.2%)	
2.78	42 (60.9%)	14 (20.3%)	10 (14.5%)	2 (2.9%)			1 (1.4%)
2.87	47 (57.3%)	15 (18.3%)	16 (19.5%)	2 (2.4%)		2 (2.4%)	
2.54	39 (48.1%)	20 (24.7%)	14 (17.3%)	5 (6.2%)	2 (2.5%)	1 (1.2%)	

Spanish

Table C.5 Irene's development of agreement morphology

MLU	3SG	1SG	2SG	3PL	1PL	INF	*Err*
1.42							
1.08							
1.05							
1.27	1 (100%)						
1.32	1 (100%)						
1.0							
1.59	3 (100%)						
1.53							
1.43							
1.57	3 (100%)						
1.61	2 (100%)						

2.15						
1.84	2 (100%)					
1.69	7 (100%)					
1.8	2 (100%)					
1.94	21 (95.4%)	1 (4.5%)				
1.94	7 (100%)					
3.0	1 (100%)					
1.88	11 (91.7%)					1 (8.3%)
2.24	8 (88.9%)	1 (11.1%)				
2.28	16 (84.2%)	2 (10.5%)	1 (5.3%)			
3.28	20 (76.9%)	3 (11.5%)		3 (11.5%)		
2.53	38 (92.7%)	2 (4.9%)			1 (2.4%)	
3.09	7 (100%)					
2.62	14 (77.8%)	3 (16.7%)			1 (5.6%)	
2.36	8 (80%)					2 (20%)
2.22	4 (100%)					
3.09	21 (84%)	3 (12%)	1 (4%)			

Table C.6 Koki's development of agreement morphology

MLU	3SG	1SG	2SG	3PL	1PL	INF	*Err*
1.96	46 (70.8%)	15 (23.1%)	4 (6.2%)				
2.54	46 (83.6%)	4 (7.3%)	1 (1.8%)	2 (3.6%)	1 (1.8%)	1 (1.8%)	
2.47	46 (67.6%)	8 (11.8%)		5 (7.4%)		3 (4.4%)	6 (8.8%)
2.51	49 (65.3%)	10 (13.3%)	1 (1.3%)	8 (10.7%)	4 (5.3%)	1 (1.3%)	2 (2.7%)
2.47	54 (64.3%)	18 (21.4%)	1 (1.2%)	10 (11.9%)	1 (1.2%)		
2.07	16 (76.2%)	3 (14.3%)	1 (4.8%)	1 (4.8%)			

C.1.2 Germanic

German

Table C.7 Kerstin's development of agreement morphology

MLU	3SG	1SG	2SG	3PL	1PL	INF	*Err*
1.09	1 (100%)						
1.31							
1.38	1 (100%)						
1.38	1 (100%)						
1.58	9 (90%)	1 (10%)					
1.36							
1.45	3 (100%)						
1.46	3 (60%)					2 (40%)	
1.37	3 (50%)					3 (50%)	
1.58	6 (66.7%)	1 (11.1%)				2 (22.2%)	
1.39	3 (33.3%)					6 (66.7%)	
1.34	12 (85.7%)	1 (7.1%)				1 (7.1%)	
1.2	1 (33.3%)			1 (33.3%)		1 (33.3%)	
1.4	7 (77.8%)					2 (22.2%)	
1.12						3 (100%)	
1.28	4 (50%)					4 (50%)	
1.38	7 (58.3%)					5 (41.7%)	
1.53	8 (47.1%)	2 (11.8%)				7 (41.2%)	
1.76	22 (73.3%)	3 (10%)		1 (3.3%)	1 (3.3%)	3 (10%)	
1.68	12 (48%)	3 (12%)	2 (8%)		1 (4%)	6 (24%)	1 (4%)
1.59	9 (47.4%)	4 (21.1%)				4 (21.1%)	2 (10.5%)
1.58	10 (58.8%)	1 (5.9%)	1 (5.9%)		1 (5.9%)	4 (23.5%)	
1.72	18 (69.2%)		1 (3.9%)			7 (26.9%)	
1.65	12 (52.2%)	2 (8.7%)	1 (4.3%)		1 (4.3%)	7 (30.4%)	
1.81	12 (60%)	2 (10%)				5 (25%)	1 (5%)
1.86	14 (35%)	6 (15%)	1 (2.5%)	1 (2.5%)		16 (40%)	2 (5%)
1.98	10 (45.5%)			1 (4.5%)		11 (50%)	
1.67	24 (40.7%)	12 (20.3%)	3 (5.1%)	4 (6.8%)	3 (5.1%)	13 (22%)	
1.8	8 (15.4%)	4 (7.7%)	6 (11.5%)			34 (65.4%)	
1.78	6 (27.3%)	4 (18.2%)	1 (4.5%)			11 (50%)	

2.25	28 (60.9%)	5 (10.9%)	2 (4.3%)	1 (2.2%)		10 (21.7%)	
1.62	17 (37%)	8 (17.4%)	1 (2.2%)			19 (41.3%)	1 (2.2%)
2.13	43 (41.3%)	26 (25%)	1 (1%)		1 (1%)	21 (20.2%)	12 (11.5%)

Table C.8 Simone's development of agreement morphology

MLU	3SG	1SG	2SG	3PL	1PL	INF	*Err*
1.54	6 (75%)	1 (12.5%)				1 (12.5%)	
1.62	44 (47.3%)	5 (5.4%)				44 (47.3%)	
1.69	32 (69.6%)	2 (4.3%)	1 (2.2%)			11 (23.9%)	
1.71	22 (38.6%)	3 (5.3%)	1 (1.8%)			31 (54.4%)	
1.52	2 (100%)						
1.94	8 (36.4%)					14 (63.6%)	
1.53	11 (23.9%)	3 (6.5%)				32 (69.6%)	
1.88	15 (45.5%)					18 (54.5%)	
2.21	16 (51.6%)			1 (3.2%)		14 (45.2%)	
2.27	48 (57.8%)			1 (1.2%)	1 (1.2%)	33 (39.8%)	
2.28	20 (34.5%)	3 (5.2%)				35 (60.3%)	
2.31	19 (57.6%)					14 (42.4%)	
2.31	37 (63.8%)	3 (5.2%)	2 (3.4%)	3 (5.2%)		13 (22.4%)	
1.98	43 (51.2%)	2 (2.4%)	2 (2.4%)	4 (4.8%)		33 (39.3%)	
1.93	33 (61.1%)	3 (5.6%)				18 (33.3%)	
1.63	27 (46.6%)	2 (3.5%)				27 (46.6%)	2 (3.5%)
1.72	27 (62.8%)	2 (4.7%)				14 (32.6%)	
1.78	35 (54.7%)	2 (3.1%)			1 (1.6%)	25 (39.1%)	1 (1.6%)
2.03	9 (69.2%)	1 (7.7%)				3 (23.1%)	
1.79	36 (62.1%)	2 (3.4%)				20 (34.5%)	
1.71	29 (63%)	2 (4.3%)		2 (4.3%)		12 (26.1%)	1 (2.2%)
1.94	65 (75.6%)	3 (3.5%)		3 (3.5%)		14 (16.3%)	1 (1.2%)
1.66	15 (44.1%)	3 (8.8%)				16 (47.1%)	
2.22	23 (79.3%)	1 (3.4%)				4 (13.8%)	1 (3.4%)
2.0	3 (60%)					2 (40%)	
2.09	92 (55.4%)	19 (11.4%)	7 (4.2%)	1 (0.6%)		47 (28.3%)	
1.99	73 (62.4%)	9 (7.7%)	4 (3.4%)	2 (1.7%)		29 (24.8%)	

1.82	52 (52%)	27 (27%)	1 (1%)	3 (1%)		17 (17%)	
1.89	44 (60.3%)	20 (27.4%)		4 (5.5%)		5 (6.8%)	
1.96	30 (44.8%)	18 (26.9%)	1 (1.5%)	5 (7.5%)		13 (19.4%)	
1.92	61 (67.8%)	13 (14.4%)		4 (4.4%)		10 (11.1%)	2 (2.2%)
2.52	72 (62.6%)	18 (15.7%)	4 (3.5%)	6 (5.2%)		13 (11.3%)	2 (1.7%)
2.35	19 (59.4%)	4 (12.5%)	1 (3.1%)			8 (25%)	
2.62	48 (71.6%)	4 (6%)	3 (4.5%)	4 (6%)		8 (11.9%)	
2.67	12 (66.7%)	1 (5.6%)		4 (22.2%)		1 (5.6%)	
3.35	42 (64.6%)	6 (9.2%)	3 (4.6%)	4 (6.2%)	1 (1.5%)	7 (10.8%)	2 (3.1%)
4.04	37 (52.9%)	10 (14.3%)	4 (5.7%)	5 (7.1%)		14 (20%)	

Dutch

Table C.9 Josse's development of agreement morphology

MLU	3sg	1sg	2sg	3pl	1pl	inf	*Err*
1.2	1 (3.3%)					29 (96.7%)	
1.44	1 (9.1%)	2 (18.2%)				8 (72.7%)	
1.55	10 (29.4%)	5 (14.7%)				19 (55.9%)	
1.59	7 (24.1%)	2 (6.9%)				20 (69%)	
1.69	11 (47.8%)	3 (13%)				9 (39.1%)	
1.74	7 (17.1%)	2 (4.9%)	1 (2.4%)			31 (75.6%)	
1.94	24 (53.3%)	10 (22.2%)				11 (24.4%)	
2.14	8 (21.1%)	7 (18.4%)	6 (15.8%)		1 (2.6%)	16 (42.1%)	
1.9	17 (56.7%)	2 (6.7%)			2 (6.7%)	9 (30%)	
1.59	5 (45.4%)					6 (54.5%)	
2.17	29 (53.7%)	7 (13%)		2 (3.7%)		16 (29.6%)	
2.11	33 (51.6%)	13 (20.3%)	6 (9.4%)		1 (1.6%)	11 (17.2%)	
2.19	30 (35.3%)	27 (31.8%)	11 (12.9%)	2 (2.4%)	3 (3.5%)	11 (12.9%)	1 (1.2%)
2.46	93 (66.4%)	14 (10%)	8 (5.7%)	2 (1.4%)	4 (2.9%)	17 (12.1%)	2 (1.4%)
2.25	47 (39.5%)	26 (21.8%)	19 (16%)		6 (5%)	20 (16.8%)	1 (0.8%)
2.55	68 (63.6%)	21 (19.6%)	8 (7.5%)	1 (0.9%)	2 (1.9%)	5 (4.7%)	2 (1.9%)
2.42	66 (66.7%)	16 (16.2%)	8 (8.1%)	7 (7.1%)		2 (2%)	
2.8	48 (54.5%)	9 (10.2%)	20 (22.7%)	5 (5.7%)	3 (3.4%)	2 (2.3%)	1 (1.1%)

Table C.10 Sarah's development of agreement morphology

MLU	3sg	1sg	2sg	3pl	1pl	inf	*Err*
1.12	2 (22.2%)					7 (77.8%)	
1.17	1 (7.7%)					12 (92.3%)	
1.07	1 (7.1%)	1 (7.1%)				12 (85.7%)	
1.17	1 (8.3%)	4 (33.3%)				7 (58.3%)	
1.09	1 (6.7%)	4 (26.7%)				10 (66.7%)	
1.17	1 (16.7%)	1 (16.7%)				3 (50%)	1 (16.7%)
1.25	1 (4%)	1 (4%)				23 (92%)	
1.37	4 (11.1%)	1 (2.8%)	1 (2.8%)	1 (2.8%)		29 (80.6%)	
1.68	8 (22.9%)	10 (28.6%)				17 (48.6%)	
1.88	8 (18.6%)	18 (41.9%)				17 (39.5%)	
2.11	38 (50.7%)	12 (16%)	7 (9.3%)			16 (21.3%)	2 (2.7%)
2.05	16 (35.6%)	14 (31.1%)	2 (4.4%)	1 (2.2%)		11 (24.4%)	1 (2.2%)
2.53	42 (40%)	29 (27.6%)	19 (18.1%)	3 (2.9%)		12 (11.4%)	
2.34	24 (33.3%)	36 (50%)	2 (2.8%)			10 (13.9%)	
2.46	40 (38.8%)	22 (21.4%)	3 (2.9%)	4 (3.9%)		24 (23.3%)	10 (9.7%)
2.47	21 (23.6%)	45 (50.6%)	4 (4.5%)			17 (19.1%)	2 (2.2%)
2.59	20 (19.8%)	55 (54.4%)	20 (19.8%)			6 (5.9%)	
2.74	45 (45.9%)	28 (28.6%)	11 (11.2%)	1 (1%)	1 (1%)	11 (11.2%)	1 (1%)
2.45	31 (44.3%)	21 (30%)	5 (7.1%)	1 (1.4%)	1 (1.4%)	8 (11.4%)	3 (4.3%)
2.8	27 (35.5%)	42 (55.3%)				6 (7.9%)	1 (1.3%)
2.51	29 (40.8%)	28 (39.4%)	8 (11.3%)			6 (8.5%)	
2.66	51 (42.1%)	35 (28.9%)	17 (14%)	2 (1.7%)	3 (2.5%)	12 (9.9%)	1 (0.8%)
2.97	49 (36%)	45 (33.1%)	17 (12.5%)	7 (5.1%)	4 (2.9%)	11 (8.1%)	3 (2.2%)
2.59	40 (55.6%)	24 (33.3%)	1 (1.4%)	1 (1.4%)	2 (2.8%)	4 (5.6%)	
3.15	91 (65.5%)	25 (18%)	10 (7.2%)	4 (2.9%)	3 (2.2%)	3 (2.2%)	3 (2.2%)
2.88	44 (43.1%)	43 (42.2%)	8 (7.8%)	1 (1%)		3 (2.9%)	3 (2.9%)
2.87	51 (55.4%)	31 (33.7%)	6 (6.5%)	1 (1.1%)		3 (3.3%)	
3.64	27 (22.7%)	29 (24.4%)	7 (5.9%)		3 (2.5%)	51 (42.9%)	2 (1.7%)

C.2 Tense and agreement morphology until Stages 1 and 2

C.2.1 Romance

Catalan

Table C.11 Laura's use of tense and agreement marking

MLU	Agreement	Imperatives	Tense
1.03		3	
1.09	3SG: 3 (*estar* 'to be')	1	
1.15	3SG: 8 (*caure* 'to fall', estar 'to be') 2SG: 2 (*treure* 'to take out')	3	
1.15	3SG: 21 (*caure* 'to fall', *estar* 'to be', *cremar* 'to burn', *ser* 'to be', *plorar* 'to cry') INF: 1 (*fer* 'to do/make')	3	
1.35	3SG: 4 (*cremar* 'to burn', *estar* 'to be', *haver-hi* 'there is/are') 1SG: 6 (*saber* 'to know', *cremar* 'to burn', *voler* 'to want') INF: 2 (*dormir* 'to sleep')	1	
1.3	3SG: 4 (*caure* 'to fall', *estar* 'to be', *ser* 'to be')	17	
1.44	3SG: 6 (*anar* 'to go', *haver-hi* 'there is/are', *haver*$_{aux}$ 'have', *ser* 'to be', *plorar* 'to cry') 2SG: 1 (*fer* 'to do/make')	3	PRES PERF: 1
1.64	3SG: 17 (*ser* 'to be', *estar* 'to be', *rentar* 'to wash', *pentinar* 'to comb', *caure* 'to fall') 1SG: 6 (*saber* 'to know', *tallar*, 'to cut', *agafar*, 'to get/pick up') 2SG: 2 (*posar* 'to put', *veure*, 'to see') 3PL: 2 (*ser* 'to be')	11 (SBJV: 1)	
1.76	3SG: 10 (*ser*, 'to be', *caure* 'to fall', *haver*$_{aux}$ 'have') 1SG: 4 (*caure* 'to fall', *anar*, 'to go', *poder*, 'can') INF: 1 (*escopir* 'to spit')	4	PRES PERF: 2

1.78	3SG: 20 (*haver*$_{aux}$ ‘have’, *cremar* ‘to burn’, *deure* ‘must’, *ser* ‘to be’, *estar* ‘to be’) 1SG: 18 (*haver*$_{aux}$ ‘have’, *voler*, ‘to want’, *venir* ‘to come’, *aguantar* ‘to hold’, *portar* ‘to carry’) 2SG: 7 (*veure* ‘to see’, *venir*, ‘to come’, *saber* ‘to know’, *estar* ‘to be’) 3PL: 3 (*ser* ‘to be’) INF: 5 (*jugar* ‘to play’, *prendre*, ‘to take’, *portar* ‘to bring’, *sortir* ‘to exit/go out’, *fer* ‘to do/make’)	19	PRES PERF: 8
1.88	3SG: 41 (*portar* ‘to carry’, *treure* ‘to take out/off’, *haver-hi* ‘there is/are’, *costar* ‘to cost’, *ser* ‘to be’) 1SG: 18 (*tirar* ‘to throw’, *voler* ‘to want’, *treure* ‘to take out/off’, *veure* ‘to see’, *estar* ‘to be’) 2SG: 12 (*veure* ‘to see’, *tenir* ‘to have’, *saber* ‘to know’, *haver*$_{aux}$ ‘have’) 3PL: 1 (*haver*$_{aux}$ ‘have’) 1PL: 1 (*fer* ‘to do/make’) INF: 2 (*posar* ‘to put’)	16 (SBJV: 4)	PRES PERF: 7 FUT: 3
1.98	3SG: 52 (*estar* ‘to be’, *ser* ‘to be’, *voler* ‘to want’, *anar* ‘to go’, *sortir* ‘to exit/go out’) 1SG: 24 (*voler* ‘to want’, *haver*$_{aux}$ ‘have’, *anar* ‘to go’, *ser* ‘to be’, *mullar-se* ‘to get wet’) 2SG: 3 (*deixar* ‘to leave/let’, *voler* ‘to want’, *tenir* ‘to have’) 1PL: 1 (*anar* ‘to go’) INF: 3 (*posar* ‘to put’, *afaitar* ‘to shave’, *menjar* ‘to eat’)	11 (SBJV: 3)	PRES PERF: 3 PERIPHR PST: 1 IMPF: 5
2.42	3SG 55 (*estar* ‘to be’, *menjar* ‘to eat’, *ser* ‘to be’, *haver*$_{aux}$ ‘have’, *anar*$_{aux}$ ‘to go’) 1SG: 49 (*posar* ‘to put’, *poder* ‘can’, *fer* ‘to do/make’, *tenir* ‘to have’, *ser* ‘to be’) 2SG: 29 (*menjar* ‘to eat’, *veure* ‘to see’, *haver*$_{aux}$ ‘have’, *tenir* ‘to have’, *aguantar* ‘to hold’) 3PL: 3 (*estar* ‘to be’, *ser* ‘to be’, *quedar* ‘to remain/stay’)	33 (SBJV: 2)	PRES PERF: 11 PERIPHR PST: 1 IMPF: 2 FUT: 5

	1PL: 5 (*poder* 'can', *menjar* 'to eat', *tenir* 'to have', *dinar* 'to have lunch') *Err:* 1		

Table C.12 Gisela's use of tense and agreement marking

MLU	Agreement	Imperatives	Tense
1.04			
1.02	3SG: 1 (*caure* 'to fall')		
1.13	3SG: 7 (*caure* 'to fall')	1	
1.16	3SG: 3 (*ser* 'to be', *haver-hi* 'there is/are')		
1.14	3SG: 12 (*ser* 'to be', *estar* 'to be')		
1.09	3SG: 4 (*estar* 'to be')		
1.53	3SG: 4 (*ser* 'to be', *estar* 'to be', *caure* 'to fall', *anar* 'to go')		
1.5	3SG: 12 (*haver-hi* 'there is/are', *estar* 'to be', *ser* 'to be', *agradar* 'to like', *cremar* 'to burn') 1SG: 7 (*voler* 'to want', *tenir* 'to have') 3PL: 1 (*ser* 'to be')	1	
1.58	3SG: 15 (*estar* 'to be', *caure* 'to fall', *ser* 'to be', $haver_{aux}$ 'have') 1SG: 1 (*voler* 'to want') 3PL: 1 (*volar* 'to fly')	3	PRES PERF: 1
2.32	3SG: 4 (*ser* 'to be', *menjar* 'to eat', *haver-hi* 'there is/are', *fer* 'to do/make') 1SG: 11 (*voler* 'to want', *poder* 'can', *tenir* 'to have', *tocar* 'to touch') 3PL: 1 (*ser* 'to be')	3 (SBJV: 2)	

Italian

Table C.13 Martina's use of tense and agreement marking

MLU	Agreement	Imperatives	Tense
1.26	3SG: 3 (*essere* 'to be') 1SG: 1 (*avere$_{aux}$* 'have')	2	PRES PERF: 1
1.9	3SG: 14 (*piangere* 'to cry', *essere* 'to be', *fare* 'to do/make') 1SG: 1 (*avere* 'to have') *Err:* 1	6	
1.57	3SG: 14 (*bruciare* 'to burn', *levare* 'to remove/raise', *andare* 'to go', *essere* 'to be', *girare* 'to turn/stir') 1SG: 3 (*toccare* 'to touch', *mettere* 'to put', *aprire* 'to open') 2SG: 1 (*aprire* 'to open') INF: 1 (*mettere* 'to put')	2	
1.59	3SG: 3 (*vedere* 'to see', *essere* 'to be') 1SG: 2 (*avere* 'to have', *rimettere* 'to put back') 2SG: 1 (*devere* 'should') INF: 1 (*levare* 'to remove/raise')	4	
1.66	3SG: 14 (*volare* 'to fly', *essere* 'to be', *avere* 'to have', *fare* 'to do/make', *girare* 'to turn/stir') 1SG: 9 (*mettere* 'to put', *avere$_{aux}$* 'have', *leggere* 'to read', *volere* 'to want', *aiutare* 'to help') 2SG: 3 (*aprire* 'to open', *essere* 'to be') 3PL: 1 (*dormire* 'to sleep') INF: 1 (*aprire* 'to open') *Err:* 1	5	PRES PERF: 1
1.99	3SG: 29 (*essere* 'to be', *chiamarsi* 'to be called', *dormire* 'to sleep', *scrivere* 'to write', *chiudare* 'to close') 1SG: 8 (*pigliare* 'to take/grab', *prendere* 'to take/get', *girare* 'to turn/stir', *mettere* 'to put', *volere* 'to want')	4	PRES PERF: 1

	2SG: 6 (*pigliare* 'to take/grab', *ricordare* 'to remember', *avere* INF: 3 (*mangiare* 'to eat') *Err:* 1		
1.99	3SG: 34 (*essere* 'to be', *chiamarsi* 'to be called', *leggere* 'to read', *dormire* 'to sleep', *aprire* 'to open') 1SG: 20 (*volere* 'to want', $avere_{aux}$ 'have', *pigliare* 'to take/grab', *mettere* 'to put', *mangiare* 'to eat') 2SG: 2 (*guidare* 'to drive/guide', *mettere* 'to put') INF: 10 (*lavorare* 'to work', *vedere* 'to see', *chiudere* 'to close', *aprire* 'to open', *gridare* 'to shout') *Err:* 1	8	PRES PERF: 1
1.86	3SG: 14 (*portare* 'to carry/bring', *essere* 'to be', *piovere* 'to rain', *fare* 'to do/make', *chiamarsi* 'to be called') 1SG: 4 (*fare* 'to do/make', *mettere* 'to put', *dare* 'to give', *sapere* 'to know') 2SG: 2 (*aprire* 'to open', *girare* 'to turn/stir') 3PL: 1 (*dicere* 'to say') INF: 1 (*fare* 'to do/make')	13	PRES PERF: 1 IMPF: 1
2.55	3SG: 26 (*essere* 'to be', *avere* 'to have', *chiamarsi* 'to be called', *piovere* 'to rain', *scrivere* 'to write') 1SG: 18 (*fare* 'to do/make', *tirare* 'to throw', *avere* 'to have', *volere* 'to want', *sapere* 'to know') 2SG: 6 (*volere* 'to want', *aprire* 'to open', *piangere* 'to have') 1PL: 1 (*giocare* 'to play')	9	IMPF: 2
2.64	3SG: 49 (*essere* 'to be', *dicere* 'to say', *lavorare* 'to work', *piangere* 'to cry', *mettere* 'to put') 1SG: 23 (*conoscere* 'to know/meet', *avere* 'to have', *fare* 'to do/make', *asciugare* 'to dry', *mangiare* 'to eat') 2SG: 16 (*leggere* 'to read', *prendere* 'to take/get', *mettere* 'to put', *guardare* 'to watch', *vedere* 'to see')	14	PRES PERF: 2 IMPF: 1 PROG: 1

	3PL: 2 (*essere* 'to be')		
	1PL: 1 (*andare* 'to go')		
	Err: 1		

Table C.14 Rosa's use of tense and agreement marking

MLU	Agreement	Imperatives	Tense
1.27	3SG: 4 (*essere* 'to be')		
1.5	3SG: 1 (*essere* 'to be')		
1.44	3SG: 3 (*essere* 'to be', *avere* 'to have') INF: 1 (*sedere* 'to sit')	1	
1.39	3SG: 7 (*essere* 'to be', $essere_{aux}$ 'be', *fumare* 'to smoke') INF: 1 (*levare* 'to remove/raise')	4	PRES PERF: 1 IMPF: 1
1.41	3SG: 15 (*essere* 'to be', *stare* 'to be', *riuscire* 'to be able to/succeed')	4	
1.5	3SG: 11 (*essere* 'to be', $essere_{aux}$ 'be', *riuscire* 'to be able to/succeed') 2SG: 2 (*portare* 'to bring/take', *vedere* 'to see') INF: 1 (*vedere* 'to see')	5	PRES PERF: 1
1.54	3SG: 8 (*essere* 'to be', *riuscire* 'to be able to/succeed') 1SG: 2 (*portare* 'to bring/take', *vedere* 'to see') 2SG: 2 (*reggere* 'to hold', *dare* 'to give')	4	IMPF: 1
1.75	3SG: 45 (*essere* 'to be', *riuscire* 'to be able to/succeed', *chiudere* 'to close', *mettere* 'to put', *fare* 'to do/make', *mangiare* 'to eat') 1SG: 6 (*chiudere* 'to close', *volere* 'to want', *tirare* 'to throw', *vedere* 'to see', *sparare* 'to shoot') 2SG: 12 (*aprire* 'to open', *mettere* 'to put', *chiudere* 'to close') 3PL: 6 (*fare* 'to do/make', *essere* 'to be', *avere* 'to have') INF: 1 (*vedere* 'to see')	4	

1.78	3sg: 41 (*essere* ‘to be’, *chiamarsi* ‘to be called’, *fare* ‘to do/make’, $essere_{aux}$ ‘be’, *stare* ‘to stand/stay’) 1sg: 2 (*volere* ‘to want’) 2sg: 4 (*aprire* ‘to open’) 3pl: 2 (*essere* ‘to be’) inf: 1 (*toccare* ‘to touch’)	2	pres perf: 1
2.14	3sg: 18 (*essere* ‘to be’, *chiamare* ‘to call’, *andare* ‘to go’, *fare* ‘to do/make’, $essere_{aux}$ ‘be’) 1sg: 10 (*volere* ‘to want’, *scartare* ‘to discard’, *portare* ‘to bring/take’, *mettere* ‘to put’, $avere_{aux}$ ‘have’) 2sg: 8 (*aprire* ‘to open’, *essere* ‘to be’, *mettere* ‘to put’) *Err:* 1	1	pres perf: 3
2.6	3sg: 43 (*chiamarsi* ‘to be called’, *essere* ‘to be’, *piangere* ‘to cry’, *scappare* ‘to escape’, $essere_{aux}$ ‘be’) 1sg: 8 (*mettere* ‘to put’, *volere* ‘to want’, $avere_{aux}$ ‘have’, *chiudere* ‘to close’) 2sg: 4 (*sapere* ‘to know’, *fare* ‘to know’, *andare* ‘to go’, *vedere* ‘to see’) 3pl: 1 (*essere* ‘to be’) *Err:* 1	4	pres perf: 3 impf: 1
1.34	3sg: 7 (*essere* ‘to be’, *rompere* ‘to break’) 1sg: 1 (*volere* ‘to want’) 2sg: 1 (*tenere* ‘to keep’) inf: 2 (*sedere* ‘to sit’)	4	
2.78	3sg: 42 (*essere* ‘to be’, *piangere* ‘to cry’, *chiudere* ‘to close’, *volere* ‘to want’, *spolverare* ‘to dust’) 1sg: 14 (*volere* ‘to want’, *girare* ‘to turn/stir’, *mettere* ‘to put’, *essere* ‘to be’, *cascare* ‘to fall/drop’) 2sg: 10 ($avere_{aux}$ ‘be’, *mettere* ‘to put’, *agganciare* ‘to hook/attach’, *aprire* ‘to open’) 3pl: 2 (*vendere* ‘to sell’, *cascare* ‘to fall/drop’) *Err:* 1	2	pres perf: 2

2.87	3SG: 47 (*andare* 'to go', *essere* 'to be', *mettere* 'to put', *riuscire* 'to be able to/succeed', *aprire* 'to open') 1SG: 15 (*volere* 'to want', *reggere* 'to hold', *mettere* 'to put', *avere*$_{aux}$ 'have', *levare* 'to remove/raise') 2SG: 16 (*aprire* 'to open', *volere* 'to want', *andare* 'to go', *mettere* 'to put', *fare* 'to do/make') 3PL: 2 (*essere* 'to be') INF: 2 (*lavare* 'to wash', *sedere* 'to sit')	20	PRES PERF: 4 IMPF: 2
2.54	3SG: 39 (*essere* 'to be', *avere* 'to have', *andare* 'to go', *mettere* 'to put', *essere*$_{aux}$ 'be') 1SG: 20 (*accendere* 'to turn on', *volere* 'to want', *lavare* 'to wash', *andare* 'to go', *mettere* 'to put') 2SG: 14 (*agganciare* 'to hook/attach', *aprire* 'to open', *andare* 'to go', *fare* 'to do/make', *venire* 'to come') 3PL: 5 (*avere* 'to have', *fare* 'to do/make', *essere* 'to be') 1PL: 2 (*girare* 'to turn/stir', *guardare* 'to watch') INF: 1 (*andare* 'to go')	16	PRES PERF: 2

Spanish

Table C.15 Irene's use of tense and agreement marking

MLU	Agreement	Imperatives	Tense
1.42			
1.08		1	
1.05			
1.27	3SG: 1 (*haber* 'there is/are')		
1.32	3SG: 1 (*ser* 'to be')	3	
1.0			

1.59	3sg: 3 (*estar* ‘to be’, *acabar* ‘to finish’, *marchar* ‘to leave’)	6	prfv: 2
1.53			
1.43			
1.57	3sg: 3 (*caer* ‘to fall’, *estar* ‘to be’)		prfv: 2
1.61	3sg: 2 (*acabar* ‘to finish’)		prfv: 2
2.15			
1.84	3sg: 2 (*acabar* ‘to finish’, *haber* ‘there is/are’)		prfv: 1
1.69	3sg: 7 (*cremar* ‘to burn’, *haber* ‘there is/are’, *llover* ‘to rain’, *caer* ‘to fall’)	4	prfv: 1
1.8	3sg: 2 (*caer* ‘to fall’, *acabar* ‘to finish’)	1	prfv: 2
1.94	3sg: 21 (*acabar* ‘to finish’, *marchar* ‘to finish’, *oír* ‘to hear’, *ser* ‘to be’) 1sg: 1 (*vivir* ‘to live’)		prfv: 18
1.94	3sg: 7 (*ser* ‘to be’, *haber* ‘there is/are’)	1	
3.0	3sg: 1 (*quitar* ‘to remove’)	3 (sbjv: 2)	
1.88	3sg: 11 (*romper* ‘to break’, *poder* ‘can’, *caer* ‘to break’, *acabar* ‘to finish’, *ser* ‘to be’) *Err*: 1		prfv: 3
2.24	3sg: 8 (*caer* ‘to fall’, *acabar* ‘to finish’, *ser* ‘to be’, *estar* ‘to be’) 1sg: 1 (*echar* ‘to throw’)	1	prfv: 3
2.28	3sg: 16 (*saber* ‘to know’, *caer* ‘to fall’, *volar* ‘to fly’, *quitar* ‘to remove’, *ser* ‘to be’) 1sg: 2 (*poder* ‘can’, *caer* ‘to fall’) 2sg: 1 (*poder* ‘can’)	3	prfv: 6
3.28	3sg: 20 (*estar*$_{aux}$ ‘be’, *estar* ‘to be’, *saber* ‘to know’, *llamarse* ‘to be called’, *llamar* ‘to call’) 1sg: 3 (*chupar* ‘to suck’, *quedar* ‘to remain/stay’) 3pl: 3 (*estar* ‘to be’, *llamar* ‘to call’)	4 (sbjv: 1)	prfv: 6 impf: 1 prog: 1
2.53	3sg: 38 (*poner* ‘to put’, *cansarse* ‘to get tired’, *caer* ‘to fall’, *tener* ‘to have’, *ser* ‘to be’) 1sg: 2 (*coger* ‘to get/take’, *echar* ‘to throw’) 1pl: 1 (*poder* ‘can’)	3	prfv: 18 impf: 2

3.09	3sg: 7 (*ser* ‘to be’, *querer* ‘to want/love’, *caer* ‘to fall’, *tener* ‘to have’, *haber* ‘there is/are’)	2 (sbjv: 2)	
2.62	3sg: 14 (*haber* ‘there is/are’, *comer* ‘to eat’, *parar* ‘to stop’, *estar*$_{aux}$ ‘be’, *ser* ‘to be’) 1sg: 3 (*reñir* ‘to quarrel/argue’, *coger* ‘to get/take’, *ir* ‘to go’) 1pl: 1 (*ir* ‘to go’)	6 (sbjv: 1)	prfv: 5 impf: 2 prog: 1
2.36	3sg: 8 (*poner* ‘to put’, *traer* ‘to bring/take’, *venir* ‘to come’, *estar* ‘to be’, *ser* ‘to be’) *Err*: 2		prfv: 6 impf: 1
2.22	3sg: 4 (*estar* ‘to be’, *llamarse* ‘to be called’, *marchar* ‘to leave’, *pasar* ‘to pass’)	1	prfv: 2
3.09	3sg: 21 (*estar* ‘to be’, *ser* ‘to be’, *saber* ‘to know’, *ir* ‘to go’, *haber* ‘there is/are’) 1sg: 3 (*coger* ‘to get/take’, *entender* ‘to understand’) 2sg: 1 (*ver* ‘to see’)	7	prfv: 3 impf: 1

Table C.16 Koki’s use of tense and agreement marking

MLU	Agreement	Imperatives	Tense
1.96	3sg: 46 (*ser* ‘to be’, *estar* ‘to be’, *querer* ‘to want/love’, *poder* ‘can’, *picar* ‘to sting’) 1sg: 15 (*poner* ‘can’, *decir* ‘to say’, *peinar* ‘to comb’) 2sg: 4 (*mirar* ‘to look’, *ver* ‘to see’)	24 (sbjv: 4)	prfv: 7
2.54	3sg: 46 (*estar*$_{aux}$ ‘be’, *estar* ‘to be’, *explotar* ‘to explode’, *querer* ‘to want/love’, *sacar* ‘to take out’) 1sg: 4 (*saber* ‘to know’, *poner* ‘to put’, *poder* ‘can’, *ir* ‘to go’) 2sg: 1 (*dar* ‘to give’) 3pl: 2 (*estar* ‘to be’) 1pl: 1 (*dar* ‘to give’)	9	prfv: 10 prog: 2

	INF: 1 (*arreglar* ‘to fix’)		
2.47	3SG: 46 (*estar*$_{aux}$ ‘be’, *estar* ‘to be’, *dar* ‘to give’, *venir* ‘to come’, *ser* ‘to be’) 1SG: 8 (*saber* ‘to know’, *dibujar* ‘to draw’, *ir* ‘to go’) 3PL: 5 (*hacer* ‘to do/make’, *venir* ‘to come’, *ser* ‘to be’, *salir* ‘to exit/go out’) INF: 3 (*arreglar* ‘to fix’) *Err:* 6	19[2]	PRFV: 2 PROG: 3
2.51	3SG: 49 (*estar* ‘to be’, *ser* ‘to be’, *abrir* ‘to open’, *explotar* ‘to explode’, *tener* ‘to have’) 1SG: 10 (*ir* ‘to go’, *ver* ‘to see’, *poder* ‘can’, *querer* ‘to want/love’, *estar*$_{aux}$ ‘be’) 2SG: 1 (*ver* ‘to see’) 3PL: 8 (*estar* ‘to be’, *poder* ‘can’, *ensuciar* ‘to dirty’) 1PL: 4 (*ir* ‘to go’, *cerrar* ‘to close’) INF: 1 (*tomarse* ‘to take’) *Err:* 2	18	PRFV: 4 IMPF: 3 PROG: 3
2.47	3SG: 54 (*ser* ‘to be’, *estar* ‘to be’, *hacer* ‘to do/make’, *haber* ‘there is/are’, *decir* ‘to say’) 1SG: 18 (*ir* ‘to go’, *poder* ‘can’, *estar*$_{aux}$, ‘be’, *querer* ‘to want/love’) 2SG: 1 (*ver* ‘to see’) 3PL: 10 (*estar* ‘to be’, *ser* ‘to be’, *comer* ‘to eat’) 1PL: 1 (*ir* ‘to go’)	21	PRFV: 8 IMPF: 3 PROG: 5
2.07	3SG: 16 (*estar* ‘to be’, *ser* ‘to be’, *dar* ‘to give’, *hacer* ‘to do/make’, *estar*$_{aux}$, ‘be’) 1SG: 3 (*querer* ‘to want’) 2SG: 1 (*dar* ‘to give’) 3PL: 1 (*ser* ‘to be’)	6	PROG: 1

[2]In this file, and in the following one (MLU 2.51), there are 9 and 2 instances of imperatives with the form for *usted/vos* (the formal counterpart of *tú* ‘you’), respectively. Although these are morphosyntactically tagged as subjunctive forms in the corpus (the 3SG present subjunctive is syncretic with the imperative form for *usted*), I deviate from the original transcription and treat them as imperatives, with *usted* as the addressee. An imperative analysis appears more likely for two reasons: firstly, the utterances do not present

the appropriate morphosyntactic contexts for subjunctive mood (e.g., licensing via negation, as in negative imperatives, or embedded clauses with certain predicates). Instead, they have clear imperative force (e.g., *tenga esto* 'have this'). Additionally, since Koki speaks a Mexican dialect, these examples are unlikely to be erroneous overgeneralisations of subjunctive: *usted* is used in a wider range of pragmatic contexts than in Peninsular Spanish, and, in some regions, to also signal affection (Dumitrescu, 2015). Importantly, Dumitrescu (2015, p. 7) notes that *usted* is also commonly used in child-directed speech among Central American families (indeed, a surface, automatic search with CLAN shows that 84 child-directed utterances by Koki's mother contain *usted*).

C.2.2 Germanic

German

Table C.17 Kerstin's use of tense and agreement marking

MLU	Agreement	Imperatives	Tense
1.09	3SG: 1 (*sein* 'to be')		
1.31			
1.38	3SG: 1 (*sein* 'to be')		
1.38	3SG: 1 (*sein* 'to be')		
1.58	3SG: 9 (*sein* 'to be')		
	1SG: 1 (*wollen* 'to want')		
1.36			
1.45	3SG: 3 (*sein* 'to be')		
1.46	3SG: 3 (*sein* 'to be')		
	INF: 2 (*sitzen* 'to sit')		
1.37	3SG: 3 (*sein* 'to be')		
	INF: 3 (*sitzen* 'to sit')		
1.58	3SG: 6 (*sein* 'to be')	1	
	1SG: 1 (*wollen* 'to want')		
	INF: 2 (*sitzen* 'to sit')		
1.39	3SG: 3 (*sein* 'to be')		
	INF: 6 (*sitzen* 'to sit')		
1.34	3SG: 12 (*sein* 'to be')	1	
	1SG: 1 (*nehmen* 'to take/get')		
	INF: 1 (*sitzen* 'to sit')		
1.2	3SG: 1 (*sein* 'to be')	1	
	3PL: 1 (*sein* 'to be')		
	INF: 1 (*sitzen* 'to sit')		
1.4	3SG: 7 (*sein* 'to be')	1	
	INF: 2 (*sitzen* 'to sit')		
1.12	INF: 3 (*sitzen* 'to sit')	1	
1.28	3SG: 4 (*sein* 'to be', *werden* 'to become')	3	
	INF: 4 (*sitzen* 'to sit')		
1.38	3SG: 7 (*sein* 'to be', *wollen* 'to want')	6	

	INF: 5 (*sitzen* ‘to sit’, *gehen* ‘to go’, *gucken* ‘to look’)		
1.53	3SG: 8 (*sein* ‘to be’, *wollen* ‘to want’) 1SG: 2 (*wollen* ‘to want’) INF: 7 (*sitzen* ‘to sit’, *hinsetzen* ‘to sit down’, *anziehen* ‘to put on’, *gehen* ‘to go’, *mitnehmen* ‘to take along’)	7	
1.76	3SG: 22 (*sein* ‘to be’, *gehen* ‘to go’, *aufmachen* ‘to open up’, *kommen* ‘to come’, *wollen* ‘to want’) 1SG: 3 (*wollen* ‘to want’) 3PL: 1 (*sein* ‘to be’) 1PL: 1 (*gehen* ‘to go’) INF: 3 (*fahren* ‘to drive’, *gehen* ‘to go’, *weitergehen* ‘to move/go on’)	3	
1.68	3SG: 12 (*runterfallen* ‘to fall down’, *fahren* ‘to drive’, *gehen* ‘to go’, *weinen* ‘to cry’, *sein* ‘to be’) 1SG: 3 (*wollen* ‘to want’) 2SG: 2 (*sehen* ‘to see’, *wollen* ‘to want’) 1PL: 1 (*gehen* ‘to go’) INF: 6 (*gehen* ‘to go’, *hinsetzen* ‘to sit down’, *anziehen* ‘to put on’, *draufsetzen* ‘to sit on’) *Err:* 1	10	
1.59	3SG: 9 (*weinen* ‘to cry’, *haben*$_{aux}$, *gehen* ‘to go’, *sein* ‘to be’, *haben* ‘to have’) 1SG: 4 (*wollen* ‘to want’) INF: 4 (*fahren* ‘to drive’, *aufmachen* ‘to open up’, *wiedergeben* ‘to give back’, *machen* ‘to do/make’) *Err:* 2	10	PRES PERF: 1
1.58	3SG: 10 (*sein* ‘to be’, *beissen* ‘to bite’, *machen* ‘to do/make’, *haben* ‘to have’) 1SG: 1 (*holen* ‘to fetch’) 2SG: 1 (*wollen* ‘to want’) 1PL: 1 (*nehmen* ‘to take’) INF: 4 (*trinken* ‘to drink’, *anziehen* ‘to put on’, *gehen* ‘to go’, *hinsetzen* ‘to sit down’)	18	

1.72	3SG: 18 (*sein* ‘to be’, *haben* ‘to have’, *nehmen* ‘to take’, *machen* ‘to do/make’, *gehen* ‘to go’) 2SG: 1 (*fallen* ‘to fall’) INF: 7 (*gucken* ‘to look’, *halten* ‘to hold’, *holen* ‘to fetch’, *ausziehen* ‘to take off’, *hinsetzen* ‘to sit down’)	7	
1.65	3SG: 12 (*sein* ‘to be’, *gehen* ‘to go’, *kämmen* ‘to comb’, *haben* ‘to have’, *sehen* ‘to see’) 1SG: 2 (*wollen* ‘to want’, *trinken* ‘to drink’) 2SG: 1 (*sollen* ‘should’) 1PL: 1 (*haben* ‘to have’) INF: 7 (*holen* ‘to fetch’, *sitzen* ‘to sit’, *sehen* ‘to see’, *abwischen* ‘to wipe off’, *mitnehmen* ‘to take along’)	16	
1.81	3SG: 12 (*machen* ‘to do/make’, *kommen* ‘to come’, *sein* ‘to be’, *gehen* ‘to go’, *haben* ‘to have’) 1SG: 2 (*wissen* ‘to know’, *wollen* ‘to want’) INF: 5 (*gucken* ‘to look’, *fahren* ‘to drive’, *gehen* ‘to go’) *Err:* 1	16	
1.86	3SG: 14 (*gehen* ‘to go’, *sein* ‘to be’, *haben* ‘to have’, *machen* ‘to do/make’, *gehen* ‘to go’) 1SG: 6 (*sollen* ‘should’, *zeigen* ‘to show’, *wollen* ‘to want’, *möchten* ‘would like’) 2SG: 1 (*kriegen* ‘to get’) 3PL: 1 (*wollen* ‘to want’) INF: 16 (*ausziehen* ‘to take off’, *hochziehen* ‘to pull up’, *hochstehen* ‘to stand up’, *haben* ‘to have’, *gehen* ‘to go’) *Err:* 2	20	PST: 1
1.98	3SG: 10 (*sein* ‘to be’, *schaukeln* ‘to rock’, *schwimmen* ‘to swim’, *fliegen* ‘to fly’, *gehen* ‘to go’) 3PL: 1 (*sein* ‘to be’) INF: 11 (*festhalten* ‘to hold onto’, *halten* ‘to hold’, *abküssen* ‘to kiss off’, *pfeifen* ‘to whistle’, *spielen* ‘to play’)	9	

1.67	3sg: 24 (*sein* 'to be', *klingeln* 'to ring', *sehen* 'to see', *schwimmen* 'to swim', *haben* 'to have') 1sg: 12 (*holen* 'to fetch', *haben* 'to have', *wollen* 'to want', *malen* 'to paint') 2sg: 3 (*gehen* 'to go', *müssen* 'must', *sollen* 'should') 3pl: 4 (*müssen* 'must', *machen* 'to do/make', *sein* 'to be') 1pl: 3 (*gehen* 'to go') inf: 13 (*holen* 'to fetch', *abnehmen* 'to lose weight', *machen* 'to do/make', *essen* 'to eat', *gucken* 'to look')	10	pres perf: 1 pst: 1
1.8	3sg: 8 (*sein* 'to be', *machen* 'to do/make', *schlafen* 'to sleep', *gehen* 'to go', *kommen* 'to come') 1sg: 4 (*sein* 'to be', *haben*$_{aux}$ 'have', *brauchen* 'to need') 2sg: 6 (*können* 'can', *spinnen* 'to be crazy', *wollen* 'to want') inf: 34 (*haben* 'to have', *spielen* 'to play', *putzen* 'to clean', *liegen* 'to lay', *anziehen* 'to put on')	20	pres perf: 1
1.78	3sg: 6 (*liegen* 'to lay', *sein* 'to be', *kommen* 'to come', *machen* 'to do/make', *essen* 'to eat') 1sg: 4 (*holen* 'to fetch', *machen* 'to do/make', *wollen* 'to want') 2sg: 1 (*können* 'can') inf: 11 (*gucken* 'to look', *haben* 'to have', *holen* 'to fetch', *aufpassen* 'to watch out', *trinken* 'to drink')	26	
2.25	3sg: 28 (*haben* 'to have', *sein* 'to be', *kommen* 'to come', *dürfen* 'may/to be allowed to', *können* 'can') 1sg: 5 (*haben* 'to have', *kaufen* 'to buy', *holen* 'to fetch') 2sg: 2 (*sehen* 'see') 3pl: 1 (*sein* 'to be')	42	

	inf: 10 (*pfeifen* ‘to whistle’, *holen* ‘to fetch’, *aufstehen* ‘to stand up’, *machen* ‘to do/make’, *gucken* ‘to look’)		
1.62	3sg: 17 (*sein* ‘to have’, *haben* ‘to have’, *wollen* ‘to want’, *können* ‘can’) 1sg: 8 (*wollen* ‘to want’, *zeigen* ‘to show’, *rufen* ‘to call’, *kaufen* ‘to buy’, *holen* ‘to fetch’) 2sg: 1 (*spinnen* ‘to be crazy’) inf: 19 (*putzen* ‘to clean’, *haben* ‘to have’, *runtergehen* ‘to go down’, *aufmachen* ‘to open up’, *fahren* ‘to drive’) *Err:* 1	22	
2.13	3sg: 43 (*stinken* ‘to stink’, *kommen* ‘to come’, *sein* ‘to be’, *heißen* ‘to be called’, *machen* ‘to do/make’) 1sg: 26 (*müssen* ‘must’, *schenken* ‘to give a gift’, *sein* ‘to be’, *wollen* ‘to want’, *können* ‘can’) 2sg: 1 (*sein* ‘to be’) 1pl: 1 (*brauchen* ‘to need’) inf: 21 (*ausziehen* ‘to take off’, *essen* ‘to eat’, *holen* ‘to fetch’, *zusammenzulegen* ‘to fold’, *haben* ‘to have’) *Err:* 12	41	pres perf: 9

Table C.18 Simone’s use of tense and agreement marking

MLU	Agreement	Imperatives	Tense
1.54	3sg: 6 (*gehen* ‘to go’, *sein* ‘to be’) 1sg: 1 (*sein* ‘to be’) inf: 1 (*haben* ‘to have’)		
1.62	3sg: 44 (*gehen* ‘to go’, *sein* ‘to be’, *klingeln* ‘to ring’, *essen* ‘to eat’, *schlafen* ‘to sleep’) 1sg: 5 (*mögen* ‘to like/want’)	7	pst: 2

	INF: 44 (*mitnehmen* ‘to take along’, *haben* ‘to have’, *trinken* ‘to drink’, *heilen* ‘to hail’, *machen* ‘to do/make’)		
1.69	3SG: 32 (*sein* ‘to be’, *machen* ‘to do/make’, *essen* ‘to eat’, *gehen* ‘to go’, *schlafen* ‘to sleep’) 1SG: 2 (*essen* ‘to eat’, *mögen* ‘to like/want’) 2SG: 1 (*sein* ‘to be’) INF: 11 (*haben* ‘to have’, *machen* ‘to do/make’, *waschen* ‘to wash’, *gucken* ‘to look’, *abmachen* ‘to take off’)	6	
1.71	3SG: 22 (*gehen* ‘to go’, *sein* ‘to be’, *weinen* ‘to cry’, *schlafen* ‘to sleep’, *kommen* ‘to come’) 1SG: 3 (*haben* ‘to have’, *wollen* ‘to want’) 2SG: 1 (*sein* ‘to be’) INF: 31 (*haben* ‘to have’, *umdrehen* ‘to turn around’, *backen* ‘to bake’, *saubermachen* ‘to clean up’, *aufstehen* ‘to stand up’)	4	
1.52	3SG: 2 (*sein* ‘to be’, *gehen* ‘to go’)		
1.94	3SG: 8 (*sein* ‘to be’, *schlafen* ‘to sleep’, *malen* ‘to paint’) INF: 14 (*malen* ‘to paint’, *abmachen* ‘to take off’, *haben* ‘to have’, *putzen* ‘to clean’, *waschen* ‘to wasch’)		
1.53	3SG: 11 (*gehen* ‘to go’, *sein* ‘to be’, *haben* ‘to have’, *weinen* ‘to cry’) 1SG: 3 (*wollen* ‘to want’, *brauchen* ‘to need’) INF: 32 (*holen* ‘to fetch’, *haben* ‘to have’, *fahren* ‘to drive’, *angucken* ‘to look at’, *waschen* ‘to clean’)		
1.88	3SG: 15 (*sein* ‘to be’, *gehen* ‘to go’) INF: 18 (*holen* ‘to fetch’, *haben* ‘to have’)		
2.21	3SG: 16 (*sein* ‘to be’, *schlafen* ‘to sleep’, *werden*$_{aux}$ PASS.AUX, *scheinen* ‘to seem/shine’, *gehen* ‘to go’) 3PL: 1 (*sein* ‘to be’)	7	PASS: 1

	INF: 14 (*frühstücken* ‘to have breakfast’, *bauen* ‘to build’, *nehmen* ‘to take’, *klettern* ‘to climb’, *suchen* ‘to search’)		
2.27	3SG: 48 (*gehen* ‘to go’, *bauen* ‘to build’, *sein* ‘to be’, *kleben* ‘to stick’, *sein*$_{aux}$ ‘be’)	2	PRES PERF: 4
	3PL: 1 (*sein* ‘to be’)		
	1PL: 1 (*bauen* ‘to build’)		
	INF: 33 (*reinlegen* ‘to trick’, *hierbleiben* ‘to stay here’, *haben* ‘to have’, *reinstecken* ‘to put in’, *machen* ‘to do/make’)		
2.28	3SG: 20 (*gehen* ‘to go’, *schlafen* ‘to sleep’, *sein* ‘to be’, *kommen* ‘to come’, *weinen* ‘to cry’)	12	
	1SG: 3 (*wollen* ‘to want’, *mögen* ‘to like/want’)		
	INF: 35 (*holen* ‘to fetch’, *fangen* ‘to catch’, *ziehen* ‘to pull’, *ausziehen* ‘to take off’, *haben* ‘to have’)		
2.31	3SG: 19 (*sein* ‘to be’, *haben* ‘to have’, *riechen* ‘to smell’, *gehen* ‘to go’, *kommen* ‘to come’)	5	
	INF: 14 (*aufwachen* ‘to wake up’, *haben* ‘to have’, *außchückeln* ‘to chill out’, *backen* ‘to bake’, *setzen* ‘to set/put’)		
2.31	3SG: 37 (*haben* ‘to have’, *sein* ‘to be’, *gehen* ‘to go’, *machen* ‘to do/make’, *schmecken* ‘to taste (good)’)	2	
	1SG: 3 (*brauchen* ‘to need’, *wollen* ‘to want’, *wissen* ‘to know’)		
	2SG: 2 (*sein* ‘to be’, *dürfen* ‘may/to be allowed to’)		
	3PL: 3 (*sein* ‘to be’)		
	INF: 13 (*haben* ‘to have’, *spielen* ‘to play’, *ziehen* ‘to pull’, *machen* ‘to do/make’, *putzen* ‘to clean’)		
1.98	3SG: 43 (*gehen* ‘to go’, *sein* ‘to be’, *kleben* ‘to stick’, *müssen* ‘must’, *machen* ‘to do/make’)	15	
	1SG: 2 (*wollen* ‘to want’, *möchten* ‘would like’)		
	2SG: 2 (*wollen* ‘to want’)		
	3PL: 4 (*sein* ‘to be’, *kommen* ‘to come’)		

	INF: 33 (*machen* ‘to do/make’, *haben* ‘to have’, *hinlegen* ‘to lie down’, *aufmachen* ‘to open up’, *abmachen* ‘to take off’)		
1.93	3SG: 33 (*gehen* ‘to go’, *spucken* ‘to spit’, *gehoren* ‘to belong’, *sein* ‘to be’, *machen* ‘to do/make’) 1SG: 3 (*wollen* ‘to want’, *machen* ‘to do/make’, *müssen* ‘must’) INF: 18 (*gehen* ‘to go’, *aufmachen* ‘to open up’, *haben* ‘to have’, *hinsetzen* ‘to sit down’, *machen* ‘to do/make’)	17	PRES PERF: 2 PST: 1
1.63	3SG: 27 (*machen* ‘to do/make’, *sein* ‘to be’, *gehen* ‘to do’, *kriegen* ‘to get’, *sein*$_{aux}$ ‘be’) 1SG: 2 (*wollen* ‘to want’) INF: 27 (*machen* ‘to do/make’, *haben* ‘to have’, *aufeßen* ‘to eat up’, *aufstehen* ‘to stand up’, *aufmachen* ‘to open up’) *Err*: 2	13	PRES PERF: 2
1.72	3SG: 27 (*gehen* ‘to go’, *sein* ‘to be’, *müssen* ‘must’, *machen* ‘to do/make’, *kommen* ‘to come’) 1SG: 2 (*haben*$_{aux}$ ‘have’, *möchten* ‘would like’) INF: 14 (*abmachen* ‘to take off’, *schneiden* ‘to cut’, *haben* ‘to have’, *abwaschen* ‘to wash off’, *aufmachen* ‘to open up’)	14	PRES PERF: 2
1.78	3SG: 35 (*sein* ‘to be’, *machen* ‘to do/make’, *fallen* ‘to fall’, *fangen* ‘to catch’, *kriegen* ‘to get’) 1SG: 2 (*brauchen* ‘to need’, *wollen* ‘to want’) 1PL: 1 (*malen* ‘to paint’) INF: 25 (*essen* ‘to eat’, *zumachen* ‘to close’, *haben* ‘to have’, *spielen* ‘to play’, *messen* ‘to measure’) *Err*: 1	14	PRES PERF: 1
2.03	3SG: 9 (*sein* ‘to be’, *machen* ‘to do/make’, *weinen* ‘to cry’, *kommen* ‘to come’, *gucken* ‘to look’) 1SG: 1 (*wollen* ‘to want’) INF: 3 (*haben* ‘to have’, *spritzen* ‘to splash’)		
1.79	3SG: 36 (*halten* ‘to hold’, *sehen* ‘to see’, *sein* ‘to be’, *machen* ‘to do/make’, *gehen* ‘to go’)	11	

	1SG: 2 (*kriegen* ‘to get’, *müssen* ‘must’) INF: 20 (*hinlegen* ‘to lie down’, *haben* ‘to have’, *trinken* ‘to drink’, *zumachen* ‘to close’, *spielen* ‘to play’)		
1.71	3SG: 29 (*essen* ‘to eat’, *haben*$_{aux}$ ‘have’, *kleben* ‘to stick’, *sein* ‘to be’, *gehen* ‘to go’) 1SG: 2 (*haben* ‘to have’) 3PL: 2 (*sein* ‘to be’) INF: 12 (*ausmachen* ‘to turn off’, *essen* ‘to eat’, *anmachen* ‘to turn on’, *reinstecken* ‘to put in’, *drehen* ‘to turn’) *Err:* 1	12	PRES PERF: 1
1.94	3SG: 65 (*laufen* ‘to run’, *haben* ‘to have’, *sein* ‘to be’, *fallen* ‘to fall’, *weinen* ‘to cry’) 1SG: 3 (*glauben* ‘to believe’, *haben*$_{aux}$ ‘have’, *wollen* ‘to want’) 3PL: 3 (*sein* ‘to be’) INF: 14 (*abmachen* ‘to take off’, *sehen* ‘to see’, *fliegenlaßen* ‘to let fly’, *totmachen* ‘to kill’, *machen* ‘to do/make’) *Err:* 1	20	PRES PERF: 3
1.66	3SG: 15 (*sein* ‘to be’, *gehen* ‘to go’, *haben* ‘to have’, *haben*$_{aux}$ ‘have’, *machen* ‘to do/make’) 1SG: 3 (*wissen* ‘to know’, *haben* ‘to have’, *wollen* ‘to want’) INF: 16 (*machen* ‘to do/make’, *spielen* ‘to play’, *aufpaßen* ‘to watch out’, *umkippen* ‘to tip over’, *haben* ‘to have’)	6	PRES PERF: 1 PST: 2
2.22	3SG: 23 (*sein* ‘to be’, *können* ‘can’, *gehen* ‘to go’, *wollen* ‘to want’, *machen* ‘to do/make’) 1SG: 1 (*sein* ‘to be’) INF: 4 (*gehen* ‘to go’, *gucken* ‘to look’, *ausmachen* ‘to turn off’, *machen* ‘to do/make’) *Err:* 1	11	PRES PERF: 1 PST: 1
2.0	3SG: 3 (*sein* ‘to be’)	1	

	INF: 2 (*hinsetzen* 'to sit down', *rausholen* 'to get out')		
2.09	3SG: 92 (*sein* 'to be', *machen* 'to do', *gehen* 'to go', *sein*$_{aux}$ 'be', *festhalten* 'to hold onto') 1SG: 19 (*müssen* 'must', *kriegen* 'to get', *wollen* 'to want', *gehen* 'to go') 2SG: 7 (*müssen* 'must', *dürfen* 'may/to be allowed to', *haben* 'to have') 3PL: 1 (*sein* 'to be') INF: 47 (*putzen* 'to clean', *aufmachen* 'to open up', *haben* 'to have', *ausziehen* 'to take off', *aufstehen* 'to stand up')	29	PRES PERF: 3 PST: 1
1.99	3SG: 73 (*sein*$_{aux}$ 'be', *sollen* 'should', *sehen* 'to see', *machen* 'to do/make', *haben*$_{aux}$ 'have') 1SG: 9 (*wollen* 'to want', *haben* 'to have') 2SG: 4 (*dürfen* 'may/to be allowed to') 3PL: 2 (*sein* 'to be') INF: 29 (*hinstellen* 'to put up', *schaukeln* 'to rock', *anstoßen* 'to bumb', *schubsen* 'to push', *fliegenlaßen* 'to let fly')	14	PRES PERF: 5 PST PERF: 1
1.82	3SG: 52 (*halten* 'to hold', *kommen* 'to come', *sein* 'to be', *aufziehen* 'to raise', *schlafen* 'to sleep') 1SG: 27 (*brauchen* 'to need', *machen* 'to do/make', *wollen* 'to want', *aufpusten* 'to blow up', *essen* 'to eat') 2SG: 1 (*sollen* 'should') 3PL: 3 (*sein* 'to be') INF: 17 (*kaputtmachen* 'to break', *umschmeißen* 'to knock over', *baden* 'to bathe', *waschen* 'to wash', *festhalten* 'to hold onto')	13	PST: 1
1.89	3SG: 44 (*machen* 'to do/make', *sein* 'to be', *gehen* 'to go', *können* 'can', *sein*$_{aux}$ 'be') 1SG: 20 (*wollen* 'to want', *paßen* 'to fit', *machen* 'to do/make', *zeigen* 'to show', *brauchen* 'to need') 3PL: 4 (*sein* 'to be')	22	PRES PERF: 2 PST: 1

	INF: 5 (*haben* 'to have', *abmachen* 'to take off', *drehen* 'to turn', *zumachen* 'to close', *machen* 'to do/make')		
1.96	3SG: 30 (*kommen* 'to come', *sein* 'to be', *gehen* 'to go', *schmecken* 'to taste (good)', *können* 'can') 1SG: 18 (*machen* 'to do/make', *wollen* 'to want', *mögen* 'to like/want', *holen* 'to fetch', *essen* 'to eat') 2SG: 1 (*können* 'can') 3PL: 5 (*sein* 'to be', *mögen* 'to like/want') INF: 13 (*gehen* 'to go', *haben* 'to have', *wehtun* 'to hurt', *auskippen* 'to pour out/empty', *machen* 'to do/make')	5	PRES PERF: 1 PST: 2
1.92	3SG: 61 (*sein* 'to be', *machen* 'to do/make', *laufen* 'to run', *kommen* 'to come', *klingeln* 'to ring') 1SG: 13 (*helfen* 'to help', *reinschütten* 'to pour in', *kriegen* 'to get', *bauen* 'to build', *wollen* 'to want') 3PL: 4 (*sein* 'to be') INF: 10 (*runterfallen* 'to fall down', *machen* 'to do/make', *bauen* 'to build', *kaputtmachen* 'to break', *hinsetzen* 'to sit down') *Err:* 2	12	PRES PERF: 5
2.52	3SG: 72 (*sein* 'to be', *wollen* 'to want', *kommen* 'to come', *gehen* 'to go', *stimmen* 'to be right') 1SG: 18 (*wollen* 'to want', *holen* 'to fetch', *angeln* 'to fish', *machen* 'to do/make', *kriegen* 'to get') 2SG: 4 (*trinken* 'to drink', *kriegen* 'to get', *dürfen* 'may/to be allowed to', *machen* 'to do/make') 3PL: 6 (*sein* 'to be', *machen* 'to do/make', *gucken* 'to look') INF: 13 (*haben* 'to have', *abmachen* 'to take off', *suchen* 'to search', *kaputtmachen* 'to break', *shaukeln* 'to rock') *Err:* 2	11	PRES PERF: 2 PST: 2

2.35	3sg: 19 (*dürfen* ‘may/to be allowed to’, *gehen* ‘to go’, *kommen* ‘to come’, *machen* ‘to do/make’, *kriegen* ‘to get’) 1sg: 4 (*wollen* ‘to want’, *haben* ‘to have’, *holen* ‘to fetch’) 2sg: 1 (*sein* ‘to be’) inf: 8 (*haben* ‘to have’, *knacken* ‘to crack’, *anpacken* ‘to handle’)	11	
2.62	3sg: 48 (*sein* ‘to be’, *können* ‘can’, *gehen* ‘to go’, *werden* ‘to become’, *wollen* ‘to want’) 1sg: 4 (*wollen* ‘to want’, *holen* ‘to fetch’, *gucken* ‘to look’) 2sg: 3 (*kriegen* ‘to get’) 3pl: 4 (*wollen* ‘to want’, *sein* ‘to be’, *machen* ‘to do/make’) inf: 8 (*reinsteigen* ‘to get in’, *schlafen* ‘to sleep’, *klettern* ‘to climb’, *umschmeißen* ‘to knock over’, *bringen* ‘to bring’)	17	pres perf: 5 pst: 1
2.67	3sg: 12 (*sein*$_{aux}$ ‘be’, *sein* ‘to be’, *gehen* ‘to go’, *können* ‘can’, *wachen* ‘to wake up’) 1sg: 1 (*machen* ‘to do/make’) 3pl: 4 (*sein* ‘to be’) inf: 1 (*trinken* ‘to drink’)	6	pres perf: 2
3.35	3sg: 42 (*können* ‘can’, *gehen* ‘to go’, *sein* ‘to be’, *machen* ‘to do/make’, *dürfen* ‘may/to be allowed to’) 1sg: 6 (*wollen* ‘to want’, *haben* ‘to have’, *fallen* ‘to fall’) 2sg: 3 (*haben* ‘to have’, *kennen* ‘to know’, *sein* ‘to be’) 3pl: 4 (*fliegen* ‘to fly’, *sein*$_{aux}$ ‘be’, *haben* ‘to have’, *sein* ‘to be’) 1pl: 1 (*tanzen* ‘to dance’) inf: 7 (*angucken* ‘to look at’, *aufmachen* ‘to open up’, *verkaufen* ‘to sell’, *waschen* ‘to wash’, *haben* ‘to have’)	14	pres perf: 5

	Err: 2		
4.04	3SG: 37 (*haben*$_{aux}$ 'have', *haben* 'to have', *halten* 'to hold', *kommen* 'to come', *sein* 'to be') 1SG: 10 (*machen* 'to do/make', *haben* 'to have', *bauen* 'to build', *haben*$_{aux}$ 'have', *wollen* 'to want') 2SG: 4 (*müssen* 'must', *sein* 'to be', *dürfen* 'may/to be allowed to') 3PL: 5 (*sollen* 'should', *wollen* 'to want', *sein* 'to be') INF: 14 (*stehen* 'to stand', *haben* 'to have', *bauen* 'to build', *hauen* 'to hit', *essen* 'to eat')	7	PRES PERF: 3 PST: 2

Dutch

Table C.19 Josse's use of tense and agreement marking

MLU	Agreement	Imperatives	Tense
1.2	1SG: 1 (*kunnen* 'can') INF: 29 (*kijken* 'to look', *vallen* 'to fall', *hoesten* 'to cough', *doen* 'to do', *blazen* 'to blow')	2	
1.44	3SG: 1 (*kloppen* 'to hit') 1SG: 2 (*mogen* 'may/to be allowed to', *kunnen* 'can') INF: 8 (*drinken* 'to drink', *schrijven* 'to write', *pakken* 'to take', *spelen* 'to play', *hebben* 'to have')	3	
1.55	3SG: 10 (*doen* 'to do', *moeten* 'must', *gaan* 'to go', *kommen* 'to come') 1SG: 5 (*kunnen* 'can') INF: 19 (*kijken* 'to look', *geven* 'to give', *kussen* 'to kiss', *zitten* 'to sit', *blazen* 'to blow')	3	
1.59	3SG: 7 (*gaan* 'to go', *kunnen* 'can', *mogen* 'may/to be allowed to', *doen* 'to do')		

	1SG: 2 (*kunnen* 'can', *willen* 'to want') INF: 20 (*drinken* 'to drink', *bellen* 'to call', *roeren* 'to stir', *halen* 'to fetch', *regenen* 'to rain')		
1.69	3SG: 11 (*gaan* 'to go', *doen* 'to do', *zijn* 'to be', *mogen* 'may/to be allowed to', *slapen* 'to sleep') 1SG: 3 (*kunnen* 'can', *hebben* 'to have') INF: 9 (*neerzetten* 'to put down', *doen* 'to do', *vallen* 'to fall', *schrijven* 'to sit', *kopen* 'to buy')	1	PST: 1
1.74	3SG: 7 (*zijn* 'to be', *gaan* 'to do', *kommen* 'to come') 1SG: 2 (*willen* 'to want', *kunnen* 'can') 2SG: 1 (*moeten* 'must') INF: 31 (*verkouden* 'to catch a cold', *roeren* 'to stir', *doen* 'to do', *kijken* 'to look', *krijgen* 'to get')	5	
1.94	3SG: 24 (*zijn* 'to be', *kunnen* 'can', *gaan* 'to go', *doen* 'to do', *moeten* 'must') 1SG: 10 (*willen* 'to want', *gaan* 'to go', *kunnen* 'can', *doen* 'to do', *moeten* 'must') INF: 11 (*drinken* 'to drink', *zitten* 'to sit', *afscheuren* 'to tear off', *tekenen* 'to draw', *schrijven* 'to write')	13	
2.14	3SG: 8 (*zijn* 'to be', *zijn*$_{aux}$ 'be', *moeten* 'must', *mogen* 'may/to be allowed to') 1SG: 7 (*willen* 'to want', *kunnen* 'can') 2SG: 6 (*doen* 'to do', *hebben* 'to have', *gaan* 'to go', *mogen* 'may/to be allowed to', *moeten* 'must') 1PL: 1 (*gaan* 'to go') INF: 16 (*doen* 'to do', *drinken* 'to drink', *zitten* 'to sit', *maken* 'to do/make', *pakken* 'to take')	2	
1.9	3SG: 17 (*zijn* 'to be', *gaan* 'to go', *zullen* 'should', *staan* 'to stand') 1SG: 2 (*willen* 'to want') 1PL: 2 (*gaan* 'to go') INF: 9 (*roeren* 'to stir', *werken* 'to work', *kijken* 'to look', *slapen* 'to sleep', *vallen* 'to fall')	6	
1.59	3SG: 5 (*moeten* 'must', *zijn* 'to be', *gaan* 'to go')	6	

	INF: 6 (*maken* ‘to do/make’, *springen* ‘to jump’, *kijken* ‘to look’, *trekken* ‘to pull’, *doen* ‘to do’)		
2.17	3SG: 29 (*moeten* ‘must’, *zijn* ‘to be’, $zijn_{aux}$ ‘be’, *zitten* ‘to sit’) 1SG: 7 (*mogen* ‘may/to be allowed to’, *willen* ‘to want’, *kunnen* ‘to want’) 3PL: 2 (*zijn* ‘to be’) INF: 16 (*doen* ‘to do’, *laten* ‘to let/leave’, *zien* ‘to see’, *eten* ‘to eat’, *kijken* ‘to look’)	6	PRES PERF: 1
2.11	3SG: 33 (*zijn* ‘to be’, *gaan* ‘to go’, *doen* ‘to do’, *kunnen* ‘can’, *moeten* ‘must’) 1SG: 13 (*kunnen* ‘can’, *willen* ‘to want’, *moeten* ‘must’, *hebben* ‘to have’, *doen* ‘to do’) 2SG: 6 (*gaan* ‘to go’, *zijn* ‘to see’, *doen* ‘to do’, *moeten* ‘must’, *durven* ‘to dare’) 1PL: 1 (*gaan* ‘to go’) INF: 11 (*uitspugen* ‘to spit out’, *rijden* ‘to ride’, *maken* ‘to do/make’, *doen* ‘to do’, *vliegen* ‘to fly’)	6	PST: 4
2.19	3SG: 30 (*zijn* ‘to be’, *gaan* ‘to go’, *kunnen* ‘can’, *zitten* ‘to sit’, *moeten* ‘must’, *vallen* ‘to fall’) 1SG: 27 (*willen* ‘to want’, *gaan* ‘to go’, *mogen* ‘may/to be allowed to’, *weten* ‘to know’, *hebben* ‘to have’) 2SG: 11 (*hebben* ‘to have’, *vallen* ‘to fall’, *moeten* ‘must’, *kunnen* ‘can’, *zeggen* ‘to say’) 3PL: 2 (*doen* ‘to do’, *zijn* ‘to be’) 1PL: 3 (*gaan* ‘to go’, *doen* ‘to do’) INF: 11 (*zitten* ‘to sit’, *rijden* ‘to ride’, *blijven* ‘to remain/stay’, *schrijven* ‘to write’, *maken* ‘to do/make’) *Err:* 1	11	PRES PERF: 1
2.46	3SG: 93 (*zijn* ‘to be’, *moeten* ‘must’, $zijn_{aux}$ ‘be’, *staan* ‘to stand’, *kunnen* ‘can’, *kommen* ‘to come’) 1SG: 14 (*gaan* ‘to go’, *kunnen* ‘can’, *moeten* ‘must’, *hebben* ‘to have’, *wetten* ‘to know’)	13	PRES PERF: 2 PST: 1

	2SG: 8 (*moeten* 'must', *kunnen* 'can', *gaan* 'to go', *mogen* 'may/to be allowed to') 3PL: 2 (*zijn* 'to be') 1PL: 4 (*moeten* 'must', *gaan* 'to go') INF: 17 (*kijken* 'to see', *doen* 'to do', *schrijven* 'to write', *kauwen* 'to chew', *eten* 'to eat') *Err:* 2		
2.25	3SG: 47 (*moeten* 'must', *zijn* 'to be', *vallen* 'to fall', *gaan* 'to go', *kommen* 'to come', *kunnen* 'can') 1SG: 26 (*willen* 'to want', *kunnen* 'can', *gaan* 'to go', *zitten* 'to sit', *mogen* 'may/to be allowed to') 2SG: 19 (*moeten* 'must', *kunnen* 'can', *kommen* 'to come', *mogen* 'may/to be allowed to', *zijn* 'to be') 1PL: 6 (*gaan* 'to go', *maken* 'to do/make') INF: 20 (*zitten* 'to sit', *bouwen* 'to build', *uitstappen* 'to get off', *vallen* 'to fall', *maken* 'to do/make') *Err:* 1	6	PRES PERF: 1
2.55	3SG: 68 (*moeten* 'must', *zijn* 'to be', *hebben* 'to have', *staan* 'to stand', *zitten* 'to sit', *gaan* 'to go') 1SG: 21 (*willen* 'to want', *moeten* 'must', *mogen* 'may/to be allowed to', *zullen* 'should', *hebben* 'to have') 2SG: 8 (*moeten* 'must', *kommen* 'to come', *mogen* 'may/to be allowed to', *zeggen* 'to say') 3PL: 1 (*zijn* 'to be') 1PL: 2 (*gaan* 'to go') INF: 5 (*huilen* 'to cry', *blazen* 'to blow', *springen* 'to jump', *duiken* 'to dive') *Err:* 2	13	PRES PERF: 4
2.42	3SG: 66 (*zijn* 'to be', *moeten* 'must', *kijken* 'to look', *zitten* 'to sit', *hebben* 'to have', *kennen* 'to know')	1	PRES PERF: 2 PST: 6

	1sg: 16 (*hetten* ‘to be called’, *mogen* ‘may/to be called’, *hebben*$_{aux}$ ‘have’, *zullen* ‘should’, *kunnen* ‘can’) 2sg: 8 (*hebben*$_{aux}$ ‘have’, *zijn* ‘to be’, *gaan* ‘to go’, *moeten* ‘must’, *zeggen* ‘to say’) 3pl: 7 (*zijn* ‘to be’, *vallen* ‘to fall’) inf: 2 (*bouwen* ‘to build’, *zitten* ‘to sit’)		
2.8	3sg: 48 (*gaan* ‘to go’, *zijn* ‘to be’, *moeten* ‘must’, *kunnen* ‘can’, *doen* ‘to do’, *kommen* ‘to come’) 1sg: 9 (*gaan* ‘to go’, *maken* ‘to do/make’, *hebben*$_{aux}$ ‘have’, *wetten* ‘to know’, *mogen* ‘may/to be allowed to’) 2sg: 20 (*moeten* ‘must’, *zien* ‘to see’, *mogen* ‘may/to be allowed to’, *doen* ‘to do’, *hebben*$_{aux}$ ‘have’) 3pl: 5 (*moeten* ‘must’, *maken* ‘to do/make’, *zijn* ‘to be’) 1pl: 3 (*gaan* ‘to go’, *moeten* ‘must’, *maken* ‘to do/make’) inf: 2 (*zitten* ‘to sit’, *maken* ‘to do/make’) *Err:* 1	13	pres perf: 5 pst: 1

Table C.20 Sarah’s use of tense and agreement marking

MLU	Agreement	Imperatives	Tense
1.12	3sg: 2 (*moeten* ‘must’) inf: 7 (*kietelen* ‘to tickle’, *zitten* ‘to sit’, *vallen* ‘to fall’)		
1.17	3sg: 1 (*zijn* ‘to be’) inf: 12 (*keuken* ‘to cook’, *zitten* ‘to sit’, *lezen* ‘to read’, *hebben* ‘to have’, *rijden* ‘to ride’)	1	
1.07	3sg: 1 (*zijn* ‘to be’) 1sg: 1 (*zijn* ‘to be’)		

	INF: 12 (*fietsen* ‘to cycle’, *tekenen* ‘to draw’, *zitten* ‘to sit’, *eten* ‘to eat’, *kijken* ‘to look’)		
1.17	3SG: 1 (*kunnen* ‘can’) 1SG: 4 (*willen* ‘to want’) INF: 7 (*fietsen* ‘to cycle’, *kietelen* ‘to tikle’, *zitten* ‘to sit’, *drinken* ‘to drink’)	2	
1.09	3SG: 1 (*zijn* ‘to be’) 1SG: 4 (*willen* ‘to want’) INF: 10 (*zitten* ‘to sit’, *fietsen* ‘to cycle’, *huilen* ‘to cry’, *poetsen* ‘to polish’, *springen* ‘to jump’)		
1.17	3SG: 1 (*zijn* ‘to be’) 1SG: 1 (*moeten* ‘must’) INF: 3 (*eten* ‘to eat’, *kijken* ‘to look’, *huilen* ‘to cry’) *Err:* 1		
1.25	3SG: 1 (*zijn* ‘to be’) 1SG: 1 (*moeten* ‘must’) INF: 23 (*pakken* ‘to take’, *huilen* ‘to cry’, *hebben* ‘to have’, *liggen* ‘to lie’, *fietsen* ‘to cycle’)		
1.37	3SG: 4 (*zijn* ‘to be’) 1SG: 1 (*willen* ‘to want’) 2SG: 1 (*doen* ‘to do’) 3PL: 1 (*hebben* ‘to have’) INF: 29 (*draaien* ‘to turn’, *lopen* ‘to walk’, *zitten* ‘to sit’, *huilen* ‘to cry’, *eten* ‘to eat’)	2	
1.68	3SG: 8 (*zijn* ‘to be’, *mogen* ‘may/to be allowed to’, *kunnen* ‘can’, *gaan* ‘to go’) 1SG: 10 (*moeten* ‘must’, *kunnen* ‘can’, *hoeven* ‘need’, *weten* ‘to know’) INF: 27 (*zitten* ‘to sit’, *huilen* ‘to cry’, *tekenen* ‘to draw’, *lachen* ‘to laugh’, *liggen* ‘to lie’)	1	
1.88	3SG: 8 (*zijn*$_{aux}$ ‘be’, *zijn* ‘to be’, *moeten* ‘must’, *gaan* ‘to go’, *mogen* ‘may/to be allowed to’) 1SG: 18 (*willen* ‘to want’, *laten* ‘to let/leave’, *mogen* ‘may/to be allowed to’, *gaan* ‘to go’, *hebben* ‘to have’)		PRES PERF: 1

	inf: 17 (*liggen* ‘to lie’, *kleuren* ‘to paint’, *helpen* ‘to help’, *hebben* ‘to have’, *slapen* ‘to sleep’)		
2.11	3sg: 38 (*zijn* ‘to be’, *willen* ‘to want’, *mogen* ‘may/to be allowed to’, *hoeven* ‘to need’) 1sg: 12 (*willen* ‘to want’, *weten* ‘to know’, *kunnen* ‘can’, *hebben* ‘to have’) 2sg: 7 (*hoeven* ‘to need’, *zien* ‘to see’, *hebben* ‘to have’, *moeten* ‘must’) inf: 16 (*ophouden* ‘to hold up’, *omgooien* ‘to knock over’, *fietsen* ‘to cycle’, *geven* ‘to give’, *lezen* ‘to read’) *Err:* 2	4	
2.05	3sg: 16 (*zijn* ‘to be’, *doen* ‘to do’, *zitten* ‘to sit’, *mogen* ‘may/to be allowed to’, *regenen* ‘to rain’) 1sg: 14 (*hebben* ‘to have’, *hebben*$_{aux}$ ‘have’, *zijn* ‘to be’, *willen* ‘to want’, *mogen* ‘may/to be allowed to’, *moeten* ‘must’) 2sg: 2 (*hoeven* ‘to need’, *mogen* ‘may/to be allowed to’) 3pl: 1 (*moeten* ‘must’) inf: 11 (*zitten* ‘to sit’, *waaien* ‘to blow’, *wrijven* ‘to rub’, *eten* ‘to eat’, *drinken* ‘to drink’) *Err:* 1	1	pres perf: 2 pst: 1
2.53	3sg: 42 (*zijn* ‘to be’, *kommen* ‘to come’, *gaan* ‘to go’, *vallen* ‘to fall’, *zitten* ‘to sit’) 1sg: 29 (*hebben* ‘to have’, *willen* ‘to want’, *hooren* ‘to hear’, *vinden* ‘to find’, *wetten* ‘to know’) 2sg: 19 (*hebben*$_{aux}$ ‘have’, *gaan* ‘to go’, *aanzetten* ‘to turn on’, *willen* ‘to want’) 3pl: 3 (*zijn* ‘to be’) inf: 12 (*vallen* ‘to fall’, *kijken* ‘to look’, *zwemmen* ‘to swim’, *tandenpoetsen* ‘to toothbrushing’, *schoonmaken* ‘to clean’)	19	pres perf: 6 pst perf: 1
2.34	3sg: 24 (*zijn* ‘to be’, *moeten* ‘must’, *mogen* ‘may/to be allowed to’, *gaan* ‘to go’)	3	pres perf: 3

	1SG: 36 (*gaan* 'to go', *vinden* 'to finden', *willen* 'to want', *hebben* 'to have', *mogen* 'may/to be allowed to') 2SG: 2 (*moeten* 'must') INF: 10 (*eten* 'to eat', *huilen* 'to cry', *kijken* 'to look', *wassen* 'to wash', *opzoeken* 'to look up')		
2.46	3SG: 40 (*zijn$_{aux}$* 'be', *doen* 'to do', *gaan* 'to do', *zijn* 'to be', *hoeven* 'to need') 1SG: 22 (*hebben$_{aux}$* 'have', *hebben* 'to have', *vinden* 'to find', *gaan* 'to go', *willen* 'to want') 2SG: 3 (*moeten* 'must', *vinden* 'to find', *hebben* 'to have') 3PL: 4 (*gaan* 'to go', *hebben* 'to have') INF: 24 (*kommen* 'to come', *lezen* 'to read', *drinken* 'to drink', *zitten* 'to sit', *gaan* 'to go') *Err:* 10	3	PRES PERF: 10
2.47	3SG: 21 (*zijn* 'to be', *zitten* 'to sit', *kommen* 'to come', *doen* 'to do', *omvallen* 'to fall over') 1SG: 45 (*gaan* 'to go', *worden* 'to become', *zitten* 'to sit', *hebben$_{aux}$* 'have', *hebben* 'to have') 2SG: 4 (*moeten* 'must', *gaan* 'to go', *zien* 'to see') INF: 17 (*afblijven* 'to keep off', *doen* 'to do', *hebben* 'to have', *make* 'to do/make', *omgooien* 'to knock over') *Err:* 2		PRES PERF: 6
2.59	3SG: 20 (*zijn* 'to be', *zitten* 'to sit') 1SG: 55 (*zijn* 'to be', *willen* 'to want', *krijgen* 'to get', *hebben* 'to have', *kunnen* 'can') 2SG: 20 (*gaan* 'to go', *zijn* 'to be', *moeten* 'must', *kunnen* 'can', *mogen* 'may/to be allowed to') INF: 6 (*doen* 'to do', *zoeken* 'to search', *halen* 'to fetch', *vallen* 'to fall')		
2.74	3SG: 45 (*zijn* 'to be', *moeten* 'must', *horen* 'to hear', *kunnen* 'can') 1SG: 28 (*kunnen* 'can', *hebben$_{aux}$* 'have', *zien* 'to see', *gaan* 'to go', *moeten* 'must')	2	PRES PERF: 2

	2sg: 11 (*vinden* ‘to find’, *moeten* ‘must’, *helpen* ‘to help’, *gaan* ‘to go’, *mogen* ‘may/to be allowed to’) 3pl: 1 (*zijn* ‘to be’) 1pl: 1 (*hebben* ‘to have’) inf: 11 (*maken* ‘to do/make’, *wachten* ‘to wait’, *kijken* ‘to look’, *helpen* ‘to help’) *Err*: 1	
2.45	3sg: 31 (*zijn* ‘to be’, *moeten* ‘must’, *willen* ‘to want’, *zijn*$_{aux}$ ‘be’, *horen* ‘to hear’) 1sg: 21 (*willen* ‘to want’, *gaan* ‘to go’, *kunnen* ‘can’, *worden* ‘to become’, *hebben*$_{aux}$ ‘be’) 2sg: 5 (*hebben*$_{aux}$ ‘be’, *doen* ‘to do’, *gaan* ‘to go’, *liggen* ‘to lie’) 3pl: 1 (*zijn* ‘to be’) 1pl: 1 (*gaan* ‘to go’) inf: 8 (*schillen* ‘to peel’, *leggen* ‘to lay’, *huilen* ‘to cry’, *opeten* ‘to eat up’) *Err*: 3	pres perf: 8
2.8	3sg: 27 (*zijn* ‘to be’, *zitten* ‘to seat’, *doen* ‘to do’, *horen* ‘to hear’, *passen* ‘to fit’) 1sg: 42 (*zitten* ‘to sit’, *gaan* ‘to go’, *bewaren* ‘to store’, *moeten* ‘must’, *willen* ‘to want’) inf: 6 (*zitten* ‘to sit’, *doen* ‘to do’, *opruimen* ‘to tidy up’, *maken* ‘to do/make’) *Err*: 1	pres perf: 3 pst: 1
2.51	3sg: 29 (*zijn* ‘to be’, *hebben* ‘to have’, *passen* ‘to fit’, *horen* ‘to hear’, *moeten* ‘must’) 1sg: 28 (*willen* ‘to want’, *zitten* ‘to sit’, *hebben*$_{aux}$ ‘have’, *gaan* ‘go’, *vinden* ‘to find’) 2sg: 8 (*hebben* ‘to have’, *moeten* ‘must’, *mogen* ‘may/to be allowed to’, *zijn* ‘to be’) inf: 6 (*meehelpen* ‘to help’, *puzzelen* ‘to puzzle’, *maken* ‘to do/make’, *voorlezen* ‘to read aloud’)	pres perf: 5 pst perf: 1 pst: 1
2.66	3sg: 51 (*moeten* ‘must’, *passen* ‘to fit’, *zitten* ‘to sit’, *zijn* ‘to be’, *hetten* ‘to be called’)	pres perf: 5 pst: 3

	1sg: 35 (*hoeven* 'to need', *moeten* 'must', *knoeien* 'to mess', *hebben*$_{aux}$ 'have', *mogen* 'may/to be allowed to') 2sg: 17 (*hebben*$_{aux}$ 'have', *gaan* 'to go', *willen* 'to want', *moeten* 'must', *hebben* 'to have') 3pl: 2 (*zijn* 'to be', *moeten* 'must') 1pl: 3 (*worden* 'to become', *mogen* 'may/to be allowed to', *zijn* 'to be') inf: 12 (*drinken* 'to drink', *proeven* 'to taste', *maken* 'to do/make', *schommelen* 'to rock', *klimmen* 'to climb') *Err:* 1	pass: 1
2.97	3sg: 49 (*moeten* 'must', *zijn* 'to be', *hebben* 'to have', *gaan* 'to go', *zitten* 'to sit') 1sg: 45 (*hebben* 'to have', *zeggen* 'to say', *gaan* 'to go', *zetten* 'to set/put', *zijn* 'to be') 2sg: 17 (*doen* 'to do', *hebben* 'to have', *verstaan* 'to understand', *willen* 'to want', *leggen* 'to lay') 3pl: 7 (*zijn* 'to be', *gaan* 'to go') 1pl: 4 (*zitten* 'to sit', *nemen* 'to take', *gaan* 'to go', *hebben*$_{aux}$ 'have') inf: 11 (*piepen* 'to beep', *puzzelen* 'to puzzle', *eten* 'to eat', *maken* 'to do/make', *huilen* 'to cry') *Err:* 3	pres perf: 3 pst perf: 1 pst: 4
2.59	3sg: 40 (*zijn* 'to be', *willen* 'to want', *horen* 'to hear', *moeten* 'must', *zijn*$_{aux}$ 'be') 1sg: 24 (*moeten* 'must', *hebben*$_{aux}$ 'have', *hebben* 'to have', *zijn* 'to be', *kunnen* 'can') 2sg: 1 (*moeten* 'must') 3pl: 1 (*zijn* 'to be') 1pl: 2 (*gaan* 'to go', *moeten* 'must') inf: 4 (*zitten* 'to sit', *helpen* 'to help', *stricken* 'to trick', *zwemmen* 'to swim')	pres perf: 3 pst: 4
3.15	3sg: 91 (*zijn* 'to be', *zitten* 'to sit', *gaan* 'to go', *zijn*$_{aux}$ 'be', *staan* 'to stand')	pres perf: 5 pst: 7

	1SG: 25 (*hebben* ‘to have’, *hebben*$_{aux}$ ‘have’, *vinden* ‘to find’, *zullen* ‘should’, *moeten* ‘must’) 2SG: 10 (*zitten* ‘to sit’, *doen* ‘to do’, *zijn* ‘to be’, *mogen* ‘may/to be allowed to’) 3PL: 4 (*gaan* ‘to go’, *zijn* ‘to be’) 1PL: 3 (*gaan* ‘to go’) INF: 3 (*slapen* ‘to sleep’, *lezen* ‘to read’, *kleuren* ‘to paint’) *Err:* 3		
2.88	3SG: 44 (*hoeven* ‘to need’, *zijn* ‘to be’, *mogen* ‘may/to be allowed to’, *zijn*$_{aux}$ ‘be’, *kunnen* ‘can’) 1SG: 43 (*gaan* ‘to go’, *doen* ‘to do’, *zijn* ‘to be’, *hebben* ‘to have’, *hebben*$_{aux}$ ‘have’) 2SG: 8 (*worden* ‘to become’, *zullen* ‘should’, *gaan* ‘to go’, *moeten* ‘must’) 3PL: 1 (*zijn* ‘to be’) INF: 3 (*uitzoeken* ‘to sort’, *zijn* ‘to be’, *doortrekken* ‘to go through’) *Err:* 3		PRES PERF: 8 PST: 2
2.87	3SG: 51 (*moeten* ‘must’, *zijn* ‘to be’, *gaan* ‘to go’, *hetten* ‘to be called’, *staan* ‘to stand’) 1SG: 31 (*willen* ‘to want’, *krijgen* ‘to get’, *kennen* ‘to know’, *vinden* ‘to find’, *hebben*$_{aux}$ ‘have’) 2SG: 6 (*mogen* ‘may/to be allowed to’, *vinden* ‘to find’, *moeten* ‘must’, *kunnen* ‘can’) 3PL: 1 (*gaan* ‘to go’) INF: 3 (*laten* ‘to let/leave’, *doen* ‘to do’, *wachten* ‘to wait’)	2	PRES PERF: 3 PST: 1
3.64	3SG: 27 (*moeten* ‘must’, *zijn* ‘to be’, *gaan* ‘to go’, *zitten* ‘to sit’, *doen* ‘to do’) 1SG: 29 (*willen* ‘to want’, *wetten* ‘to know’, *kunnen* ‘can’, *zien* ‘to see’, *gaan* ‘to go’) 2SG: 7 (*hebben*$_{aux}$ ‘have’, *hebben* ‘to have’, *moeten* ‘must’, *mogen* ‘may/to be allowed to’, *zien* ‘to see’) 1PL: 3 (*gaan* ‘to go’)	1	PRES PERF: 1 PST: 11

INF: 51 (*voorlezen* 'to read aloud', *opdrinken* 'to drink up', *slapen* 'to sleep', *vertellen* 'to narrate', *zijn* 'to be')
Err: 2