Emergent syntactic categories and increasing granularity

Evidence from a multilingual corpus study

Núria Bosch & Theresa Biberauer

University of Cambridge

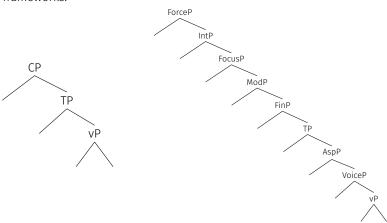
■ nb611@cam.ac.uk

nuria-bosch.github.io

BUCLD 48 (Boston University) — 3 November 2023

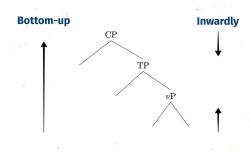
1. Introduction

■ Syntactic trees grow → differences in **granularity** across (and within) frameworks:



ACQUIRING FUNCTIONAL CATEGORIES

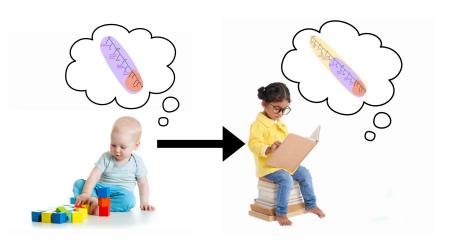
- How do children acquire these trees?
- **Prior maturational work:** focus on *directionality* of acquisition
 - Trees are acquired **bottom-up**: vP → TP → CP (i.a., Radford, 1990; Rizzi, 1993; Friedmann et al., 2021; Diercks et al., 2023).
 - Trees are acquired inwardly: vP & CP → TP (i.a., Galasso, 2003; Tsimpli, 2005; Heim and Wiltschko, 2021).



- What about granularity, though? How 'fine-grained' are children's trees at the start?
 - Implicit assumption in work thus far: granularity is fixed by UG. In cartographic approaches, as soon as a child acquires a specific domain, it is cartographic in nature.
- Existing cartographic approaches:
 - Westergaard (2009)'s micro-cues model: children have access to cartographic left-peripheral knowledge from the start.
 - Friedmann et al. (2021)'s Growing Trees: the cartographic left periphery emerges in two steps, and develops very late in its entirety.

Testable prediction: If (parts of) cartographic CP are available early, we should see (some) evidence for its distinct projections reasonably early (as noted by Westergaard, 2009; Moscati and Rizzi, 2021; Moscati, 2023)

- 2



Bigger tree, same granularity?

(Tree diagrams from Friedmann et al., 2021)

AIMS TODAY

- Granularity never changes throughout development?
- **Today:** revisiting the development of the left periphery:
 - Are categories acquired in a specific directionality?
 - But most importantly, when do children show evidence for access to an articulated CP domain?
- → Changes in granularity might be an important (unexplored) aspect of syntactic development.

- Maturation of functional categories
 - (Arguably) dominant approach so far: bottom-up approach.
 - The top of the tree (≈ CP) acquired last (Radford, 1990; Rizzi, 1993; Friedmann et al., 2021; Diercks et al., 2023).
 - Growing Trees Hypothesis (most recent, left periphery-centred proposal): two-stage development of LP.

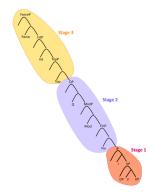


Figure 1: Stages in the Growing Trees Hypothesis (Friedmann et al., 2021, p. 12)

-4

- Maturation of functional categories
 - ► (Arguably) **dominant** approach so far: **bottom-up** approach.
 - The top of the tree (≈ CP) acquired last (Radford, 1990; Rizzi, 1993; Friedmann et al., 2021; Diercks et al., 2023).
 - Growing Trees Hypothesis (most recent, left periphery-centred proposal): two-stage development of LP.

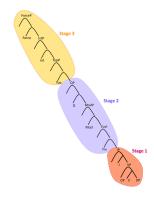


Figure 1: Stages in the Growing Trees Hypothesis (Friedmann et al., 2021, p. 12)

Bottom-up directionality, fixed granularity

- Maturation of functional categories
 - More recently revived idea: inward approach. CP emerges early! (i.a., Galasso, 2003; Tsimpli, 2005; Heim and Wiltschko, 2021).
 - Galasso (2003)'s 'Empty Middle' approach: CP>Ø>VP to CP>IP>VP.
 - Heim and Wiltschko (2021)'s Inward Growing Spine: spine matures inwardly.

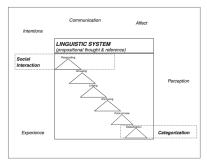


Figure 2: Bridge Model (Hinzen and Wiltschko, 2023)

Maturation of functional categories

- More recently revived idea: inward approach. CP emerges early! (i.a., Galasso, 2003; Tsimpli, 2005; Heim and Wiltschko, 2021).
- Galasso (2003)'s 'Empty Middle' approach: CP>Ø>VP to CP>IP>VP.
- Heim and Wiltschko (2021)'s Inward Growing Spine: spine matures inwardly.

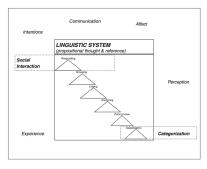


Figure 2: Bridge Model (Hinzen and Wiltschko, 2023)

Inward directionality, (generally) fixed granularity

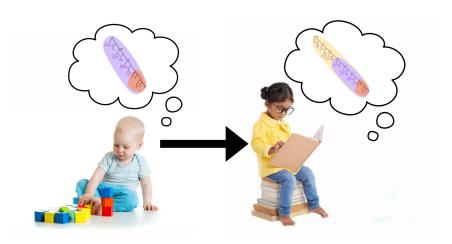
- **Continuity**: children's initial state ≈ adult's functional inventory.
 - Of various strengths: Strong Continuity, Weak Continuity (Underspecification of features, Lexical Learning, etc.) (i.a., Poeppel and Wexler, 1993; Hyams, 1992, 1996; Clahsen et al., 1994)
 - Westergaard (2009)'s micro-cues approach: sensitivity to cartographic structures early on.

¹Possible underspecification of features notwithstanding.

- **Continuity**: children's initial state ≈ adult's functional inventory.
 - Of various strengths: Strong Continuity, Weak Continuity (Underspecification of features, Lexical Learning, etc.) (i.a., Poeppel and Wexler, 1993; Hyams, 1992, 1996; Clahsen et al., 1994)
 - Westergaard (2009)'s micro-cues approach: sensitivity to cartographic structures early on.

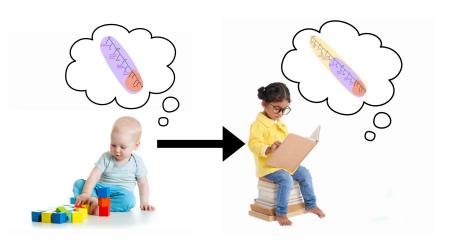
Continuity (no maturation), fixed granularity¹

¹Possible underspecification of features notwithstanding.



Bigger tree, same granularity?

(Tree diagrams from Friedmann et al., 2021)



Not in all approaches...

(Tree diagrams from Friedmann et al., 2021)

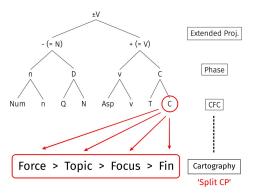
- **Neo-emergentism** (Biberauer, 2011, et seq.; Biberauer and Roberts, 2015)
 - Emergentist generative approach: **minimal UG**, no innate categories.
 - → Hypothesis relevant here: Biberauer and Roberts (2015)'s **emergent categorial**

hierarchy.

 Different levels of granularity across frameworks unified → different stages of a learning path (coarse- to fine-grained).

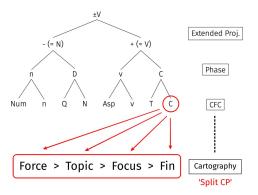
'Basic CP' before cartographic-type CP

(1) 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 (e.g., Cinque, 1999; Frascarelli and Hinterhölzl, 2007).



Syntactic categories 'granularise' (become refined) during development

(2) 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 (e.g., Cinque, 1999; Frascarelli and Hinterhölzl, 2007).



Any apparent directionality epiphenomenal, flexible granularity

Predictions for development of left periphery

Bottom up (Growing Trees):

- Late CP (two-stage).
- Fixed (cartographic) granularity: evidence for it once (or soon after) CP matures.

Inward maturation:

- Early CP.
- Fixed granularity: if cartographic, evidence for it once (or soon after) CP matures

■ Neo-emergentism (Biberauer and Roberts, 2015):

- Early CP.
- Flexible granularity: late emergence of cartographic elaboration of CP.

3. CORPUS STUDY

METHODOLOGY AND DIAGNOSTICS

 Longitudinal analysis of 10 typically-developing children in CHILDES, across five languages (Catalan, Spanish, Italian, German and Dutch)

Table 1: Children studied in the CHILDES database and summary information

Language	Corpus	Child	Files	Age	MLUw
Catalan	Serra-Solé	Laura	19	1;07-4;00	1.03-3.47
	Jerra Jole	Gisela	21	1;07-4;02	1.02-3.51
Italian	Calambrone	Martina	13	1;07-2;07	1.26-2.69
	Catambrone	Rosa	21	1;07-3;03	1.27-3.24
Spanish	Llinàs-Ojea	Irene	59	0;11-3;02	1.0-5.13
	Montes	Koki	13	1;07-2;11	1.96-3.61
German	Miller	Kerstin	37	1;03-3;04	1.09-2.89
	Miller	Simone	50	1;09-2;09	1.52-4.89
Dutch	Groningen	Josse	28	2;0-3;04	1.2-4.01
	van Kampen	Sarah	50	1;06-5;02	1.07-6.07

STRUCTURES ANALYSED: DIAGNOSTICS

- CP diagnostics:
 - 1. Wh-questions
 - Yes/no questions (Germanic only)
 - V-to-C movement (Germanic only)
 - 4. Topics/Foci
 - 5. Illocutionary (main clause) complementisers (Romance only)
 - 6. Finite embedding

- Split CP diagnostics (Romance):
 - 1. Top > Wh
 - 2. Top > Top/Foc
 - 3. Complementiser > Wh/Top
 - Quotative que 'that' > Wh (Ibero-Romance only)
 - 5. Topic > interrogative que 'that' (Catalan only)
 - Si que/sì che 'yes that' and que si que 'that yes that' structures (for the latter, Ibero-Romance only)

STRUCTURES ANALYSED: SPLIT CP DIAGNOSTICS

- (3) a. **La Júlia**, **on** ha anat? (Top > Wh, Catalan) the Júlia where AUX.HAVE.3SG go.PTCP 'Júlia, where has she gone?'
 - b. Questo, a te, ti spaventa (Top > Top/Foc, Italian) this to you CL.IO= scare.3sG 'This, it scares YOU.'
 - C. ¿Que cuánto te han costado estas that.QUOT how.much CL.IO= AUX.HAVE.3PL cost.PTCP these bambas? (Comp > Wh, Spanish) trainers

'How much have you said these trainers have cost you!?'

STRUCTURES ANALYSED: DIAGNOSTICS

- CP diagnostics:
 - 1. Wh-questions
 - Yes/no questions (Germanic only)
 - 3. V-to-C movement (Germanic only)
 - 4. Topics/Foci
 - 5. Illocutionary (main clause) complementisers (Romance only)
 - 6. Finite embedding

- Split CP diagnostics (Germanic, V3 orders):
 - 1. Frame-setters
 - 2. Hanging Topic Left-Dislocation
 - 3. Contrastive Left-Dislocation
 - 4. Conditional/temporal clauses with resumptive dann/dan 'then'

STRUCTURES ANALYSED: SPLIT CP DIAGNOSTICS

- (4) a. **In alle geval**, ik had het niet verwacht (Frame-setter, Dutch) 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. **Diesen Kuchen hier**, den möchte ich probieren the ACC cake ACC here PRON.ACC want.1sg I try.INF

 'This cake here, I want to try.'
 - c. **Als** het niet zo warm is, **dan** ga ik naar buiten (Conditional with when/if it not so hot be.3sg then go.1sg I to outside resumptive, Dutch)

'When/if it isn't so hot, then I'll go out.'

3. CORPUS STUDY

3.1. Results and generalisations

RESULTS: STAGES OBSERVED

- Transparent order of appearance of the structures analysed in the ten children. Very early CP emergence. Split CP structures systematically emerge at a later stage.
 - 1. Stage 1 CP structures: 1.38 MLUw (range 1.15-1.54).
 - 2. Stage 2 CP and TP structures: 1.64 MLUw (range 1.44-1.94).
 - 3. Stage 3 Split CP structures: 2.57 MLUw (range 2.32-2.8).
- Focus on Stages 1 and 2 (as a group) vs Stage 3 here.

MAIN GENERALISATIONS

Generalisation 1: Early Acquisition of CF

CP-structures emerge early on in the developmental data.

MAIN GENERALISATIONS

Generalisation 1: Early Acquisition of CF

CP-structures emerge early on in the developmental data.

→ 'Directionality' of emergence likely isn't bottom-up.

Some children:

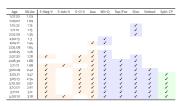


Table 2: Production of structures by Laura (Catalan)



Table 3: Production of structures by Rosa (Italian)

See Appendix for full tables of all children.

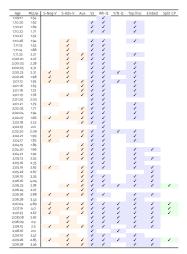


Table 4: Production of structures by Simone (German)

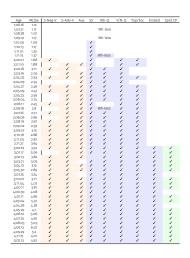


Table 5: Production of structures by Sarah (Dutch)

Table 6: CP-structures produced at Stages 1 + 2 and its length

	V2	Wh-Q	Y/N-Q	Top/Foc	Illoc	Embed	Length
Laura		15		4	42	4	1;10.22-3;03.21
							(MLUw 1.15-2.54)
Gisela		1		0	6	0	2;04.25-2;08.00
							(MLUw 1.58-2.61)
Martina		21		3	7	8	1;08.02-2;04.13
							(MLUw 1.57-2.69)
Rosa		133		12	3	8	1;07.13-2;10.14
							(MLUw 1.27-2.5)
Irene		18		3	10	4	1;04.16-1;11.13
							(MLUw 1.32-2.95)
Koki		32		7	2	4	1;07.20-2;04.18
							(MLUw 1.96-2.69)
Kerstin	/	16	21	27		1	1;10.03-2;09.11
							(MLUw 1.28-2.32)
Simone	/	166	3	105		24	1;10.03-2;06.23
							(MLUw 1.54-2.78)
Josse	/	62	37	68		1	2;00.07-2;11.09
							(MLUw 1.2-3.57)
Sarah	/	124	104	116		0	1;10.05-3;00.19
							(MLUw 1.09-3.52)

MAIN GENERALISATIONS

Generalisation 2: Some structurally very high elements emerge early

Topics Before Finite Embedding: Topicalisation emerges before finite embedding markers in most of the children studied.

Early Illocutionary Complementisers: Illocutionary complementisers emerge well before embedding complementisers (Bosch, 2023c).

 Simultaneous emergence of embedding markers and topicalisation in Friedmann et al. (2021) (their Stage 3) is, in several instances, not replicated.

Table 7: Emergence of topicalisation vs embedding markers

	Topicalisation	Embedding
Laura	2;08.03	3;00.02
	1.88 MLUw	2.42 MLUw
Gisela	2;08.00	2;08.00 (same file)
	2.61 MLUw	2.61 MLUw
Martina	1;11.20	1;11.20 (same file)
	1.99 MLUw	1.99 MLUw
Rosa	2;04.29	2;06.29
	1.77 MLUw	2.6 MLUw
Irene	1;08.09b	1;09.10
	2.24 MLUw	3.28 MLUw
Koki	1;11.25	1;11.25 (same file)
	2.47 MLUw	2.47 MLUw
Kerstin	2;00.05	2;07.23
	1.76 MLUw	2.13 MLUw
Simone	1;10.20	2;04.20
	1.62 MLUw	1.96 MLUw
Josse	2;03.28	2;09.02
	1.94 MLUw	2.42 MLUw
		3;00.19
Sarah	2;00.17	3,00.19
Sarah	2;00.1/ 1.68 MLUw	3,50 MLUw

MAIN GENERALISATIONS

 Illocutionary complementisers also emerge from the earliest files for many children (Bosch, 2023c).

- (5) a. Ai, **que** crema! (Laura, MLUw 1.35) ouch that.EXCL burn.3sG 'Ouch, it's burning!'
 - b. Que cau! (Laura, MLUw 1.3) that.EXCL fall.3sg 'It's falling!'
- → Development cannot be recapitulating a cartographic spine in a bottom-up manner.
 - ► Some of the structurally highest elements don't emerge last.

MAIN GENERALISATIONS

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).

30

Table 8: Emergence of CP- vs Split CP-structures

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

25

MAIN GENERALISATIONS

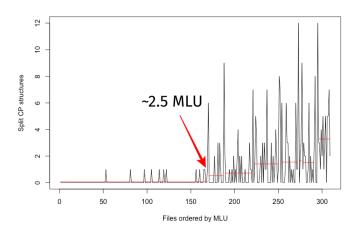
■ Emergence is not just late, but sudden and 'explosive' in the production data (z = -2.949874, p = 0.003).

Table 9: Production of Split CP-structures before and after MLUw ~ 2.5

	Before MLUw ~ 2.5	After MLUw ~ 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%

CHANGE POINT ANALYSIS

Detecting when the change occurs with change point analysis



INTERIM SUMMARY

- **So far:** Closer look at the development of left peripheral knowledge reveals two significant trends:
 - CP knowledge emerges early, and not in a way that recapitulates a cartographic spine bottom-up.
 - Evidence for articulated CP structure emerges significantly late (after TP and complex structures like subordination).

- **So far:** Closer look at the development of left peripheral knowledge reveals two significant trends:
 - CP knowledge emerges early, and not in a way that recapitulates a cartographic spine bottom-up.
 - Evidence for articulated CP structure emerges significantly late (after TP and complex structures like subordination).
- Not addressed in this talk: patterns also not explained away by the relative length of these two groups of structures: Split CP structures could have 'fitted' in earlier stages (Extra slides).
 - ▶ A performance-based account of the patterns probably does not suffice.

4. THEORETICAL ACCOUNT AND IMPLICATIONS

IMPLICATIONS

■ Most theoretical approaches → 'fixed granularity', imposed by UG. Development accounted for by 'recapitulating' this spine, either bottom-up or inwardly.

IMPLICATIONS

- Most theoretical approaches → 'fixed granularity', imposed by UG. Development accounted for by 'recapitulating' this spine, either bottom-up or inwardly.
- **Results here**: it's not (just) about directionality and fixed granularity.

8

- Most theoretical approaches → 'fixed granularity', imposed by UG. Development accounted for by 'recapitulating' this spine, either bottom-up or inwardly.
- **Results here**: it's not (just) about directionality and fixed granularity.
- Generalisations 1-3 generate a contradiction in current maturational approaches.
 - ► Early CP emergence → challenges bottom-up approaches.
 - ► Early topics and late embedding → challenges a cartographic bottom-up approach.
 - Split CP is late → challenges any account with innate functional categories (either bottom-up or inward-growing, and continuity).
 - Early CP but late cartographic-type left periphery?

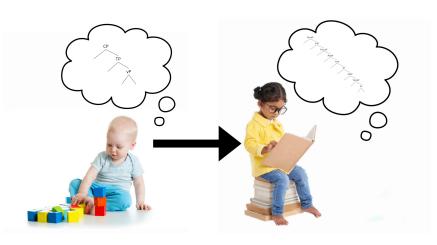
- Most theoretical approaches → 'fixed granularity', imposed by UG. Development accounted for by 'recapitulating' this spine, either bottom-up or inwardly.
- **Results here**: it's not (just) about directionality and fixed granularity.
- Generalisations 1-3 generate a contradiction in current maturational approaches.
 - ► Early CP emergence → challenges bottom-up approaches.
 - ► Early topics and late embedding → challenges a cartographic bottom-up approach.
 - Split CP is late → challenges any account with innate functional categories (either bottom-up or inward-growing, and continuity).
 - Early CP but late cartographic-type left periphery?

Innate categories (fixed granularity) and directionality-based maturation lead to this 'deadlock'

THE PROPOSAL

- Our proposed solution here: dropping innate categories.
 - → Emergent categories lend us the flexibility needed to rationalise these patterns.
- Biberauer and Roberts (2015)'s emergent categorial hierarchy:
 - ► First, children access core 'macroparametric' structural properties (see also work on 'Very Early Parameter-setting') → basic CP domain.
 - Once mastered, these enable ('unlock') more complex, increasingly 'micro-parametric' refinements → (part-)cartographic structure.
 - Input vs intake discrepancies (Tsimpli, 2014; Gagliardi, 2012; Lidz and Gagliardi, 2015).
- Understanding the **contribution** of neo-emergentism:
 - ✓ Emergent categories → expect departures from strict directionality.
 - ✓ Emergent categories and increasing granularity go hand-in-hand.
 - ✓ Discrete change in representations → 'sudden' and 'explosive' emergence expected.

- Our proposed solution here: dropping innate categories.
 - → Emergent categories lend us the flexibility needed to rationalise these patterns.
- Biberauer and Roberts (2015)'s emergent categorial hierarchy:
 - ► First, children access core 'macroparametric' structural properties (see also work on 'Very Early Parameter-setting') → basic CP domain.
 - Once mastered, these enable ('unlock') more complex, increasingly 'micro-parametric' refinements → (part-)cartographic structure.
 - Input vs intake discrepancies (Tsimpli, 2014; Gagliardi, 2012; Lidz and Gagliardi, 2015).
- Understanding the **contribution** of neo-emergentism:
 - ✓ Emergent categories → expect departures from strict directionality.
 - ✓ Emergent categories and increasing granularity go hand-in-hand.
 - ✓ Discrete change in representations → 'sudden' and 'explosive' emergence expected.
- → This not just accommodates, but crucially *predicts*, the patterns observed



Bigger tree, different granularity

5. CONCLUSION

CONCLUSION

→ **Novel generalisation**: early CP vs late Split CP knowledge, in production.

CONCLUSION

- → **Novel generalisation**: early CP vs late Split CP knowledge, in production.
- Generalisations 1-3 foreground three (largely) new challenges and requirements:
 - Early CP emergence.
 - ▶ A move away from exclusively directionality-centred approaches.
 - Potential role of granularity and categorial flexibility: neither fixed nor always fine-grained in development.
- Further work needed:
 - More children/languages, other structures and syntactic domains (work in progress!)
 - Comprehension/behavioural studies (although non-trivial to probe)
 - Alternative explanations for the patterns?

- → **Novel generalisation**: early CP vs late Split CP knowledge, in production.
- Generalisations 1-3 foreground three (largely) new challenges and requirements:
 - Early CP emergence.
 - ▶ A move away from exclusively directionality-centred approaches.
 - Potential role of granularity and categorial flexibility: neither fixed nor always fine-grained in development.
- Further work needed:
 - More children/languages, other structures and syntactic domains (work in progress!)
 - Comprehension/behavioural studies (although non-trivial to probe)
 - ► Alternative explanations for the patterns?
- More generally, *productive questions and patterns* surface when probing acquisition through a **neo-emergentist** lens.

Thank you!

Acknowledgements: Thanks in particular to Dora Alexopoulou, Cécile de Cat, Bert Vaux, Roman Feiman, Itamar Schatz, Julia Schwarz, Henry C. and reviewers for BUCLD 48 for very helpful comments and help. This work was generously supported by St John's College (Cambridge), the Cambridge Trust and the Arts and Humanities Research Council (AHRC, UKRI).



6. EXTRA SLIDES AND APPENDIX

Age is not a reliable predictive factor of timeline of emergence of structures, presenting high variance within each Stage (as in Friedmann et al., 2021). It's the stages that remain identical across children.

Table 10: 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

EXTRA SLIDES Additional case studies

- 'Basic' before 'cartographic-type' patterns repeat themselves in other work:
 - ▶ De Lisser et al. (2017) on acquisition of the **TMA** field in **Jamaican Creole**
 - Co-ocurrence of TMA markers systematically at Phase 2 (MLU 2.5-3.49) or Phase 3 (MLU > 3.5) in the data reported. No examples at Stage 1 (< MLU 2.5).
 - Development of **PPs** (Sanfelici and Gallina, 2022) in **Italian**
 - Bimorphemic prepositions (such as dentro a 'inside', sopra di 'above') only in Groups 3 (MLU 2.50-2.99) and Groups 4 (3.0-3.49).
 - ► Mitrofanova (2018)'s **Underspecification of P Hypothesis**
 - Initial stage with a coarse-grained prepositional category, but without cartographic heads encoding fine-grained meaning distinctions (such as Svenonius's, 2006, 2008, AxialPartP).

RESULTS: LONGER UTTERANCES OR EMERGENT KNOWLEDGE?

- But, do Split CP structures emerge 'late' simply because lower utterance lengths cannot accommodate these constructions (even though the child's competence does capture them)?
 - ▶ **Likely not**. Arguments come from two domains: mixed effects logistic regression and comparison of production lengths across stages.

Mixed effects logistic regression

Controlling for length of Split CP vs CP structures as the driving factor of the patterns → mixed effects logistic regression model with length of (a sample of) the CP/Split CP utterances analysed and MLU as controls². Model included the children studied as random effect (formula: 1 | child)

Results:

- → The effect of **mlu** is **highly** statistically significant and positive (beta = 1.67, 95% CI [0.99, 2.35], **p** < .001; Std. beta = 1.18, 95% CI [0.70, 1.65])
- → The effect of **length** is **moderately** statistically significant and positive (beta = 0.15, 95% CI [5.58e-03, 0.30], p = 0.042; Std. beta = 0.38, 95% CI [0.01, 0.74])

²Note that the MLU and Length variables were tested for multicollinearity with Variance Inflation Factor. Their VIF = 1.21, indicating there is no correlation between both variables. Age was dropped as a predictor variable, as it was moderately correlated with MLU (VIF = 6.16).

RESULTS: LONGER UTTERANCES OR EMERGENT KNOWLEDGE?

Patterns cannot be accounted for entirely by *length*. MLU (as an average length of *all* utterances and metric of syntactic development) is a much stronger predictor.

 NB: Importantly, length also cannot account for the 'suddenness' and 'explosiveness' with which Split CP structures emerge (growth of utterance length often isn't exponential).

RESULTS: LONGER UTTERANCES OR EMERGENT KNOWLEDGE?

Corpus data: comparison across stages

(6) a Aquest on va?

Structures at Stage 1 or Stage 2 can occasionally be as long as or even longer than those at Stage 3, raising problems for utterance length as a complete account of the patterns.

(Catalan Cicola Stago a)

(6)	d.	this where go.3sg 'This one, where does it go? / This one, where is it going?'	(Catalan, disela – Stage 3)
	b.	Jo tinc un petit suisse I have.1sG a petit suisse 'I have a petit suisse.'	(Catalan, Gisela – Stage 1)
	C.	No, jo em $$\operatorname{vull}$$ treure els patins no I CL.REFL want.1SG take.off.INF the skates 'No, I want to take off the skates.'	(Catalan, Gisela – Stage 2)
(7)	a.	Nog ik heb het gegeven yet I AUX.HAVE.1SG it give.PTCP 'Yet gave it.'	(Dutch, Josse – Stage 3)
	b.	Wat doet ie nou? what do.3sG he now 'What is he doing now?'	(Dutch, Josse – Stage 1)

c. Kan niet zo een zwembad maken van de duikplank (Josse – Stage 2) can.1sg not so a pool make.INF from the diving.board 'I can't jump to the swimming pool from the diving board.'

Corpus data: comparison across stages

- Derivational Complexity accounts (e.g., Jakubowicz, 2005, 2011) offer a partial account at best:
 - Many Split CP-structures do not require more movement (more derivational complexity) than CP-structures → they thus fall outside the definition of 'complex' in these approaches (see the Derivational Complexity Metric in Jakubowicz, 2011).
- For more in depth discussion, see the note in Bosch (2023a).

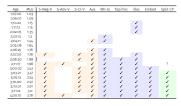


Table 11: Production of structures by Laura (Catalan)



Table 12: Production of structures by Gisela (Catalan)

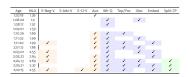


Table 13: Production of structures by Martina (Italian)



Table 14: Production of structures by Rosa (Italian)

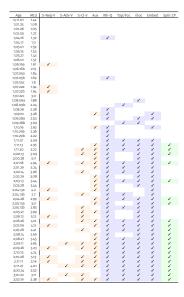


Table 15: Production of structures by Irene (Spanish)

Age	MLU	S-Neg-V	S-Adv-V	S-CI-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	/			1	1	/		/	
2;01.29	2.51	/			1	/	/			
2;02.27	2.47				1	/	/	1	/	
2,03.21	2.07	l .			1	/	/	/		
2;04.18	2.69	/		/	/	/	/		/	/
2;05.24	3.08	/	/	/	/	/	/		/	/
2;06.10	2.71	/	/	/	/	/	/	1	/	/
2:07:10	3.61	/	/	/	1	/	/	1	/	/
2;08.09	2.75	/	/	/		/	/	/	/	/
2,09.14	2.93			1	1	1	/		1	
2:11.14	3.38	/	/	1	1	1	/	1	/	1

Table 16: Production of structures by Koki (Spanish)

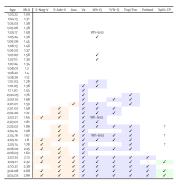


Table 17: Production of structures by Kerstin (German)

Age	MLUw	S-Neg-V	S-Adv-V	Ацх	V2	Wh-Q	Y/N-Q	Top/Foc	Embed	Split CP
1,09.11	154					/				
1,10.20	1.62	ı			1	/		/		
1,10.21	1.69				1	/		/		
1;10.22	1.71	ı			1	/		/		
1,10.27	1.52	ı								
1,10.28	1.94	ı	/		1	/		/		
1,11.13	1.53	ı	/		1	/				
1,11.14	1.88	ı	******		1	/				
1,11.23	2.21	l	/	/	1	/		/		
2,00.01	2.27	l	/	/	1	/				
2;00.03	2.28	l	/		/	1		/		
2,00.05	2.31	l	/		/	/		/		
2,00.23	2.31	/	/		1	1	1	/		
2,00.26	1.98	/	/		/	/		1		
2;01.12	1.93	. /	/	1	1	/	1	/		
2,01.16	1.63	1		1	1	/		/		
2,01.18	1.72	1		1	1	/		1		
2;01.19	1.78	ı	/	/	/	/		/		
2;01.20	2.03				1	/				
2,01.21	1.79					/		/		
2,02.03	1.71	ı		1	/	1		/		
2,02.04	1.94		/	1	1	1		/		
2;02.07	1.66	-	/	1	/	/		/		
2,02.18	2.22			1	1	/				
2,02.19	2.0	l				1				
2,02.20	2.09	/	/	1	1	/	/	/		
2,02.21	1.99	/	/	1	1	/		/		
2,04.17	1.82	/			1	/		/		
2,04.19	1.89	l		1	1	/				
2,04.20	1.96	l		1	1	/		/	/	
2,04.21	1.92	l	/	1	1	/		/		
2,0513	2.52		1	1	1	1		· ·	/	
2,05.16	2.35				/	/		/		
2,05.19	2.62			1	1	/		/	/	
2;05.22	2.67	ı		1	1				1	
2;06.10	3.35	ı		/	1	1		/	/	
2,06.16	4.04	ı		1	1	/		/	1	
2;06.23	2.78	ļ.	/	1	1	/	/	/	-	/
2,06.24	2.27			1	1					
2,05.25	2.88	/	/	1	1	/		· ·	1	
2,05.28	3.43				1			1		
2,07.04	4.89	٠.	1	1	1	1	٧.	1	1	٠.
2;07:19	4.0	1	/	1	1	1	1	1	1	1
2,07.23	2.67	. /		/	1	1	٧.	٠.	1	/
2,08.08	2.97	1	4.		1	1	1	1	1	
2,08.09	2.9		1		1	1	/	1	1	
2,08.15	2.5	/	/	1	1	/		/	/	
2;08.16	2.0	1			1		1			
2,09.10	3.47		٠.		1	1		٧.		
2,09.25	2.85	. /	1	٧.	1	1	/	1	1	/
2,09.28	3.46	I	/	1	/	-		/	/	

Table 18: Production of structures by Simone (German)

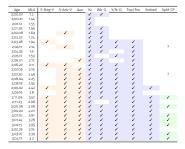


Table 19: Production of structures by Josse (Dutch)

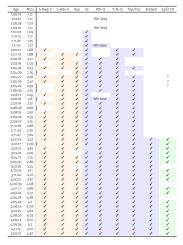


Table 20: Production of structures by Sarah (Dutch)

7. REFERENCES

- 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. and Roberts, I. (2015). Rethinking formal hierarchies: A proposed unification. *Cambridge Occasional Papers in Linguistics*, 7:1–31.
- Bosch, N. (2023a). 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. (2023b). Emergent Syntax and Maturation: a neo-emergentist approach to syntactic development. MPhil thesis, University of Cambridge.
- Bosch, N. (2023c). Not all complementisers are late: a first look at the acquisition of illocutionary complementisers in Catalan and Spanish. *Isogloss. Open Journal of Romance Linguistics*, 9:1–39.
- Cinque, G. (1999). Adverbs and Functional Heads: A Cross-Linguistic Perspective. Oxford University Press, Oxford.

REFERENCES II

- 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.
- 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.
- Diercks, M., Johns, K., Bar-Server, and Bossi, M. (2023). Developmental Minimalist Syntax. Ms. https://lingbuzz.net/lingbuzz/007134.
- 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.
- 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.
- Gagliardi, A. C. (2012). *Input and intake in language acquisition*. PhD thesis, University of Maryland, College Park.
- Galasso, J. (2003). The Acquisition of Functional Categories: A Case Study. IUCL Publications, Indiana University.

- 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.
- 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.
- 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. (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.
- Jakubowicz, C. (2005). The language faculty: (ab)normal development and interface constraints. Talk presented at *GALA*, University of Siena.

REFERENCES IV

- Jakubowicz, C. (2011). Measuring derivational complexity: New evidence from typically developing and SLI learners of L1 French. *Lingua*, 121(3):339–351.
- Lidz, J. and Gagliardi, A. (2015). How Nature Meets Nurture: Universal Grammar and Statistical Learning. *Annu. Rev. Linguist.*, 1(1):333–353.
- Mitrofanova, N. (2018). Early Underspecification of Functional Categories: Evidence from the Acquisition of Locative PPs in Russian. *Language Acquisition*, 25(4):341–365.
- Moscati, V. (2023). he observation of superiority on multiple movements to the Italian left-periphery: Intervention effects on nested dependencies and the role of information-structure features. *Journal of Child Language*, pages 1–33.
- 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.
- Poeppel, D. and Wexler, K. (1993). The Full Competence Hypothesis of Clause Structure in Early German. *Language*, 69(1):1–33.
- Radford, A. (1990). Syntactic theory and the acquisition of English syntax: The nature of early child grammars of English. Wiley Blackwell, Oxford.
- Rizzi, L. (1993). Some notes on linguistic theory and language development: The case of root infinitives. *Language Acquisition*, 3(4):371–393.

- Sanfelici, E. and Gallina, C. (2022). The timing of production: on the acquisition of Italian prepositions. *Isogloss*, 8(2):1–22.
- 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.
- 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.
- Tsimpli, I. M. (2014). Early, late or very late?: Timing acquisition and bilingualism. Linguistic Approaches to Bilingualism, 4(3):283–313.
- Westergaard, M. (2009). *The Acquisition of Word Order*. John Benjamins, Amsterdam.