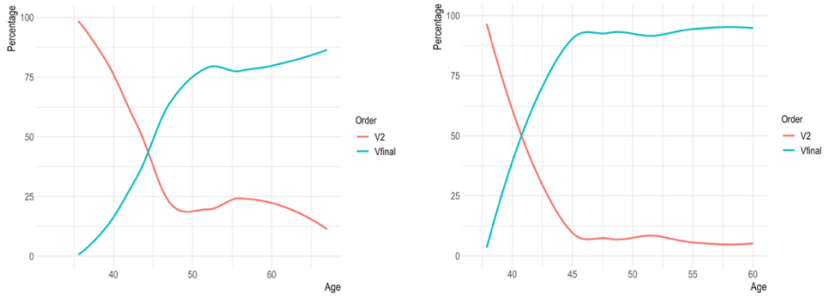
V2 all the way down: Germanic innovations in the embedded CP of German-Italian bilinguals

Núria Bosch & Theresa Biberauer
University of Cambridge

Overgeneralisation of Embedded V2 (EV2) has been observed to different degrees in monolinguals and bilinguals (see, i.a., Westergaard & Bentzen 2007; Ringstad & Kush 2021, and Jensberg et al. 2024, on Norwegian monolinguals and heritage speakers; Schönenberger 2001 on Swiss German monolinguals; Müller 1994 *et seq*. on a German-French bilingual). The source, extent and formal nature of this overgeneralised EV2 remains contentious, however, and it often seems language-specific. This paper presents a novel in-depth corpus study of the development of word order in the embedded CP of 5 German-Italian bilinguals (Müller corpus). Our contribution is three-fold: (i) we report systematic patterns of overgeneralisation of EV2 in German; (ii) we argue against an account of the data assuming transfer from Italian (*pace* Müller 1994, 2003; but cf. Müller 1996); and (iii) unlike most previous work, we argue that overgeneralised EV2 involves *L1-internal complexification* of the embedded CP. Specifically, we claim that embedded V-to-C and topic/subject movement, both extensions of German grammatical properties, evidence a more richly articulated CP in the relevant bilinguals than that in monolingual adult German. These results make diachronic predictions and furthermore align with theories advocating developmentally and crosslinguitically *variable degrees of elaboration of the CP* (e.g. Giorgi & Pianesi 1997; Biberauer & Roberts 2015; Bosch 2023).

Our **corpus study** investigated the **development of embedded clauses** (ECs) and **embedding markers** across **5 bilinguals**. We report, **firstly**, three significant broad patterns in the development of (target-like V-final) embedded clauses/ECs: (i) an initial stage, observed in 4/5 children, with overgeneralisation of embedded V2 (EV2) with complementisers and *total* or *partial* absence of embedded V-final orders; (ii) both linear EV2 (embedding marker>V>Subj) and linear embedded V3 (embedding marker>Subj/Top>V) co-existing at this stage; and (iii) in the children who initially show *total* absence of V-final orders, these crucially appearing once EV2 reduces in frequency (**Figs. 1 and 2**). EV2 with *dass*/*that*-drop emerges late in all of the children, on the other hand (mean average: 48 months, MLUw 4.25; emergence of EV2 is on average at 35.2 months, MLUw 3.3).



Figs 1-2. Development of embedded word order (V2, V-final) in the children CAR and MAR.

EV2 is abundant across *all* of wh-complements/relatives, and clauses introduced by *wenn* ('if') and *weil* ('because') (419/1124 ECs). **Secondly**, however, we report *differential patterns* across embedding markers, notably in the nature of the *pre-verbal* constituent in EV2 (e.g. (non)subject). We establish, like Schönenberger (2001) for Swiss German monolinguals, that embedded topicalisation is unattested with wh-complements/relatives and *wenn*, but is common with *weil* (2 *vs* 58 examples attested across the 5 children). *Dass*, being one of the later-acquired embedding markers, presents no EV2, like *ob* ('whether') or *als* ('as/when').

Thirdly, *wh*-V2 represents the final, novel peculiarity in our dataset: full V2 languages, including those permitting general EV2 (Icelandic, Yiddish), *disallow* V2 in embedded wh-complements; Afrikaans, a contact system, is the exception here (Biberauer 2017). *Wh*-V2 is

not just frequently attested, but has plausibly generalised in our data to predicates disallowing embedded *wh*-V2 (Vikner 1995): the children produce it with *discover*-type/'resolutive' predicates (e.g. *gucken* 'look', *wissen* 'know', *hören* 'hear', *sagen* 'say'*, erklären* 'explain'), without the illocutionary force of a true question (see McCloskey 2006 on English).

With this in place, we **discard a transfer-based analysis** (from Italian) of the data and argue that EV2 verbs move to CP: V_{FIN} systematically moves *above* negation and adverbs, it *directly follows* topicalised constituents, $V_{NON\text{-}FIN}$ always follows the object (OV) in complex structures with modals and auxiliaries, and, while linear V2/V3 orders are found, V3+ orders in embedded clauses are unattested. We interpret the data as an ***extension*** of a *Germanic* pattern (V- and XP-movement to CP) that is subsequently *formally integrated* into the bilinguals' German.

More precisely, we contend that the data points to ***elaboration/complexification of the embedded CP***. We analyse the overgeneralised EV2 stage in terms of three projections in the embedded CP ($SubP>CP_1>CP_2$), and draw on Bhatt & Yoon's (1992) two-way distinction between 'pure' and 'modal-flavoured' complementisers (the latter structurally lower) to account for differences across embedding markers. Initially, we assume acquirers posit a (minimally) expanded CP in ECs, analogous to the structure in main clauses, adopting Walkden's (2017) (non-recursive) CP_1 and CP_2 for monolingual German. We propose that wh-words in wh-complements/relatives and *wenn*, being modal-flavoured (conditional, interrogative), are located in CP_2, below SubP, (see also Roussou 2000, and subsequent work) and that subjects following embedding markers in linear V3 are hosted in the lower CP_1 at this stage. This correctly rules out embedded topicalisation with wh-complements and *wenn*. In contrast, we assume *weil* is hosted in SubP, thereby sanctioning embedded topicalisation in CP2 (see Antomo & Steinbach 2010). SubP also voids a violation of the Kayne-Rizzi-Roberts effect in *wh*-V2, which arises without this 'insulating' layer (McCloskey 2006; Biberauer 2017; cf. the analysis in Schönenberger 2001). Finally, we attribute co-existing linear EV2/3 to optionality in subject/topic-raising (which already obtains in adult German, Grewendorf 1989, Diesing 1992, Haider 1993; and in acquisition, e.g. van Kampen 2020).

The above crucially follows from **our proposed conceptualisation of learners and categorial acquisition**, which adopts the **Maximise Minimal Means** (MMM) logic in Biberauer (2018, 2019): MMM, as a third-factor principle, leads to a predilection for initially generalising structures and [*F*]s already in their grammars (see also Roberts 2007). This generalisation may be on the basis of initial *ignorance* of more complex relevant distinctions (e.g. word-order distributions in German clauses and discourse restrictions on EV2), which is then constrained as awareness of semantico-pragmatic conditioning in the input (Hooper & Thompson 1973) becomes accessible. In essence, our proposal is that these acquirers amplify the regularity of a pattern in the input (V2, "boosted" by the availability of embedded SVO in Italian), and crucially extend it to integrate with the formal structure of their existing German grammar, thereby driving innovative CP-complexification. Our conclusions support previous work emphasising the significance of *varying degrees of CP-elaboration* ('granularisation') in both (emergent) categorial acquisition and language variation (see i.a. Biberauer & Roberts 2015; Walkden 2017; Hsu 2017; Bosch 2023; Cournane & Klævik-Pettersen 2023).

Finally, this analysis also makes **diachronic predictions** regarding (E)V2: MMM predicts later-acquired properties, contingent on more complex input, to be vulnerable to change, especially under contact. Input-divergent [*F*]-overgeneralisation in favour of V2 in embedded clauses may thus lead to change if acquirers fail to retract (see also Cournane 2019 on incrementation). We provide three brief case studies of diachronic EV2 overgeneralisation to show that these expectations are borne out: Afrikaans (Biberauer 2017), Manenberg Kaaps (Van Rooi 2022), and Cimbrian, the latter displaying EV2 innovations known also *not* to be mere extensions of Romance SVO (see Bidese et al. 2014 *et seq.* on the complementiser *ké* and the rise of a two-class complementiser system, also seen in Afrikaans and Kaaps).

Overall, our work therefore has implications for categorial development in acquisition, the stability of V2, embedded word order and complementiser systems under child bilingualism-mediated crosslinguistic contact and why this differs from what is seen adult bilingual varieties (see Trudgill 2011 on 'complexification' and 'simplification'). It also speaks to the crosslinguistically attested typology of V2, and the conceptualisation of 'economy' principles.

References (selected): Bhatt & Yoon (1992) On the composition of Comp and parameters of V-2; **Biberauer (2017)** Optional V2 in Modern Afrikaans; **Biberauer & Roberts (2015)** Rethinking formal hierarchies: a proposed unification.; **Müller (1994)** Parameters Cannot Be Reset: Evidence from the Development of COMP; **Schönenberger (2001)** *Embedded V-to-C in child grammar: The acquisition of verb placement in Swiss German*.