FASL 26, ###-###
Michigan Slavic Publications
2018

The grammatical source of missing epistemic meanings for modal verbs in child BCS*

Dunja Veselinović
Ailís Cournane
New York University

Children use functional modals (e.g., *must, have to*) with root meanings (e.g., abilities, obligations) by age 2, but with epistemic meanings (i.e. knowledge-based inferences) only by age 3 (Stephany 1979; Papafragou 1998; i.a.). What can explain this Epistemic Gap (EG)? We present a corpus study of eight Bosnian/Croatian/Serbian (BCS) children and their maternal input. The BCS children's EG lasts until at least age 4, a year longer than observed for English children. We show that the EG can be accounted for by language-specific syntactic differences between epistemic and root representations of modal verbs (Cournane 2015), rather than conceptual or input-frequency differences. We argue that epistemic use of modal verbs relies on TP-embedding in English, but on later CP-embedding in BCS (Veselinović 2017).

* We thank Stephanie Harves, the ModSquad (UMD & NYU), FASL26 and NYU's Syntax Brown Bag audiences, our reviewers and editors of this volume. Errors are our own. This research was supported in part by NSF grant #BCS-1551628.

1 The Epistemic Gap

Modal verbs in many languages, including English, are functional (i.e., auxiliaries or functional verbs) and express both major modal flavors: root (1a) and epistemic meanings (1b). Lexical modals express only one of the broad flavors of modality ((2); see Hacquard 2013).

(1) a. Mary must do her homework.
b. Mary must be doing her homework.

(2) It is **probable** that Mary is **obliged** to do her homework.

Longitudinal naturalistic acquisition studies observe that root modal uses precede epistemic (e.g., Kuczaj & Maratsos 1975; Papafragou 1998), showing what we call an Epistemic Gap (EG). The EG refers to an approximately year long period from 2 to 3 years-old (with some cross-linguistic variation, see Stephany (1993) for an overview, and Smoczynska (1993) for Polish) in which children use functional modals with only root meanings. We present new results from a corpus study of eight children acquiring BCS and their maternal input, and establish that BCS children exhibit an EG for a year longer than observed in English. Our findings support the grammatical hypothesis (Cournane 2015), that during the EG children lack the grammatical representations needed to support epistemic interpretations of functional modal verbs. We show that neither the conceptual hypothesis (children lack the conceptual ability necessary for epistemic meanings) nor the frequency hypothesis (EG as an effect of input frequency) account for the protracted EG in BCS straightforwardly. We further argue that the syntax of individual languages affects acquisition of epistemic uses of modal verbs. While TP-embedding suffices in English, BCS children cannot use modal verbs epistemically until they can embed CPs (Veselinović 2017).

1.1 The Conceptual Hypothesis

A longstanding and widely-accepted account of the EG suggests that children lack the conceptual abilities necessary to support epistemic meanings (Astington et al. 1990; Shatz & Wilcox 1991; Papafragou 1998, i.a.). This account developed primarily based on naturalistic uses of English canonical functional modals, as well as functional modals cross-linguistically (Greek, Stephany 1979; German, Stephany 1993; French, Bassano 1996). Previous literature highlights several issues with this approach to explaining the EG (de Villiers 2007; Cummins 2013).

Cournane (2015) argues that the prior focus on functional modals like *must* or *can*, to the exclusion of simpler lexical modals like *maybe* or *probably*, introduced a grammatical complexity confound. As languages express epistemic modality using multiple grammatical categories (e.g., Palmer 1986; Kratzer 2012), studying only functional modals constitutes a sampling error. Lexical epistemic modals with dedicated epistemic meanings (Rett & Hyams 2014) occur during the EG. At age 2, English children use *maybe* and *probably* (O'Neill & Atance 2000, Cournane 2015), French children use the adverb *peut-être* 'maybe' (epistemic uses of the functional modal *pouvoir* occur after 4; Bassano 1996), and Polish children use the adjective *-chyba* 'probably' (Smoczynska 1993).

While we do not explicitly test conceptual development, we predict that we will see BCS lexical modals from age 2, as in other languages, providing further evidence against a solely conceptual trigger for epistemic language (see de Villiers 2007 for discussion).

1.2 The Grammatical Hypothesis for BCS

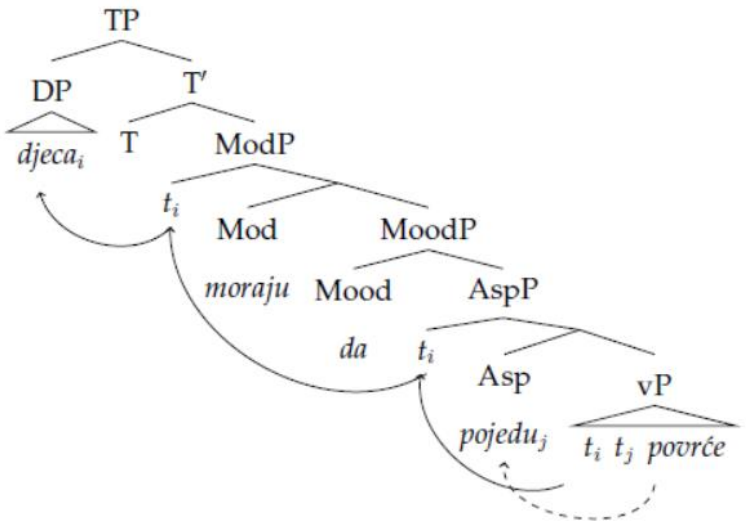
This hypothesis states that the EG occurs because children lack the grammatical representations needed to support epistemic interpretations of functional modal verbs (Cournane 2015, also Heizmann 2006; de Villiers 2007). These interpretations are argued to arise from syntactic structures more complex than needed for their root counterparts (e.g., Roberts 1985; Brennan 1993; Cinque 1999). The general consensus is that root modality is eventive, with the modal below T, while epistemic modality is propositional, and the modal is interpreted as scoping over T (e.g., Palmer 1986) and bound by the speech act event (Percus 2000). Following Hacquard (2006), we assume functional modal verbs are anaphoric to events, and have only one lexical entry (cf. Cinque 1999).

Cournane (2015) ran a corpus study of Sarah (2;3-5;1, Brown 1973; CHILDES, MacWhinney 2000) to test Hacquard (2006)'s analysis of functional modal verbs in English, where modal auxiliaries like *must* take non-finite complements. Cournane tested whether the development of TP-embedding (representative of embedding propositions) correlated with first epistemic functional modals. Sarah's first spontaneous use of such modals is at 3;0 (*must be gone*), soon after her first *to*-infinitive form on the second verb at 2;10 (*I want to see him*), and first embedded subject at 2;11 (*watch me do horsie)*. This is likely generalizable for English, as TP-embedding is reported to appear in the months leading up to 3;0 (de Villiers & Roeper 2016, i.a.) and research on the EG reports first epistemic uses of functional modals at age 3 (Papafragou 1998, i.a.).

For BCS, Veselinović (2017) argues that modal verbs, when root (3), have the structure in (4), and when epistemic (5), they have the CP-embedding structure in (6) (structures simplified). Note that (3) shows agreement on both the modal and the lexical verb, with the subject preceding the modal, and the lexical verb marked for perfective present, a form that needs a licensor (in this case, the modal) in the same clause. See Veselinović (2017) for more arguments for this analysis.

(3) Djeca$_i$ mora-ju DA t$_i$ po-jed-u povrće$_1$
children$_{NOM}$ must-$_{3PL.PRS}$ DA $_{PFV}$-eat-$_{3PL.PRS}$ vegetables
'The children must eat the vegetables.' (*root*)

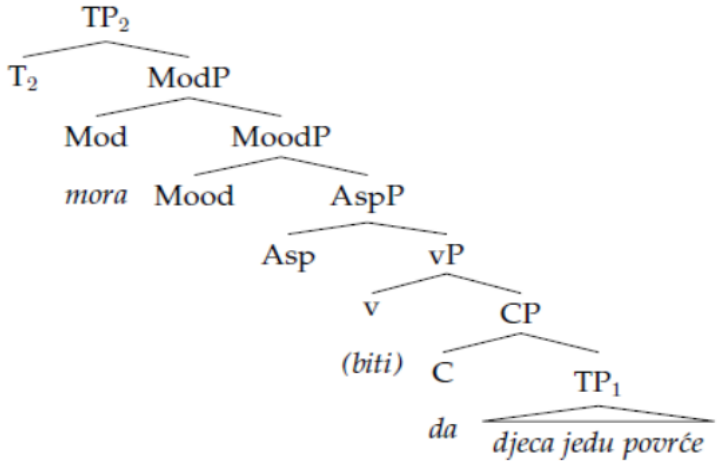
(4) Root modal structure



(5) Mora DA djeca$_i$ t$_i$ jed-u povrće
must-$_{PRS.3SG}$ DA children$_{NOMi}$ eat.$_{IPFV-PRS.3PL}$ vegetables
'The children must be eating the vegetables.' (*epistemic*)

[1] Within BCS, some dialects use infinitive MoodP here, primarily in the Northwest, including parts of Croatia and Bosnia (see Mišeska-Tomić (2006) for the distribution of infinitive and subjunctive within BCS). This does not affect the analysis, as monoclausal structures in those dialects derive root interpretations, and epistemic interpretations of *morati* 'must' and *moći* 'can' can still be derived from biclausal structures as in (5) and (6). For example, out of 56552 utterances in HrAL (*Croatian Adult Spoken Language corpus*, Kuvač Kraljević & Hržica 2016), with high dialectal variance, 4 utterances containing *morati* and 1 with *moći* are as in (6). Our analysis and discussion pertain to the dialects of BCS that use these modal verbs in both root and epistemic contexts, granting that not all dialects of the language have both uses.

(6) Epistemic modal structure



Since BCS modal verbs obligatorily show CP embedding for epistemic uses, unlike English, we can refine the Grammatical Hypothesis into two grammar-driven hypotheses. First, if representing epistemics depends on the ability to scope a modal above a proposition, represented by at least a TP in the syntax, we predict that the EG in BCS will resolve around 3;0, as in English (Cournane 2015). Second, if it depends on the input syntax, we predict that the EG in BCS children will last until CP-embedding emerges, around 4;0 cross-linguistically (de Villiers & Roeper 2016).

1.3 The Frequency Hypothesis

Finally, it is important that we test whether the EG is an effect of input frequency, as suggested by Shatz et al. (1983), Papafragou (1998) and O'Neill & Atance (2000). This is an important hypothesis, as epistemic uses form only ~8% of functional modal input in English (van Dooren et al. 2017, cf. Cournane 2015). Cournane found that the child she studied showed an EG for functional modals, and epistemic uses remain significantly lower than the input through to the end of the corpus (5;2). We test this hypothesis by examining all maternal input in the corpus. If frequency drives epistemic delay, we expect correlations between maternal rates of epistemic uses and child epistemic delay.

2 Methods

This study uses the SCECL corpus (*Serbian Corpus of Early Child Language*; Anđelković, Ševa, & Moskovljević 2001) from CHILDES (MacWhinney 2000). SCECL contains data from eight children, aged 1;6 to 4;0, gender balanced, half from Belgrade, Serbia (DAC, JEL, LUK, MIL) and half from Banja Luka, Bosnia and Herzegovina (ANA, ANE, LAZ, NIK). All children come from middle-class urban families with parents with at least secondary education. Recording occurred between 6/1998 and 12/2000, once every two months for 90 minutes, with additional 30 minutes at six month intervals. This yields 128 recordings, with 95,105 child and 72,305 mother utterances, focusing on mothers' speech as representative of the children's input.

To assess whether the EG exists in BCS, we extracted all child utterances containing any form of *moći* 'can' and *morati* 'must', with 16 lines of discourse (8 preceding, 8 following). The discourse context was examined to determine the interpretation of the modal as root or epistemic based on contextual and grammatical cues. If the discourse sampled was insufficient to determine this, we examined the situational context in the original file (i.e. non-verbal elements coded in the corpus).

To test the frequency hypothesis, we extracted all maternal input utterances with the collocations of *mora (biti) da* 'must (be.INF) DA' and *može biti da* 'can be.INF DA'. As with the child data, we use the discourse to code these uses of *moći* 'can' and *morati* 'must' as root or epistemic. We assumed that uses of *moći* and *morati* outside of these constructions have root meanings, as adult speakers find them ungrammatical in epistemic uses (Veselinović, 2017). This conservative choice may underestimate the rate of epistemic uses of modal verbs in the input.

To test the conceptual hypothesis, we extracted epistemic modal adverbs (*možda* 'maybe', *valjda* 'probably', *sigurno* 'surely') from the children's corpora, checking for epistemic contexts as with modal verbs.

To test both versions of the grammatical hypothesis, we looked for evidence of TP- and CP- embedding in BCS children and their maternal input. For evidence of TP-embedding, we looked for *V+DA* collocations, where DA is a mood marker, as in (2) (see Browne 1986, Mišeska-Tomić 2003, i.a. for arguments for non-C DA in Mod/Mood). We chose *htjeti* 'want' as the verb, as *want* is used early with TP-embedding by English-speaking children (Shatz & Wilcox 1991, a.o.). We searched the corpora

from the beginning (1;06) until we found sustained use of *htjeti* + DA, omitting the uses of *htjeti* with non-TP complements from consideration.

On strict criteria, we assume this to be minimum necessary evidence of CP-embedding: the embedding verb would need to be non-imperative and followed by an overt complementizer DA. This is not sufficient, as some TP-embedding structures meet that requirement, but given the nature of the work, we accept such structures as CP embedding in the strict sense. Again, this is conservative, as we only possibly accept non-CP-embedding structures, and do not reject CP-embedding ones.

To find evidence of CP-embedding, we found and extracted all utterances of typical CP-embedding verbs *reći* and *kazati*, both meaning 'say/tell', with 5 utterances before and after the target. We coded complement types (null, nominal, adverbial, CP, direct speech, other[2]) for each target. We coded *to* 'that', *nešto* 'something', *šta* 'what' or accusative pronouns (7), as nominal complements, and *kako* 'how', *ovako* 'this way' and manner adverbs as adverbial complements (8).

(7) *CHI: reć(i) (ć)u te tati.[3]
tell$_{INF}$ will$_{1SG}$ you$_{ACC}$ dad$_{DAT}$
'I will tell on you to Daddy.' (LAZ, 2;08)

(8) *CHI: pa kako, tako ti meni reci .
well how that-way you$_{NOM}$ me$_{DAT}$ say$_{IMP}$
'Well how, you tell me that way.' (ANA, 3;02)

We coded utterances as having a null complement to the embedding verb when there was nothing overt that could be analyzed as the verbal complement, or if only the indirect object was present. These were often imperatives, or utterances like *Rekla sam ti!* (I told you!). Utterances like (9) were also coded as having null complements, as the complement is dislocated and the utterance can be analyzed as a two sentence sequence.

[2] Not to imply we believe the adverbs or nominals are complements here. However, the children's grammar is not necessarily adult-like, and all they need to produce these constructions are adverbial adjuncts or nominal complements.

[3] Only clear spontaneous uses will be reported throughout the paper.

(9) *MAJ: a kol(i)ko me voliš nis(i) mi rekla.
and how-much Iacc love2SG NEG-be2SG meDAT tellPPT.F.SG
'You didn't tell me how much you love me.' (ANA, 2;02)

Direct speech complements were not coded as CPs because early uses of *reći*/*kazati* 'say/tell' involve utterances like *krava kaže mu* 'cow says moo' and null complements, including non-imperative forms with null complements. Utterances like (10) provide only equivocal evidence for CP embedding. We thus coded these utterances as a separate category.

(10)*CHI: rekla mi baba [:hoćeš] li na [:sankanje].
sayPPT.F.SG IDAT grandma want2SG.PRS Q on sledding
'Grandma told me: "Wanna go sledding?"' (JEL, 3;0)

Finally, we coded the complements as CPs when the verb was followed by wh-questions (11a), yes/no questions (11b), or clauses introduced by complementizer DA (11c). Utterances like (11a) and (11b) can be viewed as sequences of two CPs, especially with imperative matrix verbs. However, we wanted to err on the side of caution and find the earliest embedded CP, rather than narrow the search to utterances containing DA.

(11)a. *CHI: kaži šta si jela.
sayIMP what be2SG.PRS eatPPT.F.SG
'Say what you ate.' (JEL, 3;00)
b. *CHI: reci meni jel ti [:imaš] [:žvaku].
tellIMP meDAT is-Q you have gum
'Tell me, do you have gum?' (ANE, 2;10)
c. *CHI: mama, Ija, q:@fp, Ija kaže da sam ja glupača.
Mom Ija Ija says DA am I dummy
'Mom, Ija says that I am a dummy.' (ANA, 3;02)

To test if children acquire CP-embedding structures concurrently with *say/tell*, we follow Snyder (2007) and Cournane (2015) and use the binomial test for concurrent acquisition[4]. This tests the hypothesis that

[4] $p = (X / (X + Y))^Z$, where X is the number of times the verbs *reći*/*kazati* 'say/tell' are used with a non-CP complement in the recordings following their first use with a

the proportional use of CP-embedding structures in a child's speech after the first appearance is such that the prior zero rate of use is unsurprising. A non-null result refutes this, suggesting that the delay is unexpected if the CP-embedding uses were acquired concurrently with others.

5 Results

Of the 95,105 child utterances in SCECL, 2110 contain *moći* 'can' and 261 contain *morati* 'must'. All the children start using these modal verbs between 1;08 and 2;04, consistent with first child uses in other languages (Papafragou 1998, i.a.). Earliest uses in SCECL are mostly one- or two-word utterances with *moći* (14), which is more frequent than *morati* (15) for all children. Maternal input contains 72,305 utterances, 1958 with *moći* and 494 with *morati*. Five utterances with *moći* (2 mothers), and 18 with *morati* (4 mothers) are used in epistemic contexts.

(7) *DAR: ajde dohvat-i. *CHI: ne možem.
come-on reach-IMP NEG can1SG.PRS(overgeneralized)
'Come on, reach it.' 'I can't.' (ANA, 1;08)

(8) *CHI: mo:ram da ga popravim .
must1SG.PRS DA itACC fix1SG.PRS
'I must fix it' (*pretending to fix a toy tractor*) (LUK, 2;04)

Crucially, no child utterances contain epistemic uses of modal verbs, suggesting that in BCS the EG lasts at least until 4;0. This differs from English children, whose EG resolves around 3;0 (Papafragou 1998, i.a.).

CP complement, Y the number of times they are used with CP complements in those recordings, and Z the number of times they are used with non-CP complements in the recordings prior to the first clear use with a CP complement.

5.1 Conceptual hypothesis: lexical modal results

All the children except ANE use epistemic adverbs *možda* 'maybe' or *valjda* 'probably'; LUK uses both. ANA and NIK sporadically use *sigurno* 'surely'. Rates match those of English children for *maybe* and *probably* (O'Neill & Atance 2000; Cournane 2015). A summary of results is in Table 1, with examples in (16)-(18). Note that ANA's uses include 7 uses of variants of a semi-fixed *nije valjda* (it can't be).

Child	**First clear use (age)**	**Total uses**	**Child**	**First clear use (age)**	**Total uses**
LUK	2;04	10	JEL	3;06	1
ANA	2;06	14	DAC	3;08	3
NIK	2;10	9	MIL	4;00	1
LAZ	3;02	2			

Table 1: Child uses of epistemic modal adverbs

(16) *CHI: ko lupa?
who thumps?
*NAD: ne znam.
NEG know
*CHI: Đuđa možda.
Đuđa maybe (LUK, 2;04)

(17) *MAJ: a šta radiš ovdje, ko je ovo s tobom?
and what do2SG.PRS here who is this with you
*CHI: a:@i, moj medo valjda .
my bear probably (ANA, 3;08)

(18) *SBA: evo medvjed, a ovca nestala.
here bear and sheep disappeared
*CHI: sigurno je ovca ovamo, iza medvjeda .
surely is sheep here behind bear (NIK, 4;0)

Fig. 1 shows that children use epistemic language during the EG, but they fail to use all the strategies used by adults.

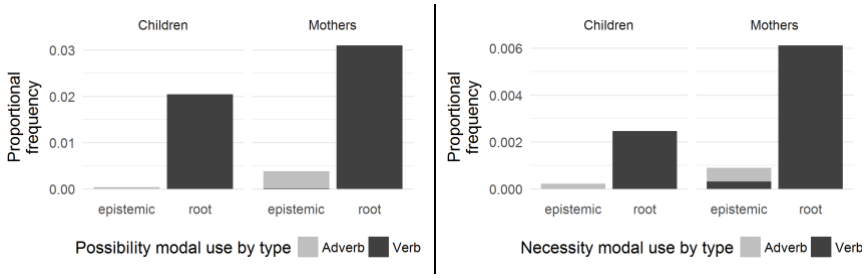


Figure 1: Aggregate mean usage of possibility (left) and necessity (right) modal verbs and adverbs by children and mothers to express epistemic and root modal flavors. NB: root modal adverbs such as *obavezno* 'obligatorily' not included.

5.1 Frequency Hypothesis: Input results

As no children in SCECL resolve their EG for the duration of the corpus, we could not use the binomial test for concurrent acquisition. We assess instead whether epistemic uses of modal verbs are less frequent in the BCS input than in English, where they form ~5% of all modal utterances. If so, the frequency could explain the cross-linguistic difference. Moreover, if the lack of epistemic uses of modal verbs in the input conditions their absence in the children's speech, we should expect the rate of root use of modal verbs in the input to be conditioning the rate at which they are acquired. As all the children have acquired root uses of modal verbs, this is a testable prediction, which we assess using mixed-effects models. Finally, if the epistemic uses of modal verbs are present, but infrequent, in the input, and the children reach adult-like frequencies of non-epistemic uses of modal verbs, we could expect adult-like frequencies of epistemic uses of modal verbs as well.

For each child corpus, the total number of utterances (TNU) is between 10,000 and 12,000, with one outlier at 17,000. Mothers in SCECL are much more variable, with TNUs ranging between 2600 and 19,000. We calculated proportional frequencies of epistemic modal verbs to total modal verbs, to assess whether mothers use modal verbs in epistemic contexts at rates similar to English adults (~5% of modal verbs). For the 5 BCS mothers who show epistemic uses of modal verbs[5], the average proportion is 6.44% for *morati* 'must' and 0.8% for *moći*

[5] These five mothers' TNUs are >9400, while TNUs of the other mothers are <5200.

'can'. It is unlikely that similarly low adult inputs differentially predict first child epistemic uses: English after 3;0, but BCS after 4;0.

The mixed-effects model for the children's frequency of use *moći* 'can', with a fixed effect for TNU and random intercept for subjects showed that age is a significant predictor, increasing the frequency by 8.51 ($\chi 2(1)=33.697$, $p<0.0001$). Having shown this, we used age as a fixed effect, along with a random intercept for subjects, to see if a mother's usage frequency of root possibility modal predicts a child's usage frequency, and found no significant effect. An increase by 1 in mother's usage frequency increases the child's usage frequency by 0.094 ± 0.052 ($\chi 2(1)=3.372$, $p=0.066$).[6] Similarly for the root necessity modal verb, where a child's age is a significant predictor, although the rate of increase is negligible ($\chi_2(1)=33.552$, $p<0.0001$), which is expected if we keep in mind that the model is applied to all the data (for the sake of uniformity) and the children use the necessity modal verb later and less frequently than the possibility modal verb. As was shown for *moći* 'can', the mother's frequency of use of the root necessity modal verb does not significantly affect the child's frequency of use of the same verb, increasing it by 0.037 ± 0.049 ($\chi 2(1)=0.6036$, $p=0.44$).

Taking the average proportional frequency across all mothers and all recordings to be the best proxy for adult-like use[7], we see in Fig. 2 that the average child proportional frequency of *moći* 'can' shows a steady increase over time, reaching the adult-like rate of 3% of all utterances at about 3;06. Fig. 3 shows that the first appearance of *morati* 'must' is delayed, and only reaches the adult-like rate of 0.6‰ of all utterances at 4;0. Figs. 2 and 3 also show the average child frequencies of the possibility and necessity modal adverbs, neither reaching adult-like rates.

[6] Just in case, we ran the same model with TNU as a fixed effect, and we found a significant, but minimal effect: an increase by 1 in mother's frequency of use increases the child's frequency of use by 0.16 ±0.057, $\chi_2(1) =7.925$, $p=0.004$.

[7] We used only root modal verbs to calculate this, as those are the only child uses.

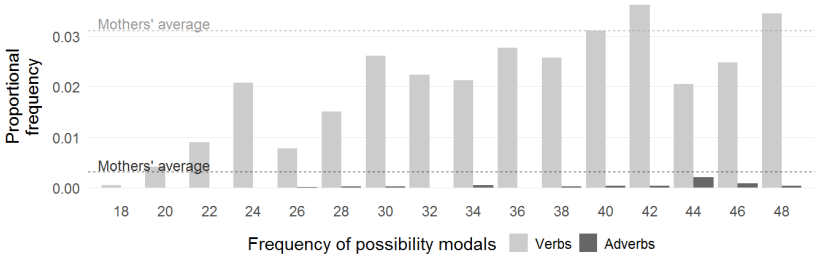


Figure 2: Average child usage frequency, at each point of recording (1;06-4;0), of the possibility modal verb *moći* 'can' and adverbs *možda* 'maybe' and *valjda* 'probably', compared to the 'adult-like frequency'.

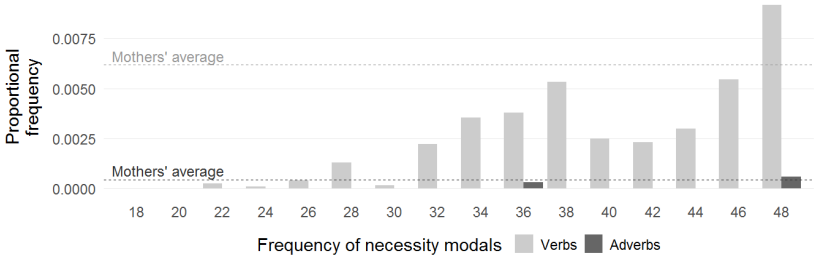


Figure 3: Average child frequency of use of the necessity modal verb *morati* 'must' and adverb *sigurno* 'surely' compared to the 'adult-like frequency' at each recording point (1;06-4;0).

5.1 Grammatical Hypothesis: syntactic results

Our starting point for evidence of TP-embedding in BCS was strict: the inflected embedding verb *htjeti* 'want' followed immediately by DA. This yielded first uses shown in (19). While earlier TP-embedding may exist, with infinitival complements or verbs other than *htjeti* 'want', we see that most children have the first of repeated uses (FRU)[8] between 2;06 and 3;02, consistent with de Villiers & Roeper's (2016) report for English, where children between 2;0 and 3;0 start using non-finite complement clauses, followed shortly by finite ones. In BCS, the exceptions to this are MIL, whose FRU occurs at 3;08, and ANE, who doesn't have

[8] Based on Snyder (2007)'s *first of repeated uses*, FRU denotes the first use followed by a repeated use in the following recording.

repeated uses across consecutive recordings, but has 8 clearly distinct uses at 3;0, followed by uses in every other recording until 4;0. If TP-embedding were sufficient for children to represent epistemic uses of modal verbs, as Cournane (2015) argues for English, we would expect the BCS children to use modal verbs in epistemic contexts shortly after first using TP-embedding. We see no epistemic modal verbs for BCS children, not even those who use TP-embedding early, which leads us to reject the TP-embedding version of the grammatical hypothesis.

(19)a. *CHI:(h)oćemo da se igramo (.) ovog?
want1PL.PRS DA SE play1PL.PRS thisGEN
'Shall we play with this?' (LUK, 2;06)

b. *CHI:mama, [:hoću] da vidim kako da nađem[?].
Mom want1SG.PRS DA see1SG.PRS how DA find1SG.PRS
'Mom, I want to see how I can find…' (NIK, 2;10)

c. *CHI:[:hoćeš] da vidiš koji bakin broj?
want2SG.PRS DA see2SG.PRS which grandma's number
'Wanna see what grandma's number is?' (ANE, 3;0)

No child produces *reći* and *kazati* 'say/tell' before 2;0, and no child uses CP-type complements before 2;04. When CP-type complements appear, child rate of use (even with broad criteria) stays at an average of 16% of utterances with *reći* and *kazati*, compared to 12% to 42% (24% avg.) of such utterances for mothers. The mean frequency of such constructions across all utterances is 0.04% for children, but 0.7% for mothers (Fig. 4).

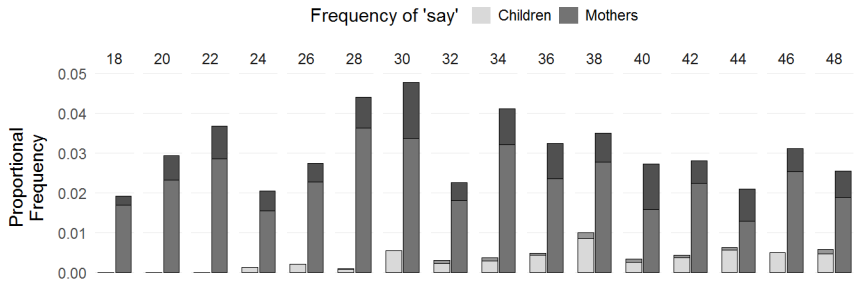


Figure 4: Use of *reći/kazati* 'say/tell' by complement type. Each pair of bars shows the average proportional frequency for children (light) and mothers (dark) at child age (x-axis). The darker top portion of each bar, if present, depicts use of CP complements, compared to other complement types combined.

			CP Embedding		
	First use	FRU	First CP-embedding	FRU CP-embedding	Total CP-embedding
ANA	2;00	2;00	2;10	2;10	17
ANE	2;04	2;04	2;04	2;08	10
LUK	2;00	2;06	2;06	2;06	9
JEL	2;08	3;00	3;04	3;04	7
LAZ	2;02	2;06	2;08	N/A	2
DAC	2;06	2;06	3;02	N/A	1
NIK	2;02	2;10	3;02	N/A	1
MIL	3;04	3;04	N/A	N/A	0

Table 2: Acquisition of *reći*/*kazati* 'say/tell' for each child in SCECL.

Table 2 shows the progression from first use of *reći*/*kazati* 'say/tell' to first repeated use of CP-embedding constructions with those verbs for each child. Only four children have repeated uses of CP-embedding constructions, and among them, ANE has uses at 2;08 and 2;10, then no uses until 3;06, and LUK has consistent use between 2;06 and 3;04, but no later utterances. For each child who shows repeated uses of CP-*say*, we ran a binomial test for concurrent acquisition (Snyder 2007), to test if the rate of use of CP-embedding 'say' before the first use is expected to be zero and found the likelihood of zero use to be $p<0.0001$.

6 Discussion

We show that the EG, which ends around 3;0 in English children, is protracted in BCS children until after 4;0. Regarding lexical epistemic modals, the results for BCS-learning children align with the reports for children learning English, French and Polish (O'Neill & Atance 2000, Cournane 2015, Bassano 1996, Smoczynska 1993). Seven of the eight BCS children use modal adverbs *možda* 'maybe', *valjda* 'probably' or *sigurno* 'surely' in contexts compatible with epistemic meanings, with use comparable to that of English-speaking children (Cournane 2015).

We see that despite BCS and English-speaking children having similar acquisition patterns of root uses of modal verbs and epistemic modal adverbs, we find different patterns for epistemic uses of modal verbs, which English-speaking children start using between 3;0 and 3;06, but BCS children do not produce before 4;0. This differential acquisition time across syntactic categories, within and across languages, provides

further evidence against the conceptual hypothesis. A purely conceptual account of epistemic uses would have trouble trying to explain why BCS children are delayed by a year compared to their English counterparts.

Concerning the frequency hypothesis, for 5 of the children the maternal input proportional frequencies of epistemic uses of modal verbs are similar to those for English, making it unlikely that they would differentially predict the children's time of acquisition of epistemic constructions by as much as a year. To further assess the frequency hypothesis, given that no BCS child used an epistemic modal verb, we tested the frequencies of the root uses alone to see if modal input rates affect acquisition time for modal verbs with root meanings. We found that mothers' usage frequencies alone cannot predict when BCS children will attain adult-like usage for root modal verbs. The BCS children reach adult-like frequencies of use of root modal verbs for both the possibility and the necessity modal verbs, but epistemic uses remain conspicuously absent. We thus rule out input frequency as explanatory of the EG.

The grammatical hypothesis, as put forward in Cournane (2015), predicts that BCS children use modal verbs epistemically as soon as they acquire TP-embedding. However, since we found TP-embedding, as in English, but no epistemic uses of the modal verbs until at least 4;0, we rule out this version of the grammatical hypothesis. Our modified grammatical hypothesis, which takes into account the syntactic differences between English and BCS epistemic uses of functional modal, can account for the data. BCS epistemic uses rely on CP-embedding, and the milestone for acquiring CP-embedding is around 4;0 (de Villiers & Roeper 2016). Further work is needed to determine when BCS children first use epistemics functional verbs. We predict that the SCECL corpus just misses first uses, which should occur soon after 4;0.

This research also speaks against an analysis where epistemic modal adverbs and epistemic modal auxiliaries and verbs are all generated as specifiers of the same functional projection (Cinque 1999). It is unclear why a child who is able to represent verbal elements elsewhere in the syntax and also able to represent Cinque's Mod_{epi}P, would be able to represent adverbial elements as specifiers of this functional head, but not verbal ones. Unlike approaches where the position of functional modals conditions their interpretation (Hacquard 2006; Veselinović, 2017), Cinque (1999)'s approach states that it is the epistemic interpretation of

modal elements that conditions their position, wrongly predicting that children should acquire epistemic modal verbs and adverbs concurrently.

An important difference between English and BCS functional modals warrants further exploration. When the English children resolve their EG, the first modal verb they use in epistemic contexts is *might* for 3 of the children Cournane (2015) examined, and *must* for the fourth. *Might* is almost exclusively epistemic in English (Hacquard & Wellwood 2012), and *must* is also largely epistemic in adult English (van Dooren et al. 2017). BCS children are faced only with functional modal verbs with predominantly root uses, potentially contributing to their prolonged EG.

Further cross-linguistic work is needed to refine the language-specific grammatical hypothesis we put forth here on the basis of BCS and English evidence. The language-specific grammatical hypothesis predicts acquisition patterns to differ depending on the syntax of the input modals, including whether the variable meaning modal verbal elements are verbs or auxiliaries. Bassano (1996) suggests this may be the case, as epistemic uses of *pouvoir* 'can' are not acquired before 4;0 in French, while *devoir* 'must' is not used epistemically before 3;3, and only 3 times after that (prop.freq: 0.0005). It is also possible that the children we do see using CP complements to *reći/kazati* 'say/tell' are treating these as TP complements (see Diessel & Tomasello 2001 for similar arguments for English sentential complements). Both types of constructions involve inflected embedding verbs followed by DA and both can involve distinct subjects of the two verbs. Experimental work is underway testing child comprehension and production of both epistemic modal verb constructions and biclausal (CP-embedding) constructions.

References

Anđelković, Darinka, Nada Ševa, & Jasmina Moskovljević, 2001. *Serbian Corpus of Early Child Language*. Laboratory for Experimental Psychology, Faculty of Philosophy, and Department of General Linguistics, Faculty of Philology, University of Belgrade.

Astington, Janet W., Paul L. Harris, & David R. Olson, 1990. *Developing Theories of Mind*. Cambridge: CUP.

Bassano, Dominique. 1996. Functional and Formal Constraints on the Emergence of Epistemic Modality: A Longitudinal Study on French. *First Language*, 77-113.

Brennan, Virginia. 1993. *Root and Epistemic Modal Auxiliary Verbs*. Amherst, MA: University of Massachusetts.

Brown, Roger W. 1973. *A First Language: The Early Stages*. Cambridge, MA: Harvard University Press.

Browne, Wayles. 1986. *Relative Clause in Serbo-Croatian in Comparison with English*. Zagreb: Institute for Linguistics, Faculty of Philosophy.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford: OUP.

Cournane, Ailís. 2015. Revisiting the Epistemic Gap: Evidence for a Grammatical Source. *BUCLD 39 Proceedings* pp. 127-140. Somerville, MA: Cascadilla Press.

Cummins, Denise. 2013. Deontic and Epistemic Reasoning in Children Revisited: Comment on Dack and Astington. *Journal of Experimental Child Psychology*: 116, 762-769.

De Villiers, Jill. 2007. The Interface of Language and Theory of Mind. *Lingua*: 117: 11, 1858-1878.

De Villiers, Jill. G. & Tom Roeper. 2016. The Acquisition of Complements. In William Snyder, & Jeffrey Lidz Eds., *Oxford Handbook of Developmental Linguistics*. OUP.

Diessel, Holger, & Michael Tomasello 2001. The Acquisition of Finite Complement Clauses in English: A Corpus-based Analysis. *Cognitive Linguistics*: 12, 97-141.

Hacquard, Valentine. 2006. *Aspects of Modality*. PhD Dissertation, MIT.

Hacquard, Valentine. 2013. Grammatical Category of Modality. *Proceedings of the 19th Amsterdam Colloquium*, pp. 19-26.

Hacquard, Valentine & Alexis Wellwood 2012. Embedding Epistemic Modals in English: A Corpus-based Study. *Semantics and Pragmatics*: 5: 4 , 1-29.

Heizmann, Tanja. 2006. Acquisition of Deontic and Epistemic Readings of Must and Müssen. In Tanja Heizmann Ed., *UMOP 35: Current issues in language acquisition*. Amherst, MA: GLSA, UMass.

Kratzer, Angelika. 2012. *Modals and Conditionals*. Oxford: OUP.

Kuczaj, Stan, & Michael Maratsos. 1975. What Children Can Say Before They Will. *Merrill-Palmer Quarterly of Behavior and Development*: 21, 89-111.

Kuvač Kraljević, Jelena & Gordana Hržica. 2016. Croatian Adult Spoken Language Corpus (HrAL). *Fluminensia: Journal for philological research*: 28: 2.

MacWhinney, Brian. 2000. *The CHILDES Project: Tools For Analyzing Talk*. Mahwah, NJ: Lawrence Erlbaum.

Mišeska-Tomić, Olga. 2003. The Syntax of the Balkan Slavic Future Tenses. *Lingua*: 114, 517-549.

Mišeska-Tomić, Olga. 2006. *Balkan Sprachbund Morpho-Syntactic Features.* Dordrecht: Springer.

O'Neill, Daniela K, & Cristina M. Atance. 2000. "Maybe my daddy give me a big piano": The Development of Children's Use of Modals to Express Uncertainty. *First Language*: 29: 52 , 29-52.

Palmer, Frank. 1986. *Mood and Modality*. Cambridge: CUP.

Papafragou, Anna. 1998. The Acquisition of Modality: Implications for Theories of Semantic Representation. *Mind & language*:13:3, 370-399.

Percus, Orin. 2000. Constraints on Some Other Variables in Syntax. *Natural Language Semantics*: 8, 173-229.

Rett, Jessica, & Nina Hyams. 2014. The Acquisition of Syntactically Encoded Evidentiality. *Language Acquisition*: 21, 173-198.

Roberts, Ian. G. 1985. Agreement Parameters and the Development of English Modal Auxiliaries. *Natural Language & Linguistic Theory*: 3, 21-58.

Shatz, Marilyn., & Sharon A. Wilcox 1991. Constraints on the Acquisition of English Modals. In Susan A. Gelman, & James P. Byrnes Eds., *Perspectives on language and thought: Interrelations in Development* pp. 319-353. Cambridge: CUP.

Shatz, Marilyn, Henry. M. Wellman, & Sharon Silber. 1983. The Acquisition of Mental Verbs: A Systematic Investigation of the First Reference to Mental State. *Cognition*: 14, 301-321.

Smoczynska, Magdalena. 1993. The Acquisition of Polish Modal Verbs. In Norbert Dittmar, & Astrid Reich Eds., *Modality in Language Acquisition* pp. 145-169. Berlin: De Gruyter.

Snyder, William. 2007. *Child Language: The Parametric Approach.* Oxford: OUP.

Stephany, Ursula. 1979. Modality. In Paul Fletcher, & Michael Garman Eds., *Language Acquisition* (2nd edition 1986) pp. 375-400. Cambridge: CUP.

Stephany, Ursula. 1993. Modality in First Language Acquisition: The State of the Art. In Norbert Dittmar, & Astrid Reich Eds., *Modality in Language Acquisition* pp. 133-144. Berlin: De Gruyter.

van Dooren, Annemarie, Anouk Dieuleveut, Ailís Cournane and Valentine Hacquard (2017). Learning What 'Must' and 'Can' Must and Can Mean. *Proceedings of the 21st Amsterdam Colloquium*.

Veselinović, Dunja. 2017. Structural Differences Between Epistemic and Root Modality: Evidence from BCS. *Proceedings of CLS52 (2016)*. Chicago: Chicago Linguistics Society.

dunja@nyu.edu
cournane@nyu.edu