
In defence of the child innovator

AILÍS COURNANE

2.1 Introduction

For over a hundred years researchers have appealed to the child innovator to explain language change (Meillet 1912; H. Paul 1920; Halle 1964; Andersen 1973; Kiparsky 1974, i.a.). Today, first language acquisition (L1A) is central to generative theories of change (e.g. Lightfoot 1979; Roberts and Roussou 2003; van Gelderen 2004, 2011a). I will call this approach the *Child Innovator Approach* (CIA), in contrast to approaches that place explanatory weight on adult speakers (Bybee and Slobin 1982a; Heine *et al.* 1991; Ziegeler 1997; Givón 2009; Diessel 2011, 2012, i.a.). Proponents of the CIA argue that innovative child analyses of the input language, if and when they survive and spread, become the changes reflected in the historical record. The CIA developed from empirical observations of historical data, but places the burden of explanation for diachronic patterns on the L1A process. Despite this, the consequences for language acquisition have rarely been considered beyond the anecdotal (though see Clark and Roberts 1993; Weerman 1993, 2011; van Gelderen 2011a: 21–6; Meisel *et al.* 2013), and even more rarely assessed using targeted L1A studies and child data (Baron 1977; Cournane 2014, 2015a).

In this chapter, I address two common criticisms of the Child Innovator Approach that persist in the (near-)vacuum of communication between researchers working on L1A and historical linguistics. The first is that the CIA is untenable since input-divergent¹ analyses that occur during L1A must survive into adulthood. While everyone agrees children make input-divergent analyses in the process of learning, critics of the CIA argue these resolve prior to a time when diffusion is possible (adulthood;

¹ I use the term *input-divergent* to refer to any analysis by the child that does not conform to the grammars of the speakers who comprise the child's input. Input-divergent analyses are normally referred to as child 'errors'. The term 'error', though convenient, is problematic for many reasons, perhaps the most egregious of which is that these apparent errors only become erroneous when compared to the 'correct' analyses of other people's grammars; when analysed within the grammar which produces them (i.e. the child grammar at its relevant state) then they are fully consistent. The term 'error' also carries a judgement of inaccuracy relative to some goal that is not consistent with a descriptive and non-teleological view of language development.

e.g. Traugott and Dasher 2005; Diessel 2012: 1600). In Section 2.2, I respond to this criticism by drawing attention to bodies of evidence for (a) peer-to-peer acquisition throughout childhood, (b) bilingual and heritage language effects on L1A that prolong input-divergent analyses, and (c) longstanding sociolinguistic evidence that the teenage years are most relevant for the spread of innovations (as well as evidence that even pre-adolescent children likely participate in advancing sociolinguistic changes).

The second common criticism comes from the assumption that parallels must hold between child innovations and diachronic innovations ('ontogeny recapitulates phylogeny') in order for the CIA to hold. We see parallel alignments in some domains (e.g. morphological regularization, semantic extensions), but not others (constructional reanalysis, as from biclausal structures to monoclausal, the development of agreement markers from pronouns). The heterogeneity of alignments between L1A pathways and diachronic pathways across linguistic domains is taken to be evidence against the CIA (e.g. Diessel 2012). In Section 2.3, I argue that the CIA does not predict that the L1A process should mimic completed diachronic changes. Rather, the CIA predicts that contemporary children make input-divergent analyses that advance diachronic cycles (which we can look back on in the future).² In an effort to capture why semantic changes show parallels but morphosyntactic ones do not, I formulate the child language predictions of the CIA in terms of the *mapping problem* (see also Courneane 2014). The mapping problem is a primary problem in L1A, referring to how children learn which linguistic forms to associate with which meanings (entities, events, properties, etc.; see E. Clark 1977, 1993; Gleitman 1990; Bloom 2000). Broadly, parallel alignments occur when the child posits a new meaning for an extant form (e.g. overextension), while oppositional alignments occur when the child posits a new form for an extant meaning (parsimonious or conservative structure postulation). This approach stresses that different necessary processes in typical L1A give rise to different types of innovations.

2.2 The survival and diffusion of L1 input-divergent properties

For the CIA to be supported, input-divergent analyses during L1 acquisition must diffuse to other speakers and become the changes we see in the historical record. For this to happen the child must either succeed in spreading her input-divergent analyses to mature speakers during childhood or fail to correct input-divergent properties and spread innovations as an adult. The first possibility is seen as problematic because children are not influential members of the speech community (older speakers will not model their speech after children) while the latter is problematic because children are excellent language acquirers who make few possible errors (Snyder 2007, i.a.) and resolve those errors that they do make prior to adulthood (see Kerswill 1996). Thus, the CIA is untenable because children cannot spread errors as children and errors resolve by adulthood when they theoretically could spread (Diessel 2012). I argue that this

² In other words, we expect children to make new changes, not repeat old ones. For example, overextend current spatial terms like *beside* to temporal meanings, rather than use, e.g., *after* first with spatial meanings then later extend it to temporal meanings.

assessment misconstrues the sociolinguistic evidence and overestimates the swiftness and accuracy of child development, especially in multilingual or contact situations. I appeal to three lines of evidence: (a) children learn from other children via peer-to-peer acquisition, (b) adolescents diffuse innovations, not adults, and (c) multilingual learning conditions may prolong divergent features of first language acquisition.

First, the notion of social prestige is important to clarify. The idea that children lack the social prestige necessary to be linguistically influential is sometimes taken as an argument against the CIA (e.g. Diessel 2012: 1600, who cites Labov 2001). However, prestige applies within all groups, not only in adulthood or with those that hold conventional power (see Labov 2001, 2012 for discussion). Children, pre-adolescents, and adolescents may all hold power roles and differing influence within their groups.

Children begin talking like their peers (instead of like their parents) from earliest socialization, approximately kindergarten age in Western societies (e.g. Labov 2012; Cekaite *et al.* 2014). Taking peer-to-peer acquisition seriously, it is possible that young siblings or school-age children diffuse and reinforce input-divergent analyses in their grammars. Children with similar input show similar acquisition patterns so reinforcement may be particularly plausible. Even if we assume a critical period for syntactic or semantic acquisition (perhaps ending at puberty), input-divergence may spread to peers, or overlapping input-divergent analyses may be reinforced among peers, within the childhood years.

Sociolinguistic studies in both real and apparent time repeatedly show that it is primarily adolescents, not adults, who embrace and spread innovations ('lead sociolinguistic changes'; e.g. Labov 2001, 2007). For example, quotative *be like* (*He was like 'X'*) was embraced first by teenage girls (e.g. Tagliamonte and D'Arcy 2004), as with most changes. If we assume innovative variants originate in L1A input-divergence, then the sociolinguistic evidence significantly decreases the timespan within which input-divergent properties of the I-language need to survive before they spread to the speech community. Further, as input-divergent properties are internally conditioned by the input and learning mechanisms, children and teens could have the same input-divergent properties and reinforce each other, as with peer-to-peer acquisition. Thus, for the CIA to be supported, input-divergence need only survive to adolescence, possibly only to pre-adolescence (see e.g. Levey 2006 for evidence that pre-adolescents may lead sociolinguistic change).

L1A is often remarkably accurate when measured against adult speakers of the input, but it is also persistently and systematically inaccurate in some areas despite the input (e.g. derivational morphology usage and comprehension continues to develop through the early teen years, Derwing and Baker 1986). Even in monolingual learners with relatively uniform linguistic input, child input-divergence has been demonstrated to persist in areas like semantic development of modals until at least 12 years old (see Papafragou 1998). More commonly, divergent patterns have been shown to persist into at least late childhood (age 7 or 8). For example, comprehension studies with 6–9-year-old children show persistent non-adult scope biases (van Koert *et al.* 2015), and overextensions of direct evidential marking (de Villiers *et al.* 2009), to give just two semantic examples. These properties of first language acquisition are a plausible source of innovations even in monolingual populations (which cannot be taken

to be without variation), but are likely exacerbated in bilingualism contexts. These should still be taken to be instances of first language acquisition (in the case of heritage languages, or 2L1). In bilingual children evidence shows that certain areas of the grammar may be susceptible to prolongation (relative to monolinguals) of divergence periods (see Gathercole and Thomas 2009). Critically, these are cases of first language acquisition that show delayed (but not deviant) development relative to monolingual L1 development. This prolonged divergence may be another means through which divergent analyses persist long enough to be reinforced and/or diffuse.³

In sum, child input-divergent analyses are plausibly mutually reinforced during childhood through peer-to-peer acquisition. Furthermore, input-divergent analyses need only persist into the early teenage years when they can be picked up by the sociolinguistic change powerhouse of teenage peer groups. The variables that diffuse in the teenage years still need to originate from somewhere (the actuation problem, Weinreich *et al.* 1968). Even in monolingual populations, input-divergence sometimes persists into late childhood or pre-adolescence, a property of L1 acquisition that is likely exaggerated in bilingualism contexts,⁴ lending further plausibility to this avenue of research. The debate over whether input-divergent child analyses may diffuse into the speech community is far from settled.

2.3 Unidirectionality of diachronic pathways and the mapping problem

Proponents and critics alike regularly see parallel development between ontogeny and phylogeny as a necessary condition for the CIA, for example: ‘if language acquisition is the source of diachronic change, child language should include *the same types of changes and developmental patterns* as the diachronic evolution of language’ (Diessel 2012: 1600, emphasis mine). However, for the CIA to be supported by child patterns, child input-divergence needs to advance new diachronic pathways, not mimic (or parallel) completed ones. While the links between L1A and change are perhaps more obvious in parallel alignments (discussed in Section 2.3.1), I show that both parallel alignments and oppositional (= nonparallel) alignments between L1 development and diachronic development can be reduced to two sides of the form \leftrightarrow meaning mapping problem. On one side, children attribute new meanings to existing forms as part of the normal process of meaning extension (and parallel pathways occur), while on the other, children attribute new forms to meanings as part of the normal process of developing grammatical complexity (and oppositional pathways occur).

The diachronic domain of inquiry that forms the empirical basis for my arguments is the development of modal expressions in child language (Kuczaj 1977; Stephany

³ Cf. Weerman (1993, 2011); Meisel *et al.* (2013), who argue that sequential bilinguals (age of onset at age 4 or older) as well as L2 speaker input in the primary linguistic data are the primary triggers for morphosyntactic changes. If L2 speakers make up much of the input then the child’s grammar will correct the inconsistencies (compare with creole formation or language creation, as with Nicaraguan Sign Language, e.g., Senghas and Coppola 2001).

⁴ Another population of interest as potential innovators is the Specific Language Impairment population (Petra Schulz, p.c.), which makes up approximately 7% of the population and exhibits prolonged and divergent L1A in several areas (Tomblin *et al.* 1997).

1979; Shepherd 1982; Wells 1979; Bassano 1996; Papafragou 1998; Heizmann 2006; Cournane 2014, 2015a,b), and along the diachronic Modal Cycle (Jespersen 1924; Lightfoot 1979; Traugott 1989; Bybee *et al.* 1994; Gergel 2009, i.a.). The modal cycle has both a semantic component (lexical > functional (root > epistemic)) and a syntactic one (V > *v* > AUX/INFL).⁵ The semantic pathway has parallels in child language, while the syntactic pathway in children is in opposition to diachrony; the Modal Cycle thus serves as a useful illustration for a discussion on diachronic directionality and the CIA.

2.3.1 *Parallel alignments: Mapping existing forms to new meanings*

Parallel alignments between acquisition and change have generated many discussions since the early days of historical linguistics (see Baron 1977: ch.1, for an overview). In parallel alignments the same descriptive pathway is charted in child development as in historical development. For example, in both domains, spatial meanings of adpositions arise before temporal (*after*_{spatial} > *after*_{temporal}; Ziegeler 1997; Clark and Carpenter 1989), motion paths arise before temporal paths (*be-going-to*_{motion} > *be-going-to*_{future}; Fleischmann 1989; Schmidtke-Bode 2009), and root modal meanings precede epistemic meanings (Traugott 1989; Papafragou 1998). In this section we focus on modal semantic development.

Modals (e.g. *must*, *can*, *have to*, *gonna*) are quantifiers (i.e. *must* is universal with a necessity meaning, while *might* is existential with a possibility meaning) that express two broad types of meanings (Kratzer 1977, 1981, 2012, i.a.): *root* meanings like abilities (1a), deontic obligations (1b), or future intentions ((1c); Copley 2002), and *epistemic* meanings (evidence-based inferences, like (1d)). Polysemous⁶ (or functional) modals are those modal expressions that express both root and epistemic meanings, depending on context and/or contributing grammatical factors (Kratzer 1981; Perkins 1983; Sweetser 1990; Hacquard 2006, i.a.). For example, English *must* can express deontic (2a) or epistemic (2b) meanings. Root meanings arise when a polysemous modal scopes below Tense (over the verbal event), while epistemic meanings arise when a polysemous modal scopes above Tense (over a propositional event) (see Hacquard 2010 for analysis). Polysemous modals are more grammatically complex than monosemous modals that consistently express only one modal meaning ((3); see Hacquard 2013).

- (1)
 - a. Olivia **can** speak Spanish.
 - b. Erin **has** to be home before 11pm.
 - c. I'm **going** to go to the shops before working.
 - d. You **must** be going skating, since you're holding your hockey stick.

⁵ I do not address any claims about sound change.

⁶ This is shorthand; the same modal word (e.g. *must*) expresses both root and epistemic meanings in combination with different modal bases and ordering sources but the modal itself is not polysemous in the Kratzerian tradition, the lexeme is specified only for quantificational force (e.g. strong necessity or weak possibility meaning; Kratzer 1977, 1981).

- (2) The baby **must** be fed...
- a. (i) ... because he's hungry.
 (ii) The baby **must**_{Root} [be fed]_{VP} Root Deontic
- b. (i) ... because he's not hungry.
 (ii) **must**_{Epistemic} [the baby be fed]_{TP} Epistemic
- (3) Monosemous modal expressions
- a. I **know** Darwin is hungry.
 b. Darwin is **probably** hungry.
 c. It's **likely** (that) Darwin is hungry.

Diachronically, lexical modal verbs with root meanings like *want* or *know how* ('pre-modals', see Lightfoot 1979) grammaticalize into functional modals. Within functional modals, root meanings dominate early on and epistemic meanings increase over time (see Traugott 1989; Tagliamonte and D'Arcy 2007). For example Old High German *musan* (> *müssen*) (4a) was a main verb with the root meaning *[[to be able]]*. In (4), we see that *müssen* has changed from ability meaning in OHG, to deontic meaning in MHG, and finally occurs with epistemic meanings in Modern German. This pattern of semantic developments for lexical items is robustly attested (Bybee *et al.* 1994). Currently, the verb *müssen* is used to express both deontic and epistemic meanings, like English cognate *must*. The meanings expressed by a lexical item go from lexical > functional (root > epistemic). Focusing here on the development from root > epistemic, *müssen* first appears with root meanings and later with epistemic. Root > epistemic reanalysis reduces to the following: the modal first scopes over the verbal event at LF (5a), expressing root meanings, and later extends to the higher scope position above the proposition (5b), expressing epistemic meanings (see Hacquard 2013; Hacquard and Cournane 2015 for details).

- (4) The Semantic Modal Cycle (German; Fritz 1997, cited in Traugott 2006: 111):
 OHG *müssen*_{ability} → MHG *müssen*_{deontic} → ModG *müssen*_{epistemic}
- a. *Sie ni musan gan so fram zi themo heidinen man.*
 they not were-able go so far to the heathen man
 '[For religious reasons] They were not able to proceed further to [the palace of] the heathen man (Pilate).' (9th c., Otfrid IV.20.4)
- b. *Tie minnera habeton die muosan gan.*
 those no.money had they had.to walk
 'Those who had no money were obliged to walk.' (c.1000 Notker I.152.1)
- c. *Du must wohl müde sein.*
 you must MOD.PRT tired be
 'You must be tired.' (S. Gargova, p.c.)
- (5) LF scope interpretations
 The baby **must** be fed.
- a. The baby **must** [be fed]_{VP} Root Deontic
 b. **must** [the baby be fed]_{TP} Epistemic

In the individual L1 learner, polysemous modals like English *must* or French *devoir* (Bassano 1996) first appear in naturalistic child speech with root meanings (6a), then later with epistemic meanings ((6b); after the third birthday). This developmental pathway, like the diachronic one in (4), is robustly attested (for Turkish, Finnish, and Greek, Stephany 1979; for French, Bassano 1996; for English, Wells 1979; Papafragou 1998; for Mandarin, Guo 1994, i.a.).

- (6) First root *must* and first epistemic *must* used by Sarah (Brown 1973)
- a. he mus(t) talk [Deontic; Sarah 2;9, referring to her broken talking-doll]
 - b. must be gone [Epistemic; Sarah 3;0, referring to missing toy plates]

If the CIA is correct and children are responsible for innovations like root > epistemic, then children are expected to overextend the meaning coverage of root lexical items. In other words, meaning extensions must create the new mapping relationships for form ↔ meaning that arise in the historical record. To develop semantic productivity, we know that the child must extend meanings learned within constrained contexts to novel contexts. For example, when learning the word ‘kitty’ with one referent (the family cat Darwin) the child is tasked with extending that meaning (to all cats, including Darwin). How far does the meaning extend, or, which set of referents is correct? Extension is necessary for productivity, however, if the child also calls dogs *kitty*, she has overextended the meaning coverage of *kitty* to include too many possible referents. Overextension happens in normal acquisition of both lexical and functional items (e.g. Bowerman 1985; E. Clark 1993). For example, even if *must* only has root meanings in the input (as was once the case in English and for West Germanic cognates), the semantic grammar still allows functional modals to merge in two positions (see Hacquard 2006, 2010; Hacquard and Cournane 2015 for details). The child’s grammar can productively extend the meaning of *must* to cover epistemic usage.⁷ The compositional semantics predicts the grammatical availability of root > epistemic extension (nothing formally restricts a functional modal from scoping over either a VP or a TP; see Hacquard 2006 for details).⁸

As a normal part of development, children must learn and constrain the extent of meaning coverage of lexemes as part of solving the form ↔ meaning mapping problem; sometimes the grammar will allow productive patterns that are input-divergent but internally consistent. If and when these diffuse, they are expected to give rise to noticeable parallels between L1A patterns and change patterns (the former begets the latter).

⁷ With modals, as with *be-going-to* or adpositions, the earlier meanings typically remain when the newer meaning arises, giving rise to one-to-many form ↔ meaning mappings (Hopper and Traugott 1993), as is relatively common in the modal domain (van der Auwera and Ammann 2011).

⁸ Why might the child do this? The child is under pressure from expressivity forces: she needs to express herself with a limited lexicon and grammar and thus may be more likely to co-opt morphemes with meanings close to the desired meaning in her thoughts (and the grammar allows it).

2.3.2 *The CIA and parallel alignments: Response to critics*

The child meaning mapping extensions from root > epistemic, space > time, or motion > future are parallel to diachronic extensions for lexemes. In both cases, more concrete meanings extend to more abstract meanings in predictable, metaphorical, ways. For Diessel (2011, 2012), parallel alignments are parallel only because in both diachrony and acquisition the same psychological mechanisms (analogy, entrenchment, and categorization) are involved, not because the child extensions are the diachronic ones (as the CIA argues). For adult language users to be responsible for parallel pathways, adults must be shown to overextend meanings in the direction predicted by the diachronic trajectories. While adult innovation is plausible (also for many generative researchers, adults may be responsible for at least some types of change), there are three hurdles the adult approach faces that the child approach does not.

First, adults have more experience with language and with existing speaker variation within and across dialects. When adult speakers produce what appear as innovations, as in over-regularization of morphology (*went* > *goed*; see Bybee and Slobin 1982a), it is always possible that the speaker acquired that variant in development or has been exposed to other speakers who have those variants natively. Adult usage of variants is more likely participation in sociolinguistic borrowing, incrementation, or diffusion of extant variants. Further, once the existent variants from the speech community are eliminated as candidates for adult innovation, the remaining candidates for adult innovations may be localized events that stand out as processing errors or poetic/playful acts; this is different from the child who systematically produces internally consistent, productive, input-divergent forms. While the child approach must demonstrate that child innovations can become fixed in their grammar and/or spread to other speakers (see Section 2.2 above), the adult approach faces this problem in addition to the burden of showing that adults are true innovators of new variants.

Second, the adult innovator approach potentially violates Ockham's razor, as it needs to demonstrate that adults also make systematic meaning extensions that align with diachronic innovations. There is agreement that children make creative extensions as in (7), in the direction compatible with diachronic patterns; but only the CIA view assumes that these are ultimately the same extensions that become diachronic innovations. In other words, the CIA hypothesizes that parallels suggest not only the same kinds of extensions in L1A and change, but also maintains that the simplest solution for the source of semantic innovations is the input-divergent meaning extensions of the developing learner. There is only one mechanism underlying overextension; we do not need to argue that the child makes psychological extensions (which, as discussed below, are suspect when not treated with enough grammatical detail) and then later adults make similar extensions. It is less clear what would motivate adults to overextend meanings along conceptual routes, while this is obvious in child development (meaning extension to the set of possible referents in the world underlies the construction of productivity in the grammar).

- (7) Can I have any reading **behind** the dinner? (= after) (Bowerman 1985: 1292)

It is commonly accepted that the L1A root > epistemic pathway reflects conceptual development in Theory of Mind (Moore *et al.* 1990; Papafragou 1998; cf. de Villiers 2007; Cournane 2015*b*); children do not use epistemic modals until they can entertain epistemic thoughts. However, previous literature, on which the Theory of Mind approach is based, focused almost exclusively on canonical English-type polysemous modals. Recall that polysemous modals enter into more complex grammatical (compositional semantic) relationships than monosemous modals. Modal meanings (root vs. epistemic) are linked to LF scope for polysemous modals (e.g. *must, can, have to, gotta*), but not for monosemous modals (*maybe, probably*). Crucially, epistemic interpretations for polysemous modal verbs depend on the ability to embed a proposition (= sentential complement) under a modal (5*b*). Cournane (2015*a,b*) shows that a child's first epistemic use of a polysemous modal (e.g. *must*) correlates strongly with first markers of sentential embedding (first uses of embedding verbs like *see* or *think*, productive infinitival marking on the second verb, disjoint subjects in matrix and embedded clauses). These grammatical complexity-based reconsiderations of the root > epistemic dogma call into question the practice of considering uses of forms like *must* as directly signifying conceptual states, without considering the mediating or confounding effects of grammatical complexity development (see also de Villiers 2007).

Even the 2-year-old child has other means of expressing reasoned guesses (grammatically simpler monosemous modals *maybe, peut-être, think*; Bassano 1996; O'Neill and Atance 2000; Cournane 2015*a,b*; see de Villiers 2007 for discussion on the lack of straightforward alignments between modal developments and Theory of Mind milestones). Thus, it is questionable to argue that *conceptual* extensions underlie child development for root > epistemic, and is likewise suspect for diachronic extensions. In both the child and diachrony, extensions are linked to local components of the grammar. The child extends the meaning coverage of polysemous modals to higher scope positions (above propositional content, minimally TP). This is a grammatical development where a form goes from a one-to-one meaning relation to a one-to-many meaning relation. As with the child, 'conceptual routes' in diachrony are localized to lexemes entering new grammatical relationships, they are not conceptual changes. In both the child and diachrony, the trajectory from root > epistemic describes grammatical development, not conceptual development; it captures the extension of the meaning coverage of polysemous forms from low scope to high scope interpretations.

I conclude that for semantic diachronic pathways where we see parallel alignment, child evidence is compatible with the CIA if we find that the child learner extends the meaning mappings of a form in order to advance the cycle. Children must necessarily extend meanings; they are 'extenders' and we know they sometimes overextend in ways consistent with their grammar. Further research is needed to explore whether functional overextensions in more domains of L1A provide the diachronically expected innovations. The assessment metric for whether children make the right kinds of overextensions is previous language changes, but we must never fall for the trap of expecting recapitulations with the same lexical items. Only renewing items (see van Gelderen 2009*a*) that will forge truly new changes are under consideration when we assess the CIA in child data.

2.3.3 *Oppositional developmental alignments: New forms for existing meanings*

Many historical changes also occur in a unidirectional manner, but fail to show a parallel in L1 acquisition. Diessel (2011) has argued that such lack of parallels in child language constitutes evidence against the CIA. Two common such changes from the historical record are (1) reanalysis of biclausal structures into monoclausal structures (e.g. $IP_{[MODAL\ VERB]} + IP_{[MAIN\ VERB]} > IP_{[MODAL\ AUX + MAIN\ VERB]}$; see Roberts 1985; or $CP_{[+Q,CLEFT]} + CP_{[+Q]} > \text{complex } CP_{[+Q]}$; see Tailleir 2012); and (2) reanalysis of independent morphemes into bound morphemes (e.g. pronoun > clitic > agreement, see Fuß 2005). These morphosyntactic changes appear to lack L1 parallels, unlike the semantic extension changes discussed above. Generalizing, we see structural simplification processes in diachrony (see Longobardi 2001; Roberts and Roussou 2003), but the child develops syntactic complexity ((8b); e.g. Brown 1973; Diessel 2004; Pérez-Leroux *et al.* 2012). In these domains we see oppositional, not parallel, developmental pathways (8). I will focus on the development from biclausal structures to monoclausal structures, as attested from the historical record for modal verbs in the history of English (Lightfoot 1979; Roberts 1985; Roberts and Roussou 2003, i.a.).

- (8) Oppositional developmental pathways:
- a. Diachrony: Complex Structures > Simpler Structures
 - b. L1A: Simpler Structures > Complex Structures

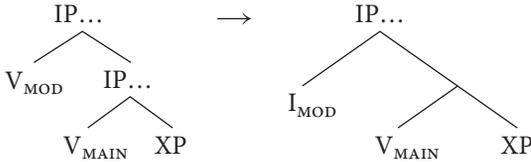
The canonical set of English auxiliary modals (e.g. *must, may, should, can, will, might*) is currently realized as inflectional markers (in INFL or T; Pollock 1989, i.a.). They underwent wholesale reanalysis from premodal verbs to inflection markers during the late OE and ME periods (see Lightfoot 1979; Roberts 1985; Warner 1993; Denison 1993, i.a.). For example, in the historical record we see that in OE and into ME these modals appeared with subject agreement and infinitival complements (9a). Currently, these modals are invariant in form⁹ and take bare verbal complements (9b). In (10a) we see an example from OE with the premodal *sceolde* (currently *shall*) taking a direct object (*hundteontig mittan hwætēs*). Later in the historical record, from ME, these modals no longer appear with direct objects, as today (10b).

- (9) Loss of inflection
- a. *some hit mæi illimp-en*
soon it may.3SG happen.INF
ME a1225 (?a1200) Lay. *Brut* 2250 (Denison 1993: 299)
 - b. It may happen soon.
- (10) Loss of object argument
- a. *he cwæð þæt he sceolde him hundteontig mittan hwætēs*
he said that he shall him hundred bushels of-wheat
'He said that he owed him a hundred bushels of wheat.'
OE *Æhom* 17:26 (Fischer and van der Wurff 2006: 147)
 - b. *I shall [DP]

⁹ There is arguably still past marking on some of these modals, for example *could, would, and should* may retain at least some reflexes of previously productive tense marking (see Cowper 2003).

These changes (among others, see Lightfoot 1979; Roberts 1985) reveal a syntactic reanalysis for the canonical modals from verbs that inflect for tense and phi-features and take arguments, to exponents of INFL, which are themselves an expression of tense features and have lost their argument structure (see Pollock 1989). The reanalysis from $V > \text{INFL}$ ¹⁰ was necessarily concomitant with a reanalysis of the entire clausal structure. Prior to reanalysis, constructions with modal verbs were biclausal, with the modal verb embedding an infinitival verb. After reanalysis, the structure is monoclausal because the (newly) INFL modal becomes the inflection marking for the main verb (11).

(11) Structural reanalysis of two clauses into one (higher $V > \text{INFL}$).



L1 development, on the other hand, progresses from simple clauses $>$ complex clauses (e.g., Brown 1973; Bloom 1991). The development process is highly contingent: with each new development, new expansions that were previously closed become possible (e.g. Gleitman *et al.* 2005). For example, Sarah (Brown 1973), like other children, begins using the premodal verb *want* in simple constructions followed by a bare noun (12a) or bare verb (12b,c), later we see *want* with multi-verb complements (12d), and later still with a possible subject for the second verbs (12e), and finally with a second verb with its own subject and inflectional marking (12f). This is the opposite trajectory from what we saw above for reanalysis from biclausal to monoclausal structures: structural reanalysis in diachrony appears as simplification but the child gradually develops structural complexity.

- (12) a. want Bobo [2;3,07]
 b. I wanna¹¹ ride my horse [2;3,07]
 c. I wan ride a horsie! [2;4,12]
 d. wan go read it [2;7,28]
 e. I want that write on [2;10,05]
 f. I want Daddy to help me [3;3,13]

What does the CIA predict for child data, if the child is responsible for innovations from complex biclausal structures to simpler monoclausal clauses? In oppositional pathways, the child needs to stay at a structurally simpler stage of analysis and never

¹⁰ I am simplifying, it is likely that these modals actually followed the more stepwise trajectory: $V > v > \text{INFL}$ (see Tollan 2013 for discussion).

¹¹ In this usage the transcriber wrote *wanna* but when assessed against all other uses, and considering Sarah's low MLU of 2.106 (relative to other typically developing English-learning children) it is likely that *wanna* at this point is an unanalysed whole. This debate over the correct analysis of early forms may equate to input opacity for the child too, not just for the linguist (thanks to an anonymous reviewer for pointing out this possibility).

posit a more complex structural analysis for the string in question (see Roberts and Roussou 2003 and van Gelderen 2004, 2011a for differing economy-based approaches). For example, children can use INFL (e.g. modal auxiliaries), as in (13a) earlier than they can embed sentences, as in (13b) (around age 2 vs. around age 3). Now, imagine that the 2-year-old child, unable to subordinate, analyses input strings like (13b) as monoclausal (*want dance*; cf. (12c) from Sarah). If the input is opaque with regards to key pieces of evidence that *want* is a subordinating verb, like the infinitival marker, then the child's monoclausal analysis may persist as her final analysis. Note this is precisely what happened in the history of English (oversimplifying for expository purposes): the infinitive marker on the second verb in biclausal modal constructions was phonologically eroded, leading to reanalysis of biclausal modal structures into monoclausal structures with INFL modals (14).

- (13) a. [(I) can_{INFL} dance_V] (emerges at approximately age 2; e.g. Wells 1979)
 b. [(I) want_V [to dance_V]] (emerges at approximately age 3; e.g. Diessel 2004)

- (14) [I want_V [to dance_V]] > [I want_{INFL} dance_V]

This type of reanalysis is fed by phonological changes that level inflectional marking paradigms, as in the history of English (see Lightfoot 1979; I. Roberts 1985 for effects of loss of infinitival and subjunctive marking on verbal reanalysis). Critically, the loss of infinitival marking makes the simplex analysis sufficient and the conservative learner will not posit more structural features than necessary (see van Gelderen 2011a). The meaning of the string remains consistent, but the syntactic composition is reduced (fewer nodes) (see von Stechow 1995 for a similar approach using semantic types; see Fuß 2005 for comparable meaning-preserving string reanalyses for the pronoun > agreement cline).

In sum, throughout development the child has interim analyses (forms) for syntactic strings consistent with the stage of her grammar; an interim analysis becomes her ultimate analysis if it: (a) captures the compositional meaning of the string, (b) is consistent with the rest of the grammar, and (c) is not sufficiently cued by the input to be complexified. In morphosyntactic reanalyses the child posits a new form (= structural analysis) for an existing meaning (the compositional meaning of a modalized verb). The new form in diachrony aligns with a simpler, earlier, stage in child development.

2.3.4 The CIA and oppositional alignments: Response to critics

Critics of the CIA argue that morphosyntactic oppositional pathways provide evidence against the CIA because these changes do not show child language parallels. For example, Diessel (2012: 1609) argues:

While the semantic developments of grammatical markers are often parallel in language acquisition and diachronic change, the morphosyntactic developments are different. There is no evidence that grammatical markers originate from lexical expressions in language acquisition as they do in diachronic change. . . . The acquisition of the English present perfect, for instance, does not originate from an attributive construction as in diachronic change; and the *be-going-to* future is acquired in the context of a simple clause, whereas the historical development

originates from a bi-clausal purposive construction. . . . If children were the main instigators of language change, one would expect a closer match between the two developments. In fact, since the categorical and constructional changes of grammaticalization have no parallels in early child language, we can be fairly certain that grammaticalization processes do not originate from changes in language acquisition. (Diessel 2012: 1609)

This is an argument against a straw man; under the CIA, the child is not expected to recapitulate completed changes (*be-going-to*), neither for parallel alignments nor for oppositional ones. The child is expected to forge new innovations (see van Gelderen 2004, 2009a, 2011a), a point which is even more critical for oppositional pathways. When a child learner progresses from one stage of development to another, her only accountability is to the present system of her grammar, to the input (remembering she has access to different information from the input as she develops, see Gleitman *et al.* 2005), and to her language-learning capacity. She is tasked with creating productivity (e.g. categorizing objects, events, and propositions in the world around her and mapping meanings to forms).

The child inevitably makes analyses that are input-divergent throughout this process (the example in (7) constitutes just such a not-yet-realized potential innovation for *behind*).¹² In order to confidently claim that child input-divergence can or cannot be responsible for changes we need to look at systematic biases in, for example, the entire modal system of learners, and look to renewing items and how they are treated by learners. For modals, renewing items are lexical verbs that denote states of obligation, potential, desire, etc., such as *to owe*, *to know how*, or *to want* (see Lightfoot 1979; Tollan 2013). If we are to study child development for innovations, we need to look at current renewing items, like *to know how*, or *to want*. These verbs currently express the meanings that *cunnan*, *willan* exhibited in OE. As these items provide renewals to the linguistic cycle, the CIA predicts that children should show directional biases with these items in development (and not with *be-going-to*). For example, the CIA predicts that the verb *want* may become a future marker, so we should look at how children treat *want*.

In sum, the existence of oppositional pathways does not refute the CIA because the CIA does not predict uniform relationships between all child phenomena and all change phenomena. The child must necessarily determine the compositional relationships between elements in a sentence; if an early (= less complex) analysis is successful at capturing the semantic composition of a sentence and is string-compatible (see von Stechow 1995; Fuß 2005), there is no internal motivation to posit a more complex analysis. Rather, different aspects of child development are predicted to underlie semantic overextensions and morphosyntactic reanalyses. This serves to illustrate that we must be domain-specific about our predictions, maintaining sensitivity to the expected renewing innovations and the actual child patterns observed in that domain of language development. As many have cautioned before me, overadherence to biology principles such as ‘ontogeny recapitulates phylogeny’ obscures the more

¹² Anecdotally, use of *behind* for *after* appears regularly in the Baltimore, Maryland-based TV show *The Wire*, e.g. *They lost their daddy behind what happened* (S1E02). I have not yet pursued whether this is a feature of native speakers from the region or AAVE more generally.

TABLE 2.1. Summary of proposed alignments between L1A and change

	Type of input divergence	Type of diachronic innovation	Means of diffusion
Parallel alignment	Overextension <i>Old form > New meaning</i>	<i>Semantic change</i>	Peer-to-peer reinforcement Persistent input-divergence
Oppositional alignment	Structural reduction <i>Old meaning > New form</i>	<i>Morphosyntactic change</i>	>> sociolinguistic diffusion

fine-grained relationships between individual development and language changes. Table 2.1 summarizes my arguments.

2.4 Conclusion

Child input-divergent analyses remain a promising source of both directional semantic changes and directional morphosyntactic changes. The data shows compatibility once we allow that different developmental mechanisms for productivity and analysis underlie different types of innovations. For parallel alignments, the child posits new meanings for existing forms, extending the meaning coverage of a lexical item. For oppositional alignments, the child posits a new form for an existing meaning, thus changing the structural relationship between compositional elements (see Hale 1998 for why reanalysis of functional elements is particularly robust).

While this approach is still relatively broad, it moves in the direction necessary for testing the validity of the CIA; generative diachronic theory can be brought to the level of domain-specific, fine-grained predictions for child language if we consider what child mechanisms align with advancements (not recapitulations) along a grammaticalization cycle. This broad approach provides opportunities for meaningful progress to generative change theory. By taking complexity in first language acquisition seriously, we can test specific hypotheses for generative change theory with data from contemporary child corpora and experiments. Furthermore, several avenues of research provide possible avenues for the reinforcement or diffusion of input-divergent analyses (peer-to-peer first language acquisition, bilingual language acquisition, and established sociolinguistic diffusion mechanisms). More targeted research is needed to assess the empirical basis of diffusion of child input-divergent analyses for these populations, but the mechanisms and extant language contact situations are present.

This chapter has responded to criticism of the generative approach to language change by recasting the CIA as part and parcel with the fundamental mapping problem in language acquisition, stressing that both child developments and diachronic

innovations reduce to redistributions of form \leftrightarrow meaning grammatical relationships. I suggest that forces that create productivity in the child's developing grammar lead to input-divergent properties which later diffuse through normal peer-to-peer acquisition or sociolinguistic mechanisms of change (diffusion). After 100 years, the child innovator is still the most compelling source for innovative grammatical analyses, though we must assess more areas of child development at the right level of targeted formal analysis.