

Ailís Cournane*

A developmental view on incrementation in language change

<https://doi.org/10.1515/tl-2019-0010>

Abstract: Acquisition is an intuitive place to look for explanation in language change. Each child must learn their individual grammar(s) via the indirect process of analyzing the output of others' grammars, and the process necessarily involves social transmission over several years. On the basis of child language learning behaviors, I ask whether it is reasonable to expect the incrementation (advancement) of new variants to be kicked off by and sustained by the acquisition process. I discuss literature on how children respond to input variation, and a series of new studies experimentally testing incrementation, and argue that at least for some phenomena, young children overgeneralize innovative variants beyond their input. I sketch a model of incrementation based on initial overgeneralization, and offer further thoughts on next steps. Much collaborative work remains to precisely link analogous dynamic phenomena in learning and change.

Keywords: language acquisition and change, innovation, actuation, incrementation, overgeneralization

1 Introduction

The language development process is rich ground for theorization about language change. Each child must learn their individual grammar(s) via the indirect process of analyzing the output of others' grammars (the input or Primary Linguistic Data; e.g. Hale 1998; Lightfoot 1979), and the process necessarily involves social transmission over several years. Children have long been entertained as primary agents of change, for both the innovation (e.g. Andersen 1973; Halle 1964; Meillet 1912; Roberts and Roussou 1999; van Gelderen 2004, i.a.) and incrementation (Labov 2001, 2007; Holmes-Elliott 2016; cf. Tagliamonte and D'Arcy 2009) of new variants in all domains of both I-languages (mental grammars) and E-languages (community languages, e.g. "Toronto English", made up of multiple individuals with social ties)

*Corresponding author: Ailís Cournane, Department of Linguistics, New York University, 10 Washington Place, New York 10003, NY, USA, E-mail: cournane@nyu.edu

(Chomsky 1986). However, while change literatures emphasize acquisition, and sociolinguists have conducted studies on variation in child language (e.g. Guy and Boyd 1990; Labov 1989; Smith et al. 2007, 2013), independently known acquisition phenomena (e.g. overregularization; Marcus et al. 1992; Hudson Kam and Newport 2005, 2009; Schuler et al. 2016), are rarely brought to bear on change problems (see Cournane 2014; Heycock and Wallenberg 2013; Yang 2000). This paper is intended to fuel this conversation, focusing on incrementation, and arguing child learning is always about extending beyond finite input and offers a testing ground for whether such extensions are in line with change patterns. Learning puts constant change pressure on E-languages (Walkden 2012), as all children face the challenge of using dynamic input to abduct productive grammars.

On the basis of child learning behaviors, is it reasonable to expect the incrementation of new variants to be kicked off by and sustained by the acquisition process? I motivate the view that all children are innovators and incrementors as an inevitable effect of learning productive grammars on the basis of finite, variable input over maturational time. On the basis of two experimental studies on Toronto English: modal structure-to-meaning mapping (Cournane 2015; Cournane and Pérez-Leroux under revision) and vowel shifts (Hall and Maddeaux 2018), I sketch an inverted U-shape model of incrementation (cf. Labov 2001, 2007; Tagliamonte and D’Arcy 2009), and argue that equating child generalization biases with incrementation gives a testing ground for exploring why incrementation begins in the first place.

This paper is organized as follows. Using modal verbs to illustrate (e.g. *must*, *have to*), in Section 2 I lay out the phenomena involved in change, broadly defined, from both “radical” I-language approaches which focus on grammatical reanalysis between individuals, and “gradual” E-language approaches, primarily variationist sociolinguistic change theory¹, which focus on change-in-progress (see Fischer 2003). In Section 3, I cover acquisition work showing children: (a) innovate input-divergent properties of grammar in the normal course of learning (Cournane 2015; Hudson Kam and Newport 2009; Pinker 1984; Schuler et al. 2016), (b) are highly sensitive to variation in their input (e.g. Labov 1989; Roberts 1994; Smith et al. 2013), and (c) overshoot the input in ways consistent with incrementation (Cournane 2015; Cournane and

¹ Cf. Functionalist views (e.g. Traugott 1989; Hopper and Traugott 2003), which do not theorize a role for the child, so are largely set aside here. See Cournane (2017) for arguments from acquisition pathways and peer-to-peer reinforcement in defense of the child innovator, critiquing functionalist approaches, which argue against children as even possible agents of morpho-syntactic changes (see Diessel 2011).

Pérez-Leroux under revision; Hall and Maddeaux 2018). Weaving together these three findings, I present an inverted U-shaped model of incrementation, defending the view that children's strategies for generalizing from their input are a good place to look for why incrementation begins and is directional.

2 Change phenomena, their terms, and where acquisition research can intervene

Productive discussion of change theory will not be possible without first clarifying terminological differences across approaches. I will use modals to illustrate, as these are well-studied in diachrony, particularly for English verbs (Lightfoot 1979; Roberts 1985; Traugott 1989), and provide one of the case studies on incrementation (Cournane and Pérez-Leroux under revision). Modal meanings divide into two broad types: root meanings (e.g. ability, bouletic, deontic) and epistemic meanings (e.g. knowledge based inferences). These meanings can be expressed via lexical categories (1) or functional categories (2) (Kratzer 1981). Functional modal verbs develop historically from lexical modal verbs (Hacquard 2013; Traugott 2011; among others), and are typically variable-meaning, with one form capable of expressing both root (2a) and epistemic meanings (2b) (cf. (1a–c) vs. (1d–f)). Interpretation is constrained by syntactic construction for variable-meaning modals (Brennan 1993; Cinque 1999; Hacquard 2006).

- | | | |
|-----|---|-----------------------------|
| (1) | a. You are able to jump very high. | Root (Ability), *Epistemic |
| | b. You want to play with the cat. | Root (Bouletic), *Epistemic |
| | c. It is permissible/necessary to swim nude. | Root (Deontic), *Epistemic |
| | d. You are likely to feed the cat. | *Root, Epistemic |
| | e. Maybe you swam nude. | *Root, Epistemic |
| | f. I think you didn't. | *Root, Epistemic |
| (2) | <i>The cat must/has to be fed...</i> | |
| | a. <i>...because he's hungry.</i> | Root (Deontic) |
| | b. <i>...because he's not hungry.</i> | Epistemic |

2.1 Variation & change

Modals illustrate two kinds of variation we see in natural language: more than one way of expressing the same meaning – variable forms (1), and more than one meaning for the same form – variable meanings (2). Note the interrelationship

between variable-form and variable-meaning, as the set of variants for a variable-meaning modal like *must* differs by its interpretation (3).

- (3) Variants in meaning-expression competition, by interpretation of *must*:
- a. Root necessity: *have to, need to, got to, have got to, ought to, should...*
 - b. Epistemic necessity: *have to, probably, know, think, guess...*

While the variants listed in (3a) and (3b) are semantically the ‘same’, they vary by syntactic category and also have fine-grained meaning differences in some cases. Variation is inseparable from change theories, as innovation introduces new variants, normally while maintaining conservative ones (Hopper and Traugott 2003), and changes accumulated over time differentiate languages from each other leading to cross-linguistic variation. Here, I am concerned primarily with how children learn intraspeaker variation not readily attributable to bilingual or bidialectal analyses². For the child, learning modal words presents with several challenges: they must work out which forms have which meanings in which grammatical and situational contexts, which of these overlap in meaning (entirely or partially), and further, which forms express more than one distinct modal meaning (van Dooren et al. 2017). Child responses to these learning challenges provides inspiration for linking theories to change (Section 3.3).

The term *change* is used by I-language historical linguists (e.g. Hale 1998; Lightfoot 1979, 1999; Roberts and Roussou 1999; van Gelderen 2004) to refer to formal differences between grammars in a descent relationship (or H-relation, Crisma and Longobardi 2009). This type of change takes place between individuals’ grammars, and is compatible with modifiers like “discrete” and “radical”, but not those like “over three centuries”. We’ll call this *reanalysis* henceforth, as it clearly refers to a comparison between pre- and post-grammatical systems, where analysis of the relevant phenomena grammatically differs. Note that if the prior representation remains possible (i.e. if variation or competition arises) this does not change the fact that the grammars in a descent relationship are grammatically substantively different.

The term *change* is used by variationist sociolinguists (Weinreich et al. 1968) to capture the dynamic properties of having variation in a language

² Universal bilingualism ideas (Roeper 1999), related to competing grammars approaches to change (Kroch 1989), are compelling even for ostensibly intraspeaker variation, though implementation is unclear.

community³ (i.e. an E-Language), comprised of individuals with social ties. Variation is structured by grammatical (e.g. phi features, verb class)⁴ and social (e.g. age, gender, social networks) factors. *Change*, in this framework, starts with variation in the phenomenon of interest, includes the period of change-in-progress, and is complete when variation ceases because one variant won out over the other in the community of study. Unless new variants have been introduced (see cyclic change approaches; van Gelderen 2009), the language to be acquired is rendered categorical for this ‘variable’. For example, many centuries ago English had the lexical verb *sceal* ‘to owe money’, which was reanalyzed as an auxiliary verb (Lightfoot 1979; Roberts 1985) and now has no verbal uses (change-in-progress complete) and even its auxiliary uses (future modal meaning) are restricted to unproductive corners for many speakers (cf. Romance languages where the *to owe money* meaning survives despite grammaticalization to modal meanings, for e.g. Fr. *devoir*). This use of *change* is compatible with modifiers like “over three centuries” and “in progress”, but not those like “discrete” or “radical”. I’ll use *change* to refer to the entire process including components not in focus in change-in-progress studies, but central to I-language theories of reanalysis, and vice versa.

2.2 Innovation & actuation

Innovation occurs when an individual posits a grammatical relation (e.g. a form-meaning mapping, a form-pronunciation mapping, a rule and its application conditions) that is divergent from the grammatical relations of the speaker(s) from whom the individual learned (Hale 1998). In previous work, I have called this input-divergence, and postulated that input-divergence is common along the learning path, varies by learning domain and input cue availability, and is the simplest source for diachronic innovations (Courname 2017). This is in keeping with the view that “change [=reanalysis] is the result of the acquirer being exposed to a PLD [=input] that differs in some way from that which the acquirer of the source grammar was exposed to.” (Hale 1998: 9).

3 Not all variation lends itself to change-in-progress analyses, as much variation appears stable in its instantiation from grammar to grammar, and over time in the community (though see Wallenberg 2013).

4 Grammatical constraints in variation are not always considered true variation because e.g. distributional information could give away when one form is used over another. For example *have to* and *must* vary in their syntax, even if their semantics overlap to a large extent. That said, speakers still have some freedom, from the grammatical perspective, as to which forms they’ll use in many contexts, where social evaluation may intervene.

Lightfoot stresses this point in his work, calling it a “triggering experience” difference (Lightfoot 1979, 1999). Divergence is not random, as innovative relations are posited on the basis of evidence available in the input, but generalized, extended beyond, or mapped differently from the input for the purpose of productivity. Productivity is a hallmark of language learning, showing inference beyond the simple tracking of input patterns. Mounting evidence suggests (and we readily intuit) that children abduct productive grammars more effectively than adults (e.g. Hudson Kam and Newport 2005; Schuler et al. 2016).

Note that on this definition innovations need not spread to other speakers or survive in the grammar of the individual, and all children are innovators along their developmental path, in line with the I-Language change view from Hale (1998: 6): “[A] general theory of change must provide an account for any change [=innovation], regardless of whether or not it diffuses.” Defining innovation via an input-divergence metric, rather than a metric based on survival or spread in diachronic records, means (at least on a strong view, like Hale (1998) or here) that there is no difference between diachronic innovations and non-target properties of children’s grammars. What differs is whether we follow innovations in diachrony by looking at descent relationships between individuals (where we only see surviving innovations), or in acquisition by looking at stages on the acquisition path within an individual. The focus is reanalysis between systems, accurately describing both states.

Actuation is one of five problems of language change (E-language) discussed in Weinreich et al. (1968). They considered it the heart of the problem, as the most mysterious and central part of language change (4):

- (4) “What factors can account for the actuation of changes? Why do changes in a structural feature take place in a particular language at a particular time, but not in other languages with the same feature, or in the same language at other times?” (Weinreich et al. 1968: 102)

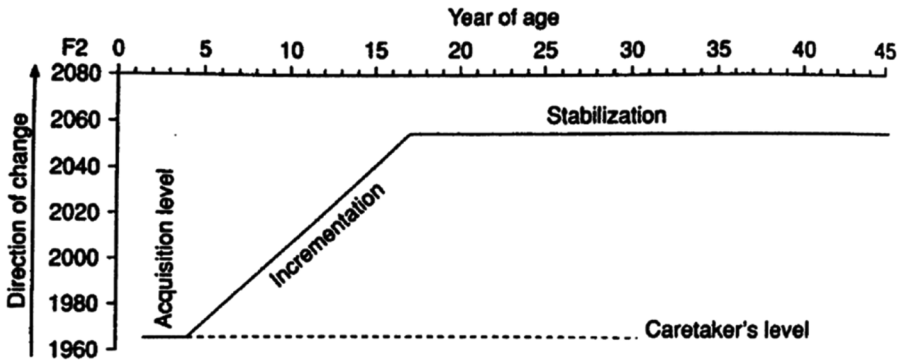
Building on the view of innovation just laid out, actuation concerns when an innovation catches on, and becomes relevant beyond the acquisition phase as more than a misstep on the path to the target (cf., Walkden 2017, who treats actuation as a spatiotemporally located instance of innovation). This involves spreading from one learner to other learners, or perhaps not from a single innovation, but by having enough innovations of the same type occur in multiple learners in the same community (as their input is similar in type) and then stay in their language through mutual reinforcement (Cournane 2017; i.e. being part of each other’s input, as with siblings).

In brief, innovation is the introduction of an input-divergent variant into the language of an individual, to be contrasted from actuation which is the introduction of a new variant into the speech community. Actuation entails innovation, but not vice versa, and continues to beg the questions laid out in (4). Actuation should draw on the pool of innovative variants, so understanding innovation is part of making the actuation problem surmountable.

2.3 Incrementation

Once an innovation has actuated by surviving learning, change-in-progress can ensue. The observed pattern in school-aged individuals in real time (Holmes-Elliott 2016), and in population usage data in apparent time (Labov 2001) is *incrementation*. Labov (2007: 346) uses this term to cover a few formally distinct processes which crucially increase the presence of one variant over the other: “Incrementation may involve increase in frequency, extent, or specificity of a sociolinguistic variable” – and he explicitly references children as agents in this process – “successive cohorts and generations of children advance a change beyond the level of their caretakers and role models”. The model in (5) is of an individual’s F2 measurement for a vowel change-in-progress over their lifetime, with a period of incrementation – advancing the change-in-progress (direction of change on y-axis) – in later childhood and adolescence. In (5), the individual first matches her caregiver’s (a member of an older generation) F2 level, then in late childhood and adolescence, advances the change (further fronting F2), and then stabilizes in adulthood.

(5) A model of incrementation (Labov 2001: 448)



Note: A linear model of incrementation for a single female speaker from 1 to 45 years of age.

Due to incrementation, when individuals in a community are viewed in the aggregate for a typical change-in-progress, younger speakers use the innovative variant more than older. Incrementation research has focused primarily on adolescents (e.g. Cedergren 1988; Tagliamonte and D'Arcy 2009), and has reported a peak innovative usage rate for individuals by age 17 (e.g. Cedergren 1988; Tagliamonte 2011). Tracking verbal forms for expressing root modal necessity (3a; *must*, *need to*, *have (got) to*), for example, there has been incrementation of *have to* since its introduction into the modal system, with particular concomitant usage loss of the older root necessity meaning of *must* (see Tagliamonte and D'Arcy 2007).

Sociolinguistic change is sometimes construed as being about diffusion of innovations through the speech community (through individuals in space and time; e.g. Hale 1998; cf. Cedergren 1988; Labov 2007), and on this view, incrementation is seen as orthogonal to diffusion as it only concerns usage effects. This is basically true on a historical view: an innovation, once actuated, then appears in the grammars of more and more individuals in a community (like tea, once actuated into boiling water, gradually spreads to fill the cup). However, even if grammars are assumed to not change in substantive ways beyond a certain age, incrementation would still affect input patterns to new generations of children, for example probabilities of hearing particular kinds of data in later generations (cf. Lightfoot 1999). And, in the approach taken here, incrementation is kicked off by and driven by repeated and reinforced input-divergence, so it arises, at least initially, from the same source. In this paper, we will focus on acquisition work up to early school years, often assumed to match caregiver variation properties (in line with the model in (5)). I'll argue the view that children first match input variation is inconsistent with much of child language evidence. I'll show that much of the time children start by overshooting rather than matching the innovative pattern (see also Holmes-Elliott 2018), and since children eventually end up matching the input closely (by learning its basis, rather than surface-matching) and change-in-progress is gradual, incrementation involves an interplay between innovations and retraction ("Two steps forward, one step back", Hall and Maddeaux 2018; see also Cournane 2015; Cournane and Pérez-Leroux under revision).

3 What children do with variable input

What do child do with variation in their input? The answer depends on the nature of the variation, as variation can be inconsistent, or consistent, and when consistent may be conditioned in various ways or simply be probabilistic in distribution (see e.g. Schuler et al. 2016). Learning presents with many

challenges. In how children respond to these challenges, what does child input-divergence look like? Most acquisition research looks at what stages children go through on the way to adult like language, exploring interdependent effects of input, prior language learning, and cognitive capacities, with the implicit assumption that they do eventually succeed at this task. Considering change, what are children doing in the face of variation? Are their behaviors consistent with how languages change, were any of the properties of their interim grammars to not get overwritten by further learning?

3.1 Rule-learning through inconsistent vs. consistent variation

How does variable input affect learning? Consider again the modals in (1) and (2): variable-form and variable-meaning input complexifies word learning (an already difficult task, see Gleitman 1990; Clark 1993), as there are many-to-one and one-to-many relations present in the input, in contrast to one-to-one mappings⁵. I discuss a few child input-divergent patterns that illustrate children's learning generalization behaviors: (a) input signal-detection in noise variation, and (b) overgeneralizations with patterned variation (overregularization of rules and overextension of meaning coverage).

We know infants are remarkable statistical learners, tracking patterns of co-occurrence in the input from early infancy and demonstrating learning by around 7 months (e.g. Saffran et al. 1996), and that they not only track surface patterns but infer algebraic rules from patterns from the first year of life (Marcus et al. 1999). When the input is inconsistent, for example, in a Pidgin or young creole, or when a caregiver is an L2 speaker, children regularize this input despite the inconsistent input, not matching the input but rather “going beyond” and creating systematicity of grammar (Hudson Kam and Newport 2005; Singleton and Newport 2004). Adults, on the other hand, probability match from this kind of inconsistently variable input. Consider also cases like Nicaraguan Sign Language where over a few generations, when a community was formed, younger speakers played a crucial role in grammaticalizing initial communicative gestures into a full-fledged sign language (Senghas and Coppola 2001).

Children readily generalize patterns from the input, and appear to over-generalize rules when they have learned them (Marcus et al. 1992). This occurs across languages and is common, contra over-irregularization (Xu and Pinker

⁵ Why posit a new meaning for a form you already have a meaning for (van Dooren et al. 2017)? Why map a second form to a meaning you already mapped a form to?

1995). Furthermore, when children have learned a rule they also can apply it in novel cases (i.e. *wug* > *wugs*; Berko 1958; see also Schuler et al. 2016). As with noise-type variation, in experimental tasks adults have been shown to not regularize from input patterns with rules with exceptions, but to probability match (Schuler et al. 2016). Child overregularization creates patterns of variation not present in their input, like the example in (6; Robinson 2019). Here Abe (MacWhinney and Snow 1985) uses the comparative *-er* morpheme on both *good* and *bett-* (suppletive base) in the same discourse. Children often produce both input-matching and input-divergent forms in the same timeframe (see Yang 2000; cf. Kroch 1989, i.a.).

- (6) Excerpt of a transcript from the Abe corpus (4;09.00); from Robinson (2019)

CHI: Mom (.) this dressing is **gooder** than Thousand_Islands.

MOT: what's its name?

CHI: French.

MOT: uhuh.

CHI: I really like it. Mom (.) why is French **better** than Thousand_Islands?

Turning to another kind of grammatical relation, to learn the meaning of a form the child needs to work out what the appropriate extension of that form is. Children are argued to use various heuristics to work out word-to-world mappings (Clark 1993), for example mutual exclusivity and the principle of contrast. To take a straightforward example, to learn the meaning of *cat* children need to work out what creatures (and representations thereof) can be labeled with *cat*. Children are known to make some errors along the way to finding just the right set, sometimes apparently under extending (only the family cat is *cat*), or more convincingly because it involves positive evidence, overextending *cat* to similar animals like foxes or dogs. This is a mapping and extension issue and not simply related to perceptual inabilities (see Booth and Waxman 2002). What forms map to what meanings and where are the category boundaries drawn? Which concepts have unique labels and which don't (the perceptual-conceptual world is much richer than language, see, e.g. Gleitman 1990)? The same under- and overextension occurs with functional morphemes; Pérez-Leroux et al. (2004) showed experimentally that English-learning children go through a phase where they treat definite plurals (e.g. *the lions*) like Spanish children do, allowing generic readings; this can be viewed as overextension of meaning coverage for English.

In summary, children must work out generalized grammatical relations on the basis of variable input, with indeterminate situational contexts and, in some cases, with exceptions present (see Schuler et al. 2016). They are quite good at

this, and use abductive strategies adults do not: finding the signal in noisy data, and revealing their dynamic generalization and extension abilities through over-regularization and overextension patterns. Learning language productivity involves generalizing, and children are particularly capable.

3.2 Learning sociolinguistic variation

Sociolinguistic research has focused primarily on adolescents and adults when studying variation. However, from early on, some research on sociolinguistic variation has studied children and preadolescents (e.g. Guy and Boyd 1990; Kerswill 1996; Labov 1970, 1989; Roberts 1994). Earliest studies looked at sound variables (e.g. -t/-d deletion, Labov 1989), culminating in the general view that “[c]hildren initially replicate their caregivers’ system and then advance the variable elements” (Labov 2007: 346) and that “the acquisition of social constraints on variation has its beginnings in early childhood, but the bulk of this learning appears to take place after the age of four” (Roberts 1994: 177) (see also, Kerswill and Williams 2000; Labov 2001). The shift from modeling on the caregiver’s system to incrementing, is called “vernacular reorganization” (Labov 2001), and is sometimes thought of as a shift to another available norm (e.g. older children, teenagers): “children must at some point focus on a norm that is different from the one they have acquired” (Tagliamonte and D’Arcy 2009: 64). Most studies follow the model in (5), adhering to monostylistic, caregiver-matching through around age 4, though there is much discussion as to type of variable and experience with the input (Guy and Boyd 1990; Kerswill 1996). Style-shifting (e.g. informal use of *you gotta* vs. formal *you must* by social situation) in particular seems to require more experience with the input in order to work out the patterns of distribution (Chevrot et al. 2011), and may also require metalinguistic development.

However, recent work has lowered the age of demonstrated sensitivity to sociolinguistic variation to below age 4, as measured by spontaneous production data⁶. This sensitivity is embedded in the most active timeframes of the first language acquisition process (Díaz-Campos 2005; Roberts 1994; Smith et al. 2007, 2013). The conclusion from even the youngest age groups is that while there is variation by caregiver-child dyad and by variable type (e.g. lexical vs.

⁶ This calls to mind bilingualism research that shows code-switching sensitive to knowledge about language associated with the interlocutor and the activity from the beginning (e.g. Pearson et al. 1995); this includes intra-sentential codeswitching and thus may be the same thing as bidialectal or intraspeaker variation.

morphosyntactic), “sociolinguistic norms are evident from the earliest stages” (Smith et al. 2013: 287). Chambers (2003: 173) also suggests that variable rules and categorical rules are acquired at the same time⁷.

Smith et al. (2013; see also 2007), focus on the community of Buckie, Scotland, where adults are bi-dialectal speakers of the local dialect and the standard dialect. They studied spontaneous production samples from 29 child-caregiver dyads, covering child ages from 2;10–4;02, looking at 6 variables, phonological, lexical and morphosyntactic. They ask, “[w]hen [do young children] start to acquire the highly complex patterns of variation, both linguistic and sociolinguistic, widely attested in adult speech?” (2013: 286). In (7) I give one of their examples of variable input for negation forms in a child-mother interaction.

- (7) (Luke) *Take it off!* (Molly) *No **don't** no! **Dinna** now! Just leave that. Just hold on. No **dinna**! You'll break it. **Don't** do that!* (Luke) *How?*

(Luke, age 3;5, and mother Molly)

Smith et al. (2013) found that children in the community are spontaneously producing both standard and local variants of variables in their input from earliest samplings, prior to the third birthday, with a tendency to match the rates in the child-directed speech for most variables, notably not the belief verbs (*know* vs. *ken*), notoriously complex to acquire from the input (e.g. Gleitman 1990; Dudley 2017), nor the most complex grammatical pattern (Northern Subject Rule). They conclude that children learn patterns of variation from the get-go, “in tandem with language acquisition more generally”, albeit with some caveats.

Naturalistic studies show children have familiarity with many social and linguistic constraints on variable patterns from as early as we can see this reflected in production. These results fit with acquisition evidence more generally: the input determines which language(s) children learn, and by and large, children are remarkable statistical learners, including over variable and

⁷ Smith et al. (2013) argue their data support this, however they do not present any categorical rules, nor compare results to children whose input contains only the standard variant. Going forward, I could see sociolinguistic studies functioning like bilingual studies where time-course and the role of amount of input are relevant to the research question. For example, to study a French-English bilingual to see if differences in the use of the definite determiner delay learning in one or both languages one must compare to English and French monolingual acquisition of the same phenomenon. Otherwise, one can't conclude whether differential acquisition is a result of less input in one language, transfer, or simply the ecology of the determiner system of French vs. English, without reference to bilingualism.

bilingual situations (e.g. Werker and Byers-Heinlein 2008). Kerswill (1996: 199) makes the point that linguistic variables vary in complexity, and consistent with a common general acquisition assumptions, more complex phenomena should be later acquired.

3.3 Learning too well not to change?

A set of studies using experimental methods has asked a question mixing insights from Sections 3.1 and 3.2. (Cournane 2015; Cournane and Pérez-Leroux under revision; Hall and Maddeaux 2018), do young children overshoot the adult pattern in some variable-learning cases? The reasoning is that if children drive incrementation, we expect them to go beyond the input in ways consistent with directional patterns in language change but motivated from something other than change phenomena themselves. One approach for why children and pre-adolescents drive incrementation is that they are aware of the distribution of variants in speakers older than them (i.e. age-stratification of variation⁸). The approach taken here contrasts with that approach (but is not mutually exclusive), as I aim to address a shortcoming of extant models (like (5)): there is no primary mover, actuation is largely inscrutable. Where does the initial directional shove off come from⁹? How does age-stratification come to be in the first place?

Cournane (2015) and Cournane and Pérez-Leroux (under revision) tested the hypothesis that learners would prefer the historically innovative interpretation of a functional modal when the structure was ambiguous. Variable-meaning modals offer us the opportunity to study incrementation in the selection of the epistemic senses over root for variable-meaning modals, known to gradually shift from root-to-epistemic in diachrony (Traugott 1989). This is not lexical variation (i.e. there is only one *must*; Kratzer 1981; Hacquard 2006); other syntactic and semantic properties of the utterance can render the modal

⁸ The age-stratification hypothesis for children is untested. It can be tested with perceptual methods like eye tracking (see Koops et al. 2008). Cournane and Mackenzie (in prep) propose using well-described phonetic and lexical community variables with 3–6 year-old children. Children would listen to audio stimuli of a voice producing words containing this variable and fixation and gaze-path would be recorded. If age-grading of linguistic variables is something young children are tracking in their input, then they will be more likely to look to older faces when hearing conservative variants, and younger faces when hearing innovative variants.

⁹ Note this is not the Actuation Problem, which is about the when of change; this is the what of change – why are grammars so prone to the same kinds of innovations catching on (e.g. Hopper and Traugott 2003; Roberts and Roussou 1999; van Gelderen 2004)?

categorical (8a), or probabilistically suggest one interpretation over the other (8b, root preference).

- (8) a. Modal+Aspect: Doggy **must've** eaten his dogfood. *Root, Epistemic
 b. ModalOnly: Doggy **must** eat his dogfood. Root, Epistemic (Habitual)

Modal interpretation is constrained by syntax (or, more precisely, combinatorial semantics; see Condoravdi 2002; Hacquard 2006; Klecha 2016; van Dooren et al. 2017): when the verbal complement has grammatical aspect marking (8a; perfect), only an epistemic interpretation is available (adding temporal adverbs like *by tonight* to forward-shift time of the prejacet can render the prejacet future-oriented and get a root interpretation; what matters is the temporal semantics the construction allows). If the verbal complement has a bare verb, the structure is indeterminate as to the interpretation of the modal (8b). In (8b) *must* can receive a root interpretation or an epistemic one (with a habitual construal). The incrementation hypothesis, with children as the prime mover, predicts that in contexts where meaning-variability is possible (8b), children will favor epistemic interpretations at higher rates than young adults from the same speech community¹⁰.

We used a picture preference task (hear one sentence, select between pictures representing interpretations) to test preferred interpretations (root vs. epistemic) of sentences like (8a) and (8b) (Figure 1). We ran 54 monolingual Toronto English children between 3;0–6;2, and 10 dialect-matched young adults (aged 18–25). Each participant gave judgments for 16 sentences, 8 with aspect and 8 without.

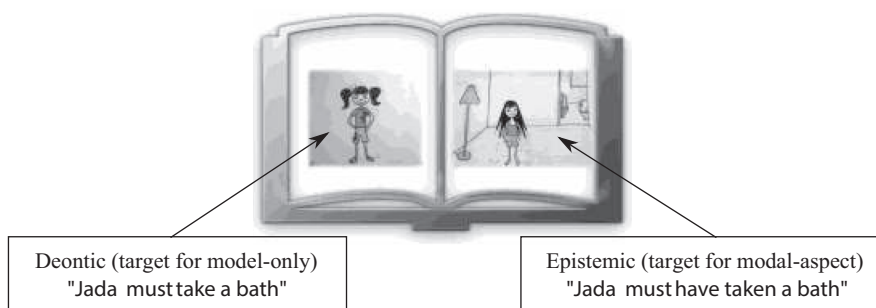


Figure 1: Target pictures for each test-sentence.

¹⁰ This is about competence: which LF interpretation of *must*, that is which syntax – root or epistemic, is ascribed to a sentence?

Results showed adults behaved as expected, showing a significant contrast by sentence condition. Aspect-marked sentences were at ceiling for epistemic interpretations, sentences with bare verbs were variable, but with a preference for root interpretations. Young children showed a slight overall root bias (in line with early usage patterns, and not matching input; van Dooren et al. 2017), but 5-year-old children had “flipped” to an across-the-board highly significant epistemic preference, several preferring epistemic interpretations 100% of the time.

These results begged the question of whether the 5-year-olds had the variable meaning of *must*, or had analyzed this modal as *lexically* epistemic (production data suggests this is not what is happening; Cournane 2015; van Dooren et al. 2017). Cournane & Pérez-Leroux ran a second study on the deontic interpretation of *must*, in comparison to the same sentence, unmodalized (Deontic: *The boys must wash their hands with soap* vs. Actual: *The boys wash their hands with soap*). All children, even 5-year-olds, were significantly sensitive to the deontic meaning contribution of *must*, and in qualitative comments some children gave explicit insights into their form-meaning relations (9). This strongly suggests they understand the bare-verb allows a root interpretation (even if by age 5 they dis-prefer this compared to an epistemic interpretation).

- (9) a. Not different because *must* means *maybe they will*. (Age 5)
 b. *Must* means *have to*. (Age 4)

Cournane & Pérez-Leroux argued that 5-year-olds showed a bias toward epistemic interpretations of *must* for ambiguous bare-verb sentences, apparently overriding grammatical constraints (these are linked to semantic contributions of the modal's preadjacent) present for adult speakers. Why? When children learn the grammatical basis for a pattern of variation, they lean in, overgeneralizing the grammatical pattern they have learned in the direction consistent with regularization (similar in spirit to rule-learning; e.g. Schuler et al. 2016). Returning to the grammatical facts, linguists have amassed a wealth of cross-linguistic evidence showing that functional modal interpretations are linked to syntactic positions in the following way: when root, these modals are represented below tense and aspect, while when epistemic they are represented above tense and aspect (Cinque 1999; Hacquard 2006). This is schematized in (10).

- (10) [Subject [*must*_{EPISTEMIC} [Tense [Aspect [*must*_{ROOT} [Verb Phrase]]]]]]

Relevant to our study, *must* sentences with grammatical aspect in English (8a) show *must* overtly preceding aspect, giving a cue to a high epistemic interpretation. By contrast, bare verb constructions (8b) are ambiguous as to the scope of *must*.

Epistemic meanings for variable-meaning modals appear to be learned later than root meanings (even for *must*, despite epistemic uses being more prevalent in the input, the syntactic complexity makes its epistemic structure later acquired; Cournane 2015; cf. van Dooren et al. 2017). Once children reliably map *must* to its higher epistemic modal position in their 4th year of life (Cournane 2015; Papafragou 1998), they overregularize this representation to ambiguous sentences.

Turning to sound variables, Hall and Maddeaux (2018) tested Toronto English-speaking children ($n = 24$; aged 4–12 years) and caregivers ($n = 20$), and non-parents as controls ($n = 16$), on a task eliciting multiple productions of words containing the /æ/ and /u/ phonemes. Toronto English has variation in the production of these vowels, with patterns of /æ/-raising and /u/-fronting. /æ/-raising is a stable part of the Canadian Shift that does not seem to be linked to speaker age (Boberg 2011), while /u/-fronting is part of a change-in-progress led by young female speakers and more advanced in Toronto (urban center) than less urban areas (Boberg 2011). Hall and Maddeaux tested (a) whether children first replicate their parents system and later shift to a different norm (in line with Labov's model in (5)), or incrementation occurs from the beginning, (b) when and how incrementation, operationalized as overgeneralization, occurs in child production, and (c) whether patterns would differ for stable variation (realizations of /æ/) compared to change-in-progress (realizations of /u/). They predicted that the youngest group should match adults most closely, but that by preadolescence the children should be more advanced than the caregivers (i.e. produce more fronted-/u/ than caregivers). These are the basic predictions if, "[b]y the age of stabilization, every cohort reaches a higher level of change than their immediate elders and younger peers." (Tagliamonte and D'Arcy 2009: 69).

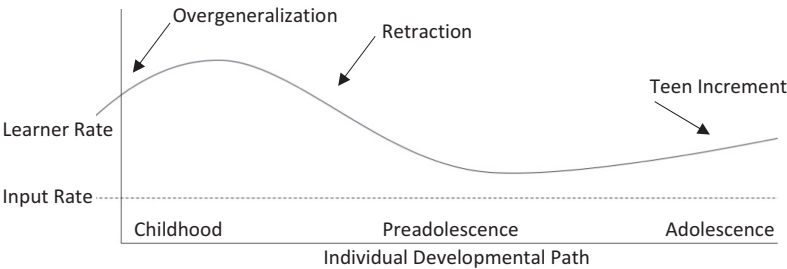
Results show that for the /u/-fronting change-in-progress, the youngest children are not matching parents but are in fact the most fronted and least adult-like. Older children are less front than younger, but more front than their caregivers. Hall and Maddeaux suggest that on the individual level this suggests that children retract, but because they started more front, they nonetheless are still more front than caregivers in preadolescence. They call this a "two steps forward, one step back model". For the stable variation, /æ/-raising, which is also a more lexically conditioned sound-change, they likewise found that the youngest group was the most raised, overshooting the input, but only before nasals (e.g. *ban*). The oldest children were adult-like, having acquired the pattern but not incremented it. They conclude that for /u/-fronting children by age 4 have a production pattern consistent with incrementation, and they are not faithfully replicating the parental pattern. By age 10 children show a pattern for a change-in-progress that will increment the change on the community level, despite the fact that these kids are themselves retracting from an L1A-driven internally driven overshooting of the

input community grammar (contra Roberts and Labov 1995). For the stable variation, which is also less productive in the input grammar (likely, on the current view, to play a role in why it is stable), children also initially overshoot somewhat (suggesting this is a thing children do, but in some cases their overgeneralizing gets overwritten promptly on the basis of further learning), but settle on matching the idiosyncrasies of their input.

4 Overgeneralization model of incrementation & discussion

Taken together, these lines of research show that children generalize when they discover the basis for a rule or other systematic relation in their language, and then gradually retract by learning sub-regularities, exceptions, blocking factors, or other factors governing the selection of one form (or meaning) over another. In production studies targeting the acquisition of variation from the input, we see that children generally input-match by 3 or 4 years old. However, experimental studies of variation have found a combination of these results, showing children command variation but with input-divergent overgeneralization of the more innovative variant. These findings for child overgeneralization are in the direction of change in progress, supporting the hypothesis that children’s overgeneralizations are of the right type to drive incrementation. Hall and Maddeaux (2018) suggest that children overgeneralize and then retract, but retraction still leaves them ahead of their input (“two steps forward, one step back”). This is consistent with the modal study as well, but Courname and Pérez-Leroux (under revision) tested only up to 6 years-old. These findings are the starting point for a Reverse-U Model of Incrementation (11), stressing the developmental path as an internal source for change directionality.

- (11) Reverse-U Model of Incrementation, plotting innovative variant in the Individual



The precise age of overgeneralization will vary by domain and complexity of the variable, so I have left it open-jaw. What matters is three things: (a) overgeneralization when the generalization underlying input patterns is discovered (e.g. modal relationship to language specific syntax, conditions for rule-application), (b) input-divergence with respect to the “frequency, extent, or specificity of a sociolinguistic variable” (Labov 2007: 346) in the input (the size of the overshooting can vary by domain and measuring time), and (c) gradual retraction. Retraction could vary in slope (what kind of learning corrects the child’s production patterns to align with the input? E.g. positive evidence that wasn’t available earlier? Indirect positive evidence?). Survival or spread to other speakers may or may not happen (imagine settling on the input line), as with the parallel view of innovation supported here, but incrementation is by definition about change in progress and presupposes actuation and in turn innovation. To get the adolescent peak (e.g. Cedergren 1988; Tagliamonte and D’Arcy 2009), following the internally driven initial shove, it may be that socially motivated incrementation takes over (see Holmes-Elliott 2018), drawing on late childhood grammatical competence.

Notice this model is a reversed U-shape model, because the interest is on input-divergence (the drop to the bottom of the U on the traditional model), rather than input-coherence. In this model, incrementation is “two steps forward” (because of acquisition overshooting), and “one step back” (because of increased learning or failure of negative evidence or social-reinforcement among peers), inspired from Hall and Maddeaux (2018), and consistent with Cournane and Pérez-Leroux (under revision). Like the U-shape model for rule learning, (11) is a simple descriptive sketch of the dynamics of learning over time relative to the target grammar, and is not expected to apply to all learning, but at least to those kinds that rely on generalization and extension (for the purposes of productivity). It seems to complicate things compared to Labov’s model given in (5), as the initial learning is not aligned to the caregiver pattern, but to the child’s discovery of a pattern’s basis in the input, which often involves overgeneralization, and then we see retraction (with further learning) and social “lift-off” from there, if the variant gets linked to social factors. This may complicate things, but it takes into consideration child learning behaviors not previously accounted for, and addresses the prime mover problem (directionality) by appealing to grammar.

Stability is not something child language is known for, as learning is by definition unstable. Why then do many prior studies of child production find input-matching? There are several potential answers that merit further discussion. The simplest is that for these variables, by the time the researchers studied the variation, the child had already worked out the patterns and, if there were

errors along the way, these had been corrected. For sound variables, much learning is done “under the hood” in the first year of life, so sound variables might be less visibly divergent in preschoolers¹¹. Another related thought is that different kinds of input variation (i.e. relative to what conditions the variation) may be easier or harder to learn (see Kerswill 1996). Finally, the methods differed, where overgeneralization is forefronted in targeted experimental work (but also exists in natural data), input-matching is observed in natural samples. This is a key contrast, as production and comprehension can be viewed as two sides of the same speaker for adults, but show sometimes drastic mismatches in development (e.g. Hendriks and Koster 2010). Why is this? Production data is conservative (not in the historical sense, in the acquisition sense): it obscures child grammar states on a whole because language is Zipfian and samples collect high frequency types, and this is exacerbated in children because they are well known to exhibit avoidance (see Snyder 2007), and therefore production samples, especially small ones, may lead to false negatives when we consider input-divergence or overgeneralizations. Experiments give concentrated, controlled, data, but run the risk of task effects.

How does the view on learning here relate to preadolescent and adolescent incrementation? Does cumulative overgeneralization over many children continue to push the change in the innovative direction? Or, does overgeneralization only provide the initial shove, from which preadolescent and adolescents advance the change-in-progress and then age-stratification is in place and child learning overgeneralizations are no longer necessary (i.e. they are incidental to already established changes in progress)? If children continue to overgeneralize in similar ways over generations, it seems plausible that this could maintain the variant’s association with younger speakers vs. older. And then, we can still look at overgeneralization for biases consistent with change patterns. Work tracking cohorts will be essential to helping answer these questions about the transition from childhood to adolescence (e.g. Holmes-Elliott 2018).

In this approach, there is no teleology of change, nor necessity for learners to be aware of language history, but there is a proposed source and it helps to ground why innovative variants are such (e.g. with modals it is linked to syntactic representation). The approach stresses the relation between the grammatical relations posited in children’s grammars, and the input-divergent productive patterns these engender. This is the link between the I-languages of children and child usage patterns: “interim” grammars are productive, and

11 This however, presents a ph- vs. syntax asymmetry that may suggest ph-variables would not increment because kids would learn before social reinforcement or spread were possible. However, Hall and Maddeaux’ data provide counterevidence here.

substantively different from the grammar of the input. This creates: (a) posited relations that overshoot the input, because learning requires going beyond the input, and (b) productive language from which others – peer listeners – could learn, biasing the input toward child-produced patterns. The learning itself in this approach fits with discrete views of language change, but learning states affect usage patterns. A major challenge remains about formally clarifying which kind of variables children may input-match from early on¹² and never overgeneralize or mis-learn otherwise, compared to which ones they overshoot, what the formal nature of the overgeneralization is, and how retraction happens (or fails to happen).

Hale (1998: 15), articulates a common I-Language view of diachronic innovation: “The real mystery in change is how the evidence provided in the [input] for recurrent and presumably richly attested elements – occurrences of a particular phoneme (e.g. *s*) or a functional head – can come to be sufficiently opaque to license change”. This needle-in-the-haystack innovator problem is solved, or more accurately, rendered a non-problem by a deeper look at acquisition. Explaining the actuation (and diffusion) of innovations can be found in “mundane” input-divergence, the bread and butter of developmental linguists. The triggering experience – e.g. the nature of the input – is still critical (in line with Lightfoot 1979, 1999) on this view, but explaining actuation is rendered not about trying to describe the perfect input-triggering experience for a child to posit a precious innovation, but about trying to describe why among rampant input-divergence, some innovations stay in the grammars of learners. Now, to find innovations that actuate or increment, we look to what didn’t get corrected to match the input, or what child output got considered as input by other children (perhaps peers or younger siblings; see Hockett 1950). The directional force should be (roughly) related to the properties of the input information (which changes over developmental time) and the child’s learning strategies and prior learning. The social lives of children and preadolescents are a good place to look for what supports vs. overwrites input-divergence (Smith et al. 2013; Holmes-Elliott 2016, 2018, i.a.).

¹² Perhaps ones that are straightforward to track early on, and e.g. do not rely heavily on word learning or prior within-system learning to bootstrap from; also bear in mind, infants spend 2 years doing input-analysis before production shows multiword utterances of the type sociolinguists have primarily examined. It’s possible most learning in phenomena that are accessible early on is complete, or mostly complete (but considering production data is a poor measure for infrequent patterns, this would not be readily apparent).

5 Conclusion

In this paper, we focused on whether, on the basis of child language learning behaviours for variable input from three lines of literature, it is reasonable to expect the incrementation of new variants to be kicked off by and sustained by the acquisition process. I have argued that when innovation is defined as input-divergent representations or analyses of form-meaning relations, we can conclude that all children innovate. Innovation is grammatical learning; when defined this way, innovation is universal, as learning involves change over time and generalization from finite input to a productive grammatical system. Furthermore, children use their languages along the learning path, contributing to the linguistic community patterns. On the basis of new experimental studies revealing how children overshoot the input, I've argued that incrementation can also arise from input-divergence and sketched a reverse-U model. A theory of what creates new variants can give incrementation too (at least, the initial thrust, and social factors may take over from there). Much work remains to formally link the analogous dynamic phenomena in learning and change, paying close mind to developmental challenges each area of language presents the child, and how children respond to these challenges, so we can further assess theories of explanation for change.

References

- Andersen, H. 1973. Abductive and deductive change. *Language* 49. 765–793.
- Berko, J. 1958. The child's learning of English morphology. *Word* 14(2–3). 150–177.
- Boberg, C. 2011. Reshaping the vowel system: An index of phonetic innovation in Canadian English. *University of Pennsylvania Working Papers in Linguistics* 17(2). 4.
- Booth, A. E. & S. R. Waxman. 2002. Word learning is 'smart': Evidence that conceptual information affects preschoolers' extension of novel words. *Cognition* 84(1). B11–B22.
- Brennan, V. 1993. *Root and epistemic modal auxiliary verbs*. University of Massachusetts PhD.
- Cedergren, H. 1988. The spread of language change: Verifying inferences of linguistic diffusion. In P. H. Lowenberg (ed.), *Language spread and language policy: Issues, implications, and case studies*, 45–60. Washington, DC: Georgetown University Press.
- Chambers, J. 2003. *Sociolinguistic theory*. Oxford: Blackwell.
- Chevrot, J.-P., A. Nardy & S. Barbu. 2011. Developmental dynamics of SES-related differences in children's production of obligatory and variable phonological alternations. *Language Sciences* 33(1). 180–191.
- Chomsky, N. 1986. *Knowledge of language: Its nature, origin, and use*. Westport, CT: Greenwood Publishing Group.
- Cinque, G. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. Oxford: OUP.
- Clark, E. V. 1993. *The lexicon in acquisition*. Cambridge: CUP.

- Condoravdi, C. 2002. Temporal interpretation of modals. *The Construction of Meaning*, 59–88.
- Cournane, A. 2014. In search of L1 evidence for diachronic reanalysis: Mapping modal verbs. *Language Acquisition* 21(1). 103–117.
- Cournane, A. 2015. *Modal development: input-divergent L1 acquisition in the direction of diachronic reanalysis*. University of Toronto (Canada) PhD thesis.
- Cournane, A. 2017. In defense of the child innovator. In E. Mathieu & R. Truswell (eds.), *Micro-change and Macro-change in Diachronic Syntax*, 10–23. Oxford: OUP.
- Cournane, A. & L. Mackenzie. in prep. SocialEyes: Testing age-stratification in child perception.
- Cournane, A. & A. T. Pérez-Leroux. under revision. Leaving obligations behind: epistemic incrementation in preschool English.
- Crisma, P. & G. Longobardi. 2009. Change, relatedness and inertia in historical syntax. In P. Crisma & G. Longobardi (eds.), *Historical syntax and linguistic theory*, 1–13. Oxford: OUP.
- Díaz-Campos, M. 2005. The emergence of adult-like command of sociolinguistic variables: a study of consonant weakening in Spanish-speaking children. *Selected proceedings of the 6th conference on the acquisition of Spanish and Portuguese as first and second Languages*, 56–65.
- Diessel, H. 2011. Grammaticalization and language acquisition. In H. Narrog & B. Heine (eds.), *Handbook of grammaticalization*, 130–141. Oxford: OUP.
- Dudley, R. E. 2017. *The role of input in discovering presuppositions triggers: Figuring out what everybody already knew*. College Park: University of Maryland PhD thesis.
- Fischer, O. 2003. The development of the modals in English: Radical versus gradual changes. In D. Hart (ed.), *English Modality in Context*, 17–32. Bern: Peter Lang.
- Gleitman, L. R. 1990. The structural sources of verb meanings. *Language Acquisition* 1(1). 3–55.
- Guy, G. R. & S. Boyd. 1990. The development of a morphological class. *Language Variation and Change* 2(1). 1–18.
- Hacquard, V. 2006. *Aspects of modality*. Cambridge, MA: M.I.T. PhD thesis.
- Hacquard, V. 2013. The grammatical category of modality. *Proceedings of the 19th Amsterdam Colloquium*, 19–26.
- Hale, M. 1998. Diachronic syntax. *Syntax* 1(1). 1–18.
- Hall, E. & R. Maddeaux. 2018. /u/-fronting and /æ/-raising in Toronto Families. Talk Presented at New Ways of Analyzing Variation (NWAV) 47. New York University.
- Halle, M. 1964. Phonology in generative grammar. In J. Fodor & J. Katz (eds.), *The structure of language*, 334–352. Englewood Cliffs, NJ: Prentice-Hall.
- Hendriks, P. & C. Koster. 2010. Production/comprehension asymmetries in language acquisition. *Lingua* 120(8). 1887–1897.
- Heycock, C. & J. Wallenberg. 2013. How variational acquisition drives syntactic change. *The Journal of Comparative Germanic Linguistics* 16(2–3). 127–157.
- Hockett, C. F. 1950. Age-grading and linguistic continuity. *Language* 26(4). 449–457.
- Holmes-Elliott, S. 2016. Ladies first? Adolescent peaks in a male-led change. *University of Pennsylvania Working Papers in Linguistics* 22(2). 81–90.
- Holmes-Elliott, S. 2018. Future Leaders: real time incrementation of sound change between childhood and adolescence. Talk presented at the Department of Linguistics and English Language, University of Manchester. December 11, 2019.
- Hopper, P. J. & E. C. Traugott. 2003. *Grammaticalization*. Cambridge: CUP.
- Hudson Kam, C. L. & E. L. Newport. 2005. Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development* 1(2). 151–195.

- Hudson Kam, C. L. & E. L. Newport. 2009. Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology* 59(1). 30–66.
- Kerswill, P. 1996. Children, adolescents, and language change. *Language Variation and Change* 8(2). 177–202.
- Kerswill, P. & A. Williams. 2000. Creating a new town koine: Children and language change in Milton Keynes. *Language in Society* 29(1). 65–115.
- Klecha, P. 2016. Modality and embedded temporal operators. *Semantics and Pragmatics* 9. 9–1.
- Koops, C., E. Gentry & A. Pantos. 2008. The effect of perceived speaker age on the perception of PIN and PEN vowels in Houston, Texas. *University of Pennsylvania Working Papers in Linguistics* 14(2). 93–101.
- Kratzer, A. 1981. The notional category of modality. In: H.-J. Eikmeyer & H. Rieser (eds.), *Words, Worlds, and Contexts*, 38–74. Berlin: Mouton de Gruyter.
- Kroch, A. S. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1(3). 199–244.
- Labov, W. 1970. The logic of nonstandard English. In F. Williams (ed.), *Language and poverty*, 153–189. Amsterdam: Elsevier.
- Labov, W. 1989. The child as linguistic historian. *Language Variation and Change* 1(1). 85–97.
- Labov, W. 2001. Principles of linguistic change, Volume 2: Social factors. New York, NY: John Wiley & Sons.
- Labov, W. 2007. Transmission and diffusion. *Language* 83(2). 344–387.
- Lightfoot, D. 1979. *Principles of diachronic syntax*. Cambridge: CUP.
- Lightfoot, D. 1999. *The development of language: Acquisition, change, and evolution*. Oxford: Wiley-Blackwell.
- MacWhinney, B. & C. Snow. 1985. The child language data exchange system. *Journal of Child Language* 12(2). 271–295.
- Marcus, G. F., S. Pinker, M. Ullman, M. Hollander, T. J. Rosen, F. Xu & H. Clahsen. 1992. Overregularization in language acquisition. *Monographs of the Society for Research in Child Development* 57. i–178.
- Marcus, G. F., S. Vijayan, S. B. Rao & P. M. Vishton. 1999. Rule learning by seven-month-old infants. *Science* 283(5398). 77–80.
- Meillet, A. 1912. L'évolution des formes grammaticales. In A. Meillet (ed.), *Linguistique Historique et Linguistique Générale*, 130–158. Paris: Champion.
- Papfragou, A. 1998. The acquisition of modality: Implications for theories of semantic representation. *Mind & Language* 13(3). 370–399.
- Pearson, B. Z., S. Fernández & D. K. Oller. 1995. Cross-language synonyms in the lexicons of bilingual infants: One language or two? *Journal of Child Language* 22(2). 345–368.
- Pérez-Leroux, A. T., A. Munn, C. Schmitt & M. Delrish. 2004. Learning definite determiners: Genericity and definiteness in English and Spanish. *BUCLD 28 Proceedings Supplement*.
- Pinker, S. 1984. *Language learnability and language learning*. Cambridge, MA: Harvard University Press.
- Roberts, I. & A. Roussou. 1999. A formal approach to “grammaticalization.”. *Linguistics* 37(6). 1011–1041.
- Roberts, I. G. 1985. Agreement parameters and the development of English modal auxiliaries. *Natural Language & Linguistic Theory* 3. 21–58.
- Roberts, J. & W. Labov. 1995. Learning to talk Philadelphian: Acquisition of short *a* by preschool children. *Language Variation and Change* 7(1). 101–112.

- Roberts, J. L. 1994. Acquisition of variable rules: (-t, d) deletion and (ing) production in preschool children. University of Pennsylvania PhD thesis.
- Robinson, M. 2019. Child L1 Acquisition of Comparatives and Superlatives: Evidence for *ABA? New York University Ms.
- Roeper, T. 1999. Universal bilingualism. *Bilingualism: Language and Cognition* 2(3). 169–186.
- Saffran, J. R., R. N. Aslin & E. L. Newport. 1996. Statistical learning by 8-month-old infants. *Science* 274(5294). 1926–1928.
- Schuler, K. D., C. Yang & E. L. Newport. 2016. Testing the Tolerance Principle: Children form productive rules when it is more computationally efficient to do so. *CogSci*.
- Senghas, A. & M. Coppola. 2001. Children creating language: How Nicaraguan sign language acquired a spatial grammar. *Psychological Science* 12(4). 323–328.
- Singleton, J. L. & E. L. Newport. 2004. When learners surpass their models: The acquisition of American Sign Language from inconsistent input. *Cognitive Psychology* 49(4). 370–407.
- Smith, J., M. Durham & L. Fortune. 2007. “Mam, my trousers is fa’in doon!”: Community, caregiver, and child in the acquisition of variation in a Scottish dialect. *Language Variation and Change* 19(1). 63–99.
- Smith, J., M. Durham & H. Richards. 2013. The social and linguistic in the acquisition of sociolinguistic norms: Caregivers, children, and variation. *Linguistics* 51(2). 285–324.
- Snyder, W. 2007. *Child language: The parametric approach*. Oxford: OUP.
- Tagliamonte, S. 2011. *Variationist sociolinguistics: Change, observation, interpretation*, vol. 40. New York, NY: John Wiley & Sons.
- Tagliamonte, S. A. & A. D’Arcy. 2007. The modals of obligation/necessity in Canadian perspective. *English World-Wide* 28(1). 47–87.
- Tagliamonte, S. A. & A. D’Arcy. 2009. Peaks beyond phonology: Adolescence, incrementation, and language change. *Language* 85(1). 58–108.
- Traugott, E. C. 1989. On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language* 65(1). 31–55.
- Traugott, E. C. 2011. Modality from a historical perspective. *Language and Linguistics Compass* 5(6). 381–396.
- van Dooren, A., A. Dieuleveut, A. Cournane & V. Hacquard. 2017. Learning what must and can must and can mean. *Proceedings of the 21st Amsterdam Colloquium*.
- van Gelderen, E. 2004. *Grammaticalization as economy*. Amsterdam: John Benjamins.
- van Gelderen, E. (ed). 2009. *Cyclical change*. Amsterdam: John Benjamins.
- Walkden, G. 2012. Against inertia. *Lingua* 122(8). 891–901.
- Walkden, G. 2017. The actuation problem. In A. Ledgeway & I. Roberts (eds.), *The Cambridge handbook of historical syntax*. Cambridge: CUP.
- Wallenberg, J. C. 2013. A unified theory of stable variation, syntactic optionality, and syntactic change. *Talk delivered at the 15th Diachronic Generative Syntax (DiGS) Conference*, University of Ottawa.
- Weinreich, U., W. Labov & M. Herzog. 1968. *Empirical foundations for a theory of language change*. Austin, TX: University of Texas Press.
- Werker, J. F. & K. Byers-Heinlein. 2008. Bilingualism in infancy: First steps in perception and comprehension. *Trends in Cognitive Sciences* 12(4). 144–151.
- Xu, F. & S. Pinker. 1995. Weird past tense forms. *Journal of Child Language* 22(3). 531–556.
- Yang, C. D. 2000. Internal and external forces in language change. *Language Variation and Change* 12(03). 231–250.