



Syntax acquisition

Stephen Crain* and Rosalind Thornton

Every normal child acquires a language in just a few years. By 3- or 4-years-old, children have effectively become adults in their abilities to produce and understand endlessly many sentences in a variety of conversational contexts. There are two alternative accounts of the course of children's language development. These different perspectives can be traced back to the nature versus nurture debate about how knowledge is acquired in any cognitive domain. One perspective dates back to Plato's dialog 'The Meno'. In this dialog, the protagonist, Socrates, demonstrates to Meno, an aristocrat in Ancient Greece, that a young slave knows more about geometry than he could have learned from experience. By extension, Plato's Problem refers to any gap between experience and knowledge. How children fill in the gap in the case of language continues to be the subject of much controversy in cognitive science. Any model of language acquisition must address three factors, *inter alia*:

1. The knowledge children accrue;
2. The input children receive (often called the PRIMARY LINGUISTIC DATA);
3. The nonlinguistic capacities of children to form and test generalizations based on the input.

According to the famous linguist Noam Chomsky, the main task of linguistics is to explain how children bridge the gap—Chomsky calls it a 'chasm'—between what they come to know about language, and what they could have learned from experience, even given optimistic assumptions about their cognitive abilities. Proponents of the alternative 'nurture' approach accuse nativists like Chomsky of overestimating the complexity of what children learn, underestimating the data children have to work with, and manifesting undue pessimism about children's abilities to extract information based on the input. The modern 'nurture' approach is often referred to as the USAGE-BASED account. We discuss the USAGE-BASED account first, and then the NATIVIST account. After that, we report and discuss the findings of several studies of child language that have been conducted with the goal of helping to adjudicate between the alternative approaches to language development. © 2011 John Wiley & Sons, Ltd.

How to cite this article:

WIREs Cogn Sci 2011. doi: 10.1002/wcs.1158

THE USAGE-BASED (NURTURE) APPROACH TO LANGUAGE ACQUISITION

This solution to Plato's Problem views language development on a par with the acquisition of knowledge in other cognitive domains: for example,

social skills, learning to count, learning to read, and so forth. So, the nurture approach invokes DOMAIN-GENERAL learning mechanisms to explain language development. These domain-general learning mechanisms embody general learning processes that are not specially tailored to acquire any particular kinds of facts about the world. Like knowledge in other cognitive domains, knowledge of language is accrued in a piecemeal fashion, based on statistical regularities in the input.

*Correspondence to: Stephen.Crain@mq.edu.au

ARC Centre of Excellence in Cognition and its Disorders, Macquarie University, North Ryde, New South Wales, Australia

Advocates of the usage-based account highlight the availability of relevant cues in the input to children. These cues serve as the basis for the generalizations that children form about language. These generalizations are formed using general purpose learning mechanisms including distributional analysis, analogy, cut and paste operations, and the like. The products of these learning algorithms are ‘shallow’ records, which children keep of their linguistic experience. These are piecemeal records of construction types (a.k.a. templates/schemas/constructs) that encode linguistic patterns that are displayed in the input. Construction types are concatenated sequences of words, combined with category labels, such as *NP*, *V*, *neg*, *INF*, *P*, that are drawn from an intuitively simple typology. Constructions are learned solely from positive evidence.^{1–3} When children’s generalizations extend beyond their experience, the supposition is that this is just an instance of a completely general problem of induction, one that arises for all learning which involves projecting beyond one’s experience.⁴ According to Pullum and Scholz³ linguists need not suppose that children are innately endowed with ‘specific contingent facts about natural languages’. If the data available to children are rich enough for them to determine the structures of human languages, given the right inferential techniques, then appeals to innately specified principles are at best a useful crutch for theorists—and at worst a source of erroneous claims about alleged ‘gaps’ between the facts concerning particular human languages and the evidence that is available to children.

According to the usage-based account, all human languages contain a wide range of semi-idiosyncratic constructions that cannot be accounted for by universal, or innate linguistic principles. On any account of language development, these ‘peripheral’ constructions must be learned. According to the usage-based account, the same mechanisms that children use to learn these constructions are also used to learn the CORE PHENOMENA of human languages. The reasoning here is that the core phenomena of human languages are even more regular, and occur more frequently than the idiosyncratic patterns. If so, then the core phenomena should be even easier to learn,² with more frequently attested constructions being mastered earlier than less frequently attested constructions.⁵ Here is a representative quote from Goldberg²:

‘Crucially, all linguists recognize that a wide range of semi-idiosyncratic constructions exist in every language, constructions that cannot be accounted for by general, universal, or innate principles or constraints. Generative linguists argue that these constructions exist only on the “periphery” or “residue” of language—that they need not be the focus of linguistic

or learning theorists. Constructionists on the other hand have zeroed in on these constructions, arguing that whatever means we use to learn these patterns can easily be extended to account for so-called “core” phenomena. In fact, by definition the core phenomena are more regular, and tend to occur more frequently within a given language as well. Therefore, if anything, they are likely to be easier to learn.’ (p. 14)

Once children have mastered the core construction types, these are merged into more and more complex patterns, until the language of the child approximates that of an adult in the same linguistic community. On the usage-based account, then, child language is expected to match that of adults, more or less. Initially, child language will be a less articulated version of the adult language, but children will gradually converge on the target language.

On the usage-based account, linguistic generalizations are based on a combination of linguistic form and communicative function. The function of a construction type is essential in accounting for its distribution in a language. These functions are directly related to meaning. For example, Cameron-Faulkner et al.⁶ (pp. 252, 266) propose that negators (e.g., *no*, *not*) express three functions in child language: NON-EXISTENCE (e.g., *There’s no juice left*), REJECTION (e.g., *I don’t want anymore*), and PROHIBITION (e.g., *No swimming*). When the specific negator, *no*, was followed by a verb in the productions by one child, this *no V* construction was associated with four sub-functions, FAILURE (*No move*), PROHIBITION (*No touch*), REJECTION (*No apple*), and INABILITY (*No reach*). According to Goldberg,² information-theoretic notions also play a role in formulating the meanings of constructions, both for children and adults. These information-theoretic notions include TOPIC (matters of current interest), FOCUS DOMAIN (what is newly asserted), and BACKGROUNDED elements, such as pre-suppositions.

The usage-based account purports that, in tandem, form and function can also explain how children build up relations among constructions at the final stages of language development. The process is outlined by Lieven and Tomasello⁷:

‘Finally, the child has to abstract the relations between constructions. Evidence that this has occurred is that the child is able to transform an utterance in one construction into another construction, for instance a declarative into a *wh*-question or an active into a passive. This could be done by forming a semantic representation of what the speaker wishes to say, thereby allowing the production of the other construction. Whether and when the learner actually maps the form—function mappings of one construction to those

of the other is an empirically open question at the moment. It depends on the metalinguistic expertise and/or educational level of different speakers.' (p. 171)

In addition to generalizations within a language, any viable account of language development owes an explanation of cross-linguistic generalizations. On the usage-based account, cross-linguistic generalizations (recurrent patterns) are the by-product of general cognitive constraints, such as analogical processes, processing factors, and discourse-pragmatic factors.² Importantly for our purposes, the usage-based account anticipates substantial variability among the constructions that appear in different human languages. The usage-based account attempts to avoid the conclusions of nativists about the innate specification of universal linguistic principles. On this account, children only (re)produce linguistic expressions that they have experienced in the input, at least at the earliest stages of language development. This proposal is called CONSERVATIVE LEARNING. If true, conservative learning renders innate linguistic principles unnecessary for language learning. Language development consists, instead, in developing constructions based on exposure to strings of words that learners encounter in their experience.

Conservative Learning and the Truncation Hypothesis

Tomasello⁸ defends the conservative learning model of language acquisition, for verbs. Essentially, young children's productions of verb forms are limited to forms that they have previously encountered in the input, at least for children younger than 3 years. After 3 years, children start to form more abstract adult-like linguistic categories. When children make 'errors', these are purged from children's grammars by (direct or indirect) NEGATIVE EVIDENCE (corrective feedback to children when they produce ill-formed utterances), ENTRENCHMENT (being drowned out by the frequency of a different expression), and PRE-EMPTION (e.g., adult recasts using an alternative expression). These usage-based mechanisms assume the role played by innate constraints on the nativist account.^{4,9,10}

The most detailed defense of conservative learning has been advanced by Tomasello,¹¹ who contends that, before age three, children's productions of verb forms are limited to forms that they have previously encountered in the input, at least:

'...before their third birthdays children use individual verbs and syntactic constructions in just the same way they have heard and understood them being used.' (p. 71)

Concentrating on the specific verbal forms that appear in young children's speech, Tomasello⁸ argues that 'young children's creativity—productivity—with language has been grossly overestimated' (p. 210) and that '[o]ther than the categorization of nominals, nascent language learners possess no other linguistic abstractions or forms of syntactic organization'. (p. 214).

At a first look, the literature on children's spontaneous productions seems replete with examples that pose a challenge to conservative learning. For example, children learning English advance through a stage at about 24 months where they produce nonfinite verbs in simple sentences; for example, *he open it* and *she eat grapes*. These are clearly not in the input, since adult English-speakers use finite verbs in 'root' (main) clauses (as compared to embedded clauses). According to Tomasello,⁸ however, this 'Root Infinitive' phenomenon can be explained by imitative learning, based on the observation that children frequently encounter adult questions like *Should he open it? Does she eat grapes?*, and so on. Children produce so-called Root Infinitives because they omit the initial verbal element from the corresponding adult questions. So, for example, the adult's *Should he open it?* becomes the child's *he open it*.

In evaluating the arguments for conservative learning, it pays to ask whether the input from adults is the likely source of children's Root Infinitives in other languages. Example (1) is from Poeppel and Wexler,¹² who analyzed the spontaneous productions of a two-year-old German-speaking child. As in the English examples cited by Tomasello (*he open it* and *she eat grapes*), the utterances in (1) contain a nonfinite form of the verb.

- (1) a. Thorsten Caesar haben.
Thorsten C(=doll) have-Finite
'Thorsten has Caesar'.
b. Du das haben.
you that have-Finite
'You have that'.

Applying Tomasello's⁸ truncation hypothesis, one could infer that the utterances in (1) result from the truncation of adult utterances. For example, (1a) could be a truncated version of *Will Thorsten Caesar haben?* 'Does Thorsten want to have Caesar?' However, the truncation analysis would predict that children's truncated utterances would also include nonfinite verbs followed by a finite form, such as ... *gehen muß* ('to go must-2.SG') and ... *essen willst* ('to eat want-2.SG'), since these are highly frequent

combinations in the input from adults (e.g., *Mami sagt daß du ins Bett gehen mußt*. 'Mommy says you must go to bed', and *Mami fragt was du essen willst*. 'Mommy asks what you want to eat'). However, the 2-year old German-speaking child in question never produced utterances with a nonfinite verb followed by a finite verb. Instead, Poeppel and Wexler¹² report that finite verbs overwhelmingly occurred in V_2 position in the child's productions, whereas nonfinite verbs overwhelmingly occurred in final position. In short, this child's nonadult forms are better analyzed as nonadult linguistic constructions, rather than as truncated adult constructions. A related problem with the truncation hypothesis is that some child languages do not exhibit root infinitives, for example, Italian¹³ and Spanish.¹⁴ The input to children learning these languages would be expected to lead them to produce nonfinite verbs in simple sentences, contrary to fact.

During the same period that children produce Root Infinitives (i.e., nonfinite verbs in simple sentences), children also produce sentences with non-nominative subject pronouns, that is, with a pronoun in subject position marked with accusative or genitive case. The resulting utterances are deviant for adults, such as *her open it*. According to Tomasello,⁸ children's nonnominative subjects can again be explained by the truncation hypothesis. In this case, children encounter adult utterances like *let her open it*, and (re)produce a truncated version of these utterances, lacking the initial verbal element, *let*.¹⁵ The truncation account by Tomasello⁸ is also invoked to explain the absence of certain 'errors' by children, such as *Mary hit I* (with a direct object pronoun in nominative case):

'The reason they do not make this error is that they never hear adults say anything like this in any linguistic construction.' (p. 240)

Conservative learning has been extended to argument structure. Children have been reported to produce argument structure overgeneralizations in the literature, included utterances with verbs in noncanonical positions. One class of examples is the causative construction, first documented by Bowerman.¹⁶ In this construction, children were found to insert intransitive verbs like *fall*, *giggle*, and *die* into frames that are reserved for transitive verbs for adults. Tomasello⁸ acknowledges the need to explain these instances of children's nonadult behavior 'because, presumably, children have not heard such forms used in adult speech' (p. 215). Having reviewed the literature, Tomasello reaches the following conclusion:

'The main result of interest in the current context is that these children produced very few argument

structure overgeneralizations before about 3 years of age and virtually none before 2.5 years of age (see, Pinker 1989, pp. 17–26).' (p. 215)

Since Tomasello cites Pinker¹⁷ as providing relevant evidence of the dearth of argument structure overgeneralizations, it is worth pointing out that Pinker reports six instances of causative overgeneralizations by children younger than age 2 years and 5 months. These are reproduced in (2). In addition, children younger than 3 years produced 18% of the entire sample cited by Pinker (14/78).

- (2) C, 2 years and 1 month: [M: Close your eyes] No! I want be my eyes open.
 E, 2 years and 2 months: I'm talking my birdie. [Pulling string on bird-shaped music box]
 E, 2 years and 1 month: I wanta swim that. [Holding an object in the air and wiggling it as if it were swimming]
 C, 2 years and 3 months: Bottle feel my feets better. [makes them feel better]
 K, 2 years and 3 months: Kendall fall that toy.
 S, 2 years and 2 months: Tommy fall Stevie truck down.

Other examples of this construction are reported elsewhere in the literature. For example, Bowerman¹⁶ studied two children, Christy and Eva. According to Pinker's review of the findings by Bowerman, 'Christy... began to overgeneralize the causative relation at age 2;0'. (p. 283) and for Eva, '...causativization errors begin at 2;2'. (p. 336). Maratsos et al.¹⁸ also report the productive use of the causative construction for the child they studied, beginning at 2 years and 6 months. Finally, another review of the literature, by Hochberg,¹⁹ concludes that 'from about 2;0, children begin to form novel transitive verbs from intransitives' (p. 317). Only time will tell whether conservative learning, and the truncation hypothesis, will withstand the apparent counterexamples.

Statistical Learning

Recently, developmental psycholinguists have been exploring the possibility that linguistic facts can be learned without the kinds of abstract or implicit principles that have been proposed in the 'nature' approach to language development. One relevant discovery is that children are able to effectively learn certain linguistic properties based on statistical regularities

in the input. For example, Saffran et al.²⁰ showed that 8-month-old children could exploit statistical learning to extract information about transitional probabilities from the input (i.e., how likely one item is to follow another). Infants inferred the existence of word boundaries between three-syllable pseudowords (nonsensical combinations). Those three-syllable sequences that crossed a word boundary were not treated by the child subjects as a ‘word’ during the post-test phase of the study, because there was a lower probability for such sequences to be repeated if they crossed a word boundary than if they were part of a ‘word’. The second development concerns the nature of the input available to children. It has recently been argued that the input contains relevant features in sufficient abundance to support statistically based acquisition of several seemingly complex facts about language.^{3,21}

The conclusion reached by proponents of the usage-based account is that children can extract the relevant generalizations from what adults actually say, in the circumstances in which they say them. However, critics have pointed out limitations in statistical learning mechanisms (e.g., distributional analysis). While the statistical learning mechanism most often cited in the literature is capable of extracting information about the transitional probabilities of three-syllable sequences, Yang²² showed that the same mechanism *cannot* reliably segment sequences of monosyllabic words.^{23,24} This is problematic, because sequences of monosyllabic words make up the majority of the input directed to English-speaking children. The statistical learning mechanism is effective, however, if it is guided by a linguistic constraint—that each word contains a single primary stress.²²

It has also been pointed out that statistical learning mechanisms can learn things that human children cannot learn. For example, a study by Read and Schreiber²⁵ found that 7-year-old children are sensitive to structural notions like subject noun phrase, as long as such phrases contain more than one word. But the Read and Schreiber study also found that 7-year-olds *cannot* learn STRUCTURE-INDEPENDENT rules, like ‘drop the first four words of a sentence’. Similarly, Smith and Tsimpli²⁶ showed that adults are unable to learn structure-independent rules for question formation. To the extent that statistical learning mechanisms *are* able to form structure-independent generalizations, these mechanisms are apparently quite unlike human minds.

THE NATIVIST (NATURE) APPROACH TO LANGUAGE ACQUISITION

The nativist solution to Plato’s Problem supposes that children are biologically fitted, as part of the human

genome, with a UNIVERSAL GRAMMAR (UG).^{27–29} The Universal Grammar account views language acquisition as, at least in part, the by-product of a domain-specific computational mechanism. Universal Grammar contains the core PRINCIPLES of language, that is, principles that are manifested in all human languages. In addition, Universal Grammar spells out particular ways in which human languages can vary. These points of variation are called PARAMETERS. Taken together, the principles and parameters of Universal Grammar establish the boundary conditions on what counts as a possible human language. Children are seen to navigate within these boundaries in the course of language development. The universal principles enable children to rapidly and effortlessly acquire any human language without formal instruction and despite the considerable latitude in the experiences of different children. As noted earlier, according to nativists, children’s linguistic knowledge is vastly underdetermined by their experience. Concrete instances of ways in which children’s linguistic knowledge is underdetermined by their experience are called POVERTY-OF-THE-STIMULUS arguments. Based on a series of such arguments, nativists have concluded that children are innately endowed with certain linguistic knowledge, namely the principles and parameters of Universal Grammar.

Structure Dependence

There are several points about the principles of UG that are often not fully appreciated. The first point is that UG is not a theory of the grammar of particular languages. Here is an instructive quote from Chomsky.²⁷

‘The grammar of a particular language . . . is to be supplemented by a universal grammar that . . . expresses the deep-seated regularities which, being universal, are omitted from the grammar itself. *Therefore it is quite proper for a grammar to discuss only exceptions and irregularities in detail.* [Our italics] It is only when supplemented by a universal grammar that the grammar of a language provides a full account of the speaker-hearer’s competence’. (p. 6)

As Chomsky makes clear, it is usual for grammars of particular languages to concentrate on what makes each language special, and not what it has in common with other languages. By contrast, UG does not attempt to account for ‘exceptions and irregularities’, but rather those aspects that are shared by human languages. Little is gained in arguing against UG by pointing out that individual languages contain lots of irregularities and exceptions. Any challenge to UG requires more than this.

A second point is that the theory of Universal Grammar (UG) is an empirical proposal about the initial state of language learners, not a proposal about the final state of adult speakers of any human language. The principles of UG determine the kinds of analyzes that language learners can adopt. To cite a famous example, Chomsky proposed almost 40 years ago that children can hypothesize structure-dependent operations, but not structure-independent operations. To illustrate, he discussed the formation of Yes/No questions in English. There are several points to the argument. One is that Yes/No questions and declarative statements are somehow related: for any declarative statement, there is a corresponding Yes/No question, as illustrated in (3) and (4).

- (3) Declarative: Australians are friendly.
 (4) Yes/No Question: Are Australians friendly?

The question is about the nature of the linguistic mechanism that establishes this relationship. Chomsky considers two logically possible relationships, indicated in (5) and (6).

- (5) Rule A: Move the first occurrence of *are* to the front.
 (6) Rule B: Move *are* from the main clause to the front.

It turns out that only Rule B is empirically adequate, but that both Rules A and B can account for the kind of simple examples that makes up the preponderance of the input to young children, such as examples (3) and (4). Unless the young child has a predisposition toward Rule B, young children would not be expected to initially favor Rule B over Rule A, since Rule A is far simpler on any standard measure of complexity.

Input that distinguishes Rules A and B is more complex statements, such as (7). In (7), the subject phrase contains a restrictive relative clause (here, ...*who are sunburned*). Since the relative clause contains the verbal element *are*, the structure-independent Rule A would move this occurrence of *are*, resulting in the ill-formed Yes/No question (8), instead of the well-formed Yes/No question (9). The deviance of (8) is indicated by the asterisk '*':

- (7) Australians who are sunburned are friendly.
 (8) *Are Australians who sunburned are friendly?
 (9) Are Australians who are sunburned friendly?

Chomsky's proposal is that Universal Grammar restricts children's grammars to structure-dependent

hypotheses. If so, then children should never adopt the structure-independent hypothesis that would enable them to ask questions like (8). This prediction has been empirically verified by Crain and Nakayama.³⁰ Since the core linguistic principle under discussion dictates what sentences *cannot* be formed, the principle is often referred to as a CONSTRAINT on grammar-formation. Metaphorically, a relative clause is an 'island' from which nothing can escape.

Poverty-of-the-Stimulus Arguments

As mentioned at the beginning, PLATO'S PROBLEM refers to any gap between experience and knowledge. Nativists working within the generative linguistic tradition have championed an argument, called the poverty-of-the-stimulus argument, which reaches the same conclusion about language—that children know more than they could have learned from their experience. The term 'poverty-of-the-stimulus' underscores the claim that children's linguistic experience underdetermines what they come to know about human language. We will review the basic ingredients of the poverty-of-the-stimulus argument. According to Jerry Fodor³¹ (p. 258) the argument from the poverty-of-the-stimulus is '*the existence proof for the possibility of cognitive science*'.

Constraints

One component of the poverty-of-the-stimulus argument is founded on what are called linguistic constraints. The role that linguistic constraints play in the nativist account of language development is to prevent children from forming misguided hypotheses about the forms and meanings of linguistic expressions. One way that a child's hypothesis might be misguided is by being too broad, in the sense that hypothesis generates linguistic *forms* that are not in the language. Another possibility is that the hypothesis extends the *meanings* of expressions beyond those generated by adult speakers of the local language. If children make hypotheses that are broader than those of adults, in either of these ways, then they will make errors that may be difficult for them to recover from, thereby making convergence on the system of adult linguistic knowledge slow and onerous. The observation that children rapidly master many seemingly complex linguistic facts suggests that children are generally able to avoid the pitfalls of mistaken hypotheses.

This is where constraints enter into the equation. Constraints place limits on the kinds of hypotheses children can entertain, so constraints ensure that real-world experience will provide relevant data to confirm a hypothesis, or redirect the child to a new hypothesis. Constraints first assumed prominence in

in linguistic theory in the late 1960s. In the 1970s and 1980s it became standard to invoke constraints in poverty-of-the-stimulus arguments, and this remains true today.^{9,10} Prior to the introduction of constraints into linguistic theory, grammars were taken to be system of rules. Rules are ‘positive’ statements in a grammar, indicating which forms and meanings are possible in a language. In contrast to rules, constraints are negative statements. Constraints dictate which forms and meanings are *not* possible in a language. Armed with constraints, the child is prevented from producing illicit forms and from assigning illicit meanings. We have already seen an example of a constraint on sentence form. This was Chomsky’s example of the structure-dependent nature of the formation of Yes/No questions.

The other kind of constraint prevents children from assigning illicit meanings to well-formed sentences. Before we introduce one such constraint, it will be useful to consider what meanings can be assigned in its absence. Notice that the declarative sentence (10) and the *wh*-question in (11) have two distinct interpretations. On one reading, the pronoun *he* can be anaphorically linked to the name *Obama* in (10) and to the *wh*-word *who* in (11). On this reading, (10) can be paraphrased as *Obama thinks that he, Obama, will buy a Prius*. And example (11) can be paraphrased as *Which x is such that x thinks x will buy a Prius?* There is also a second reading for both (10) and (11). On the second reading of (10), the pronoun *he* is being used ‘deictically’, to refer to some male individual who is not mentioned in the sentence, say Joe Biden. Since Barack Obama is not mentioned in (11), he is a potential referent of the pronoun *he* on its second reading, so this reading can be paraphrased as *Who thinks he (Obama) will buy a Prius?*

(10) Obama thinks he will buy a Prius.

(11) Who thinks he will buy a Prius?

Now let us introduce the constraint. The constraint is called Principle C, and it is operative in examples (12) and (13). Because of Principle C, only one interpretation of these sentences is possible. Notice that the pronoun *he* cannot be related to *Obama* in (12). And in example (13), if the answer is *Obama*, then presumably Barack Obama was not the referent of the pronoun *he*.

(12) He thinks Obama will buy a Prius.

(13) Who does he think will buy a Prius? Obama.

Chomsky proposed that the same linguistic constraint, Principle C, governs both of the examples

in (12) and (13). This constraint prohibits pronouns from being linked to other expressions, when the pronoun is in a particular structural position with respect to the other expression (called c-command).^{29,32}

It has recently been argued that Principle C applies in discourse sequences. Notice that the pronoun *he* used by Speaker A cannot be anaphorically linked to the name *Obama* in the responses by Speaker B in (14). Generative linguists have suggested that Speaker B’s fragment answer involves copying and then deleting some of the linguistic material from Speaker A’s statement, as indicated in (15).³³ The constraint on anaphoric relations operates on the underlying representation, in which the pronoun *he* is present, but not pronounced.

(14) Speaker A: I know who he thinks will buy a Prius.

Speaker B: Me too. Obama.

(15) Speaker B: Me too. ~~He thinks that~~ Obama will buy a Prius.

A number of studies have been conducted to see how young children interpret structures like those illustrated in (10)–(14).^{34–36} The findings are consistent with the expectations of the nativist account of language development. Essentially, as soon as they can be tested (about 3-years of age), children have been found to adhere to several linguistic constraints, including the constraint on when pronouns cannot be semantically related to expressions in the same sequence of words, as in (12) and (13), or in the discourse example in (14). Apparently, children do not need to avail themselves of mechanisms such as entrenchment, pre-emption and indirect negative evidence in order to acquire the information encoded in constraints. The early acquisition of constraints, therefore, provides prima facie evidence for the nativist account, because children evidently have linguistic knowledge for which there is no decisive experience in the input.

Negative Evidence

Nativists invite us to ask whether children could plausibly learn the constraint of anaphoric relations between pronouns and names on the basis of evidence available to them. A critical observation about linguistic constraints is that they are negative statements. To ‘learn’ a constraint, children would need to have access to information about what meanings strings of words *cannot* have. This is called NEGATIVE EVIDENCE. As far as we know, negative evidence is not available in the primary linguistic data, much less a ‘regular’ part of the evidence available to children.^{9,10,37–39}

One possible source of direct negative evidence is corrective feedback. After careful scrutiny of parent-child interactions, however, researchers have concluded that this kind of negative evidence is not available at the right time or in sufficient abundance to promote learning of negative constraints.^{23,40–43} In the absence of negative evidence, it is difficult to see how children could ‘learn’ such constraints. Moreover, even if negative evidence were available, children may not avail themselves of it. There is no evidence that children exposed to negative evidence use it to purge their grammars of incorrect hypotheses.⁴⁴

Because the usage-based approach lacks innate constraints (e.g., Principle C), the usage-based account must postulate alternative mechanisms to unify constructions that nativists claim to be governed by abstract structural properties. In the case of anaphoric relations, van Hoek⁴⁵ offers an account within the framework of Cognitive Grammar, and Levinson⁴⁶ offers a pragmatic account. As acknowledged by Ambridge and Lieven,⁴⁷ the challenge for the usage-based approach is to explain in detail how children move from specific constructions to build up ‘patterns of relationships’ among constructions⁷ (p. 169). Recently, Goldberg² has proposed that linguistic generalizations can be explained by invoking notions of information-structure and discourse function, such as topic, background, and focus (for a critique, see Ref 48). The dispute is far from settled but, as we discuss next, the intricate and unexpected patterns among constructions, both within and across languages, are the bread and butter of the nativist approach to language development.

Cross-linguistic Research

Another component of the poverty-of-the-stimulus argument appeals to the unification of disparate linguistic phenomena across-languages. An example is DOWNWARD ENTAILMENT.^{38,49} At the broadest cut, a linguistic expression is downward entailing if it licenses inferences from general terms to specific terms. More precisely, a linguistic expression is downward entailing if it validates inferences from expressions referring to sets (e.g., *fruit*, *car*) to expressions referring to their subsets (e.g., *apples*, *Prius*). Many linguistic phenomena are governed by downward entailment. One downward entailing linguistic structure is CONDITIONAL STATEMENTS, for example, English *if... then...*, Mandarin *ruguo*-conditionals, but other downward entailing expressions include certain adverbs, propositions, comparative, quantificational expressions, verbs, and more.

It is useful to partition conditional statements into two constituents, (a) the ANTECEDENT CLAUSE

and (b) the CONSEQUENT CLAUSE. The critical observation is that different things happen in each of these clauses. One asymmetry concerns the validity of inferences from sentences that contain terms referring to sets (e.g., *car*) to ones in which the set-referring term is replaced by one that refers to a subset (e.g., *Prius*). Example (16) shows that the antecedent of conditionals meets the definition of downward entailing in English, where ‘ \Rightarrow ’ is used to indicate that the second sentence is logically entailed by the first. Example (17) shows that the consequent of conditionals does not meet the definition of downward entailment in English.

- (16) If a linguist ordered a **car**, he got a rebate.
 \Rightarrow If a linguist ordered a **Prius**, he got a rebate.
- (17) If a linguist got a rebate, he ordered a **car**.
 $*\Rightarrow$ If a linguist got a rebate, he ordered a **Prius**.

This is not just a fact about the interpretation of conditional statements in English. If we translate the English examples into Mandarin, the same asymmetry is manifested. That is, the antecedent clause of a conditional statement is downward entailing, as shown in (18), but the consequent clause of a conditional statement is not downward entailing, as the deviance of the entailment in (19) shows.

- (18) *Ruguo yi-ge yuyanxuejia mai-le qiche, ta jiu na-le huikou.*
 if one-CL linguist buy-ASP car he then get-ASP rebate
 ‘If a linguist bought a car, he got a rebate’.
 \Rightarrow *Ruguo yi-ge yuyanxuejia mai-le Puruisi qiche, ta jiu na-le huikou.*
 if one-CL linguist buy-ASP Prius car he then get-ASP rebate
 ‘If a linguist bought a Prius, he got a rebate’.
- (19) *Ruguo yi-ge yuyanxuejia na-le huikou, ta jiu mai-le qiche.*
 if one-CL linguist get-ASP rebate he then buy-ASP car
 ‘If a linguist got a rebate, he bought a car’.
 $*\Rightarrow$ *Ruguo yi-ge yuyanxuejia na-le huikou, ta jiu mai-le Puruisi qiche.*
 if one-CL linguist get-ASP rebate he then buy-ASP Prius car
 ‘If a linguist got a rebate, he bought a Prius’.

A second linguistic phenomenon governed by downward entailment involves disjunction (English *or*; Mandarin *huozhe*). In the antecedent clause of a conditional statement, disjunction generates a conjunctive entailment. This is illustrated in example (20). Example (20) entails that if Ted ordered pasta, then Max ordered pizza, and if Ted ordered sushi, then Max ordered pizza. By contrast, when disjunction appears in the consequent clause, the interpretation of disjunction is ‘disjunctive’, as illustrated in (21). Example (21) would be true, for example, if Max ordered pizza, and Ted ordered pasta, but not sushi, so disjunction does not generate a conjunctive entailment in the consequent of a conditional.

(20) If Ted ordered pasta or sushi, then Max ordered pizza. = Conjunctive

(21) If Max ordered pizza, then Ted ordered pasta or sushi. = Disjunctive

Mandarin works exactly the same way. In the antecedent clause of a conditional statement, disjunction generates a conjunctive entailment, as illustrated in (22). When disjunction appears in the consequent clause in Mandarin, the interpretation is disjunctive, as shown in (23).

(22) Ruguo Taide dian-le yidalimianshi huozhe shousi, name Maikesi dian-le pisa.

if Ted order-ASP pasta or sushi, then Max order-ASP pizza

‘If Ted ordered pasta or sushi, then Max ordered pizza’. = Conjunctive

(23) Ruguo Maikesi dian-le pisa, name Taide dian-le yidalimianshi huozhe shousi.

if Max order-ASP pizza, then Ted order-ASP pasta or sushi

‘If Max ordered pizza, then Ted ordered pasta or sushi’. = Disjunctive

A third phenomenon is governed by downward entailment. This is the licensing of Negative Polarity Items, such as English *any* and Mandarin *renhe*. A Negative Polarity Item (NPI) is defined as an expression that must be interpreted within the scope of a downward entailing linguistic context. NPIs are licensed in downward entailing linguistic environments. As (24) illustrates, the English NPI *any* can appear in the antecedent of a conditional, but the sentence is anomalous with *any* in the predicate phrase. The corresponding NPI in Mandarin, *renhe*, has the same pattern of distribution, as shown in (25).

(24) a. If John ate **any** ice cream, he became ill.

ANT[If John ate any ice cream] CONS[he became ill]

b. If John became sick, he ate ***any** ice cream.

ANT[If John became ill] CONS[he ate *any ice cream]

(25) a. ANT[Ruguo Yuehan chi-le **renhe** binjiling]
CONS[ta jiu shengbing]

if John eat-ASP any ice-cream he then sick

‘If John ate any ice cream, he became ill’

b. ANT[Ruguo Yuehan shengbing-le] CONS[ta jiu chi ***renhe** binjiling]

if John sick-ASP he then eat any ice-cream

‘If John became ill, he ate *any ice cream’

Turning to acquisition, both English-speaking children and Mandarin-speaking children must somehow figure out that the antecedent clauses of conditionals license Negative Polarity Items (NPIs such as English *any*, Mandarin *renhe*), and that the consequent clauses of conditionals do not. Perhaps the kind of piecemeal acquisition advocated by the usage-based approach could suffice for children to learn the (positive) environments in which negative polarity items can appear, such as (24a) and (25a). It is conceivable, moreover, children could avoid producing NPIs in nondownward entailing linguistic environments, such as (24b) and (25b), simply because they do not encounter them in those environments. Finally, if children are not conservative learners, and overgenerate (i.e., produce NPIs where they are not licensed), they might be able to avail themselves of pre-emption (i.e., noticing that adult use *some*) to expunge their grammars of illicit uses of NPIs.⁹

As noted by Chierchia,³⁸ however, the pressing question for the usage-based account is how children discover when disjunction licenses a conjunctive entailment and when it does not. The different truth-conditions associated with disjunction in downward and nondownward entailing contexts are not distributional facts; they are different interpretations of the same lexical items. It seems highly improbable that children, across the globe, have the requisite cognitive skills to keep detailed records of when disjunction words do and do not generate a conjunctive entailment. It is also highly unlikely that either

English-speaking children or Mandarin-speaking children have decisive evidence from the input informing them that the word for disjunction generates a conjunctive entailment when it appears in the antecedent clause of a conditional but has disjunctive truth-conditions when it appears in the consequent clause. Nevertheless, 3- to 5-year old children correctly distinguish the interpretation of disjunction in these two positions.^{50,51}

ACQUISITION OF COMPLEX STRUCTURES

This completes our review of the two competing accounts of language development. The different predictions of these accounts are probably most pronounced when they attempt to account for children's acquisition of complex syntax. Setting aside the possibility of biological maturation, the nativist account anticipates the early emergence of knowledge of complex syntax in children, both in comprehension and in production. By contrast, the usage-based approach contends that early knowledge is limited, with more abstract constructions being acquired gradually, as children are exposed to an expanded database of input. We survey some recent experimental work that has been cited by researchers from both camps, briefly considering three complex structures: passives, relative clauses, and questions.

Passives

On the nativist approach, passives (e.g., *The rock star is being chased by a fan*) are formed by movement, one of the basic building blocks of the innate language apparatus. According to Ambridge and Lieven,⁴⁷ by contrast, the usage-based approach makes the opposite claim:

'passive utterances are not formed by ... movement. Rather, the passive construction is initially acquired on an item-by-item basis, beginning with simple 'adjectival' passives (e.g., *It's broken*), then gradually accruing lexically specific formulas (e.g., *The X was broken by the Y*), and finally incorporating more abstract constructions (X BE/GET VERB by Y).' (p. 276)

Ambridge and Lieven⁴⁷ conclude from their review of the literature that:

'the two sides ... are essentially in agreement with regard to the data. English-speaking children first acquire adjectival stative passives (e.g., *It's broken*), then full actional passives (e.g., *It got broken by*

the hammer) and finally non-actional passives (e.g., *The boy was seen by the girl*). Furthermore, both sides agree that, given the right experimental circumstances, children aged as young as 3 demonstrate at least some abstract knowledge of the full passive.' (p. 276)

It is worth reviewing some of the relevant literature. Probably the most successful study of the passive to date was an elicited production study reported by Crain et al.⁵² These researchers interviewed 32 English-speaking children (from 3 years and 4 months to 4 years and 9 months), and were able to elicit full actional passives with *by*-phrases from 29 children, including the youngest child (3 years and 4 months). Twenty-four children produced three or more full passives. The majority of children's productions were *get*-passives, as illustrated in (26).

(26) She got knocked down by the Smurfie. (Phoebe 3 years and 4 months)

The car got flied over by that helicopter right there. (Kyle 3 years and 9 months)

It got pushed over by the bus, and the bus is bringing him to the hostibul. (Michael 3 years and 11 months)

Point to the car that is being crashed by the helicopter. (Joey 3 years and 11 months)

For other kinds of passives, however, the evidence is less clear cut.^{53,54} The relevant research suggests that children younger than 5-years-old experience difficulty comprehending nonactional passives that contain both some form of the copula *be* and a *by*-phrase (e.g., *The horse is seen/heard by Donald Duck*). Most 3- and 4-year-old children are generally successful in comprehending truncated versions of this structure (e.g., *The horse is seen/heard*). However, some children have difficulty with specific verb forms (including *seen/heard*), as the usage-based account anticipates.

Relative Clauses

Research has investigated children's acquisition of several two-clause structures, including sentences with relative clauses, and complex questions. A relative clause is a modifier that expresses a proposition about the noun phrase it modifies. Relative clauses may have a missing Subject or Object, and they can attach it to one or other of the lexical noun phrases in the main clause. Consider sentence (27).

(27) The boy fed the elephant that _ squirted the giraffe.

In this sentence, the Object of the main clause, *the elephant*, is modified by the relative clause, ... *that _ squirted the giraffe*. This relative clause has a 'gap' in Subject position (as indicated by the '_'). Consider the slightly different sentence in (28).

- (28) The boy that the elephant squirted _ fed the giraffe.

In this example, the Subject of the main clause, *the boy*, bears the relative clause, ... *that the elephant squirted _ ...*, and the relative clause has a 'gap' in Object position. As these examples illustrate, there are several varieties of relative clause structures. According to the nativist account, all varieties of relative clauses involve basic syntactic operations, including movement (whence the gap). Again, the nativist account anticipates the early emergence of relative clause structures, both in comprehension and in production. By contrast, the usage-based approach anticipates only gradual development of different kinds of relative clause constructions, based on properties of the input.

There is empirical support for both approaches. The nativist approach can take heart in the finding that children even younger than 3-years-old produce sentences with relative clauses. Using an elicited production technique devised by Hamburger and Crain,⁵⁵ Crain et al.⁵⁶ elicited relative clauses from 20 Italian-speaking children ranging in age from 2 years and 8 months to 3 years and 11 months. Nine of the 20 children produced relative clauses on all 12 trials, including two of the youngest children, aged 2 years and 8 months and 2 years and 10 months; every child who was interviewed produced at least six relative clause structures.

There is also consolation for the usage-based account. In many comprehension studies, children did not perform with equal success on all varieties of relative clause sentences. As Ambridge and Lieven⁴⁷ (p. 293) observe, the nativist account anticipates that 'children should be equally adept at understanding or producing all types. In fact, early comprehension studies revealed that children have considerably more difficulty with some types of relative clause than others'. In subsequent studies, where experimenters have taken care to reduce task demands to a minimum, children demonstrate improved performance.^{55,57} However, children younger than 5-years-old do not display adult-like levels of performance on all varieties of relative clause sentences.

On the nativist account, delays in the comprehension of complex syntactic structures are often attributed to the processing limitations experienced

by younger children, over-and-above the limitations experienced by older children and adults. On the usage-based account, the same findings are interpreted as evidence that children are progressing from simpler to more complex structures, based on properties of the input. In fact, these alternatives are not easy to tease apart. The null hypothesis is that the performance systems of children and adults are (qualitatively) the same.^{58,59} If the null hypothesis is correct, then adults are expected to experience difficulty in processing the same kinds of relative clause structures that prove difficult for children. This, in turn, would prevent adults from producing those relative clause structures. Consequently, both the experience-based account and the performance-limitation (nativist) account can explain children's delayed language development.

Questions

Another area of dispute between the usage-based approach and the nativist approach concerns children's productions of *wh*-questions. In *wh*-questions, either a 'bare' *wh*-word (e.g., *who*, *what*) or a 'full' *wh*-phrase (e.g., *what Smurf*, *which one of the babies*) appears at the front, and there is a missing noun phrase (or 'gap') in another position.

Our interest is in complex *wh*-questions. Complex *wh*-questions involve both a main clause and a complement clause. In these questions, the *wh*-word or *wh*-phrase appears at the front of the main clause, and the gap can be in the Subject position or in the Object position of the complement clause (or elsewhere, but we will focus on these two possibilities). Generative linguists have postulated that the *wh*-word has been 'extracted' from the site of the gap and 'moved' from there to the front of the question. Such questions are referred to as 'long distance *wh*-questions' or 'long-distance dependencies'. Example (29) is a long-distance *wh*-question in which the *wh*-word *what* has been moved from the Subject of the complement clause ... *_ is in the box* (where '_' marks the site of extraction). In example (30) the full *wh*-phrase *which turtle* has been extracted from the object position of the complement clause ... *the Troll stood on _*.

- (29) What do you think _ is in the box?

- (30) Which turtle did you say the Troll stood on _ ?

Because of their inherent complexity, generative linguists have pointed to children's early productions of long-distance *wh*-questions, as in (29) and (30), as evidence in support of the nativist account of language development. They have argued, in addition, that supporting evidence is provided by children's nonadult long-distance questions (see (31) and (32) below).

In recent work, advocates of the usage-based approach have taken up the challenge. Here is a quote from Dabrowska et al.⁶⁰:

'A number of researchers have claimed that questions and other constructions with long distance dependencies (LDDs) are acquired relatively early, by age 4 or even earlier, in spite of their complexity. . . . Analysis of LDD questions in the input available to children suggests that they are extremely stereotypical, raising the possibility that children learn lexically specific templates . . . rather than general rules of the kind postulated in traditional linguistic accounts of this construction.' (p. 571)

Citing the findings of an experiment from Thornton and Crain⁶¹ (Experiment 1), Dabrowska et al. introduce some examples of children's nonadult *wh*-questions, and they advance a usage-based account of the source of these questions:

'Interestingly, in their productions of LDD children sometimes produced questions like (31) and (32), with a WH word at the beginning of both the main clause and the subordinate clause.

- (31) What do you think what is in the box?
 (32) What way do you think how he put out the fire?

Thornton and Crain regard such 'medial WH' questions as evidence for the . . . application of movement. . . Note, however, that such utterances could also be produced by simply juxtaposing two independent questions (*what do you think? + what is in the box?*) or an independent question and an indirect question (*what way do you think? + how he put out the fire?*).' (p. 273; examples are renumbered)

As examples (31) and (32) illustrate, the Thornton and Crain⁶¹ study was designed to make a specific comparison of children's long-distance adjunct *wh*-questions (with *how*, *why*), and their argument *wh*-questions (with *what*, *who*). However, the evidential base is considerably more substantial than the findings of this one experiment. A fuller array of findings has been reported and discussed in numerous places.^{9,10,34,35,49,62-67}

The arguments advanced in support of the nativist account are mainly based on the extensive experimental findings reported by Thornton.³⁵ Using an elicited production technique, Thornton was able to coax a variety of *wh*-questions from 2-5-year-old children, including long-distance *wh*-questions from one child at 2 years and 6 months. Among the structures Thornton elicited were both Subject and Object extraction long-distance *wh*-questions. These included both bare *wh*-words like *who* and *what* and ones with

full *wh*-phrases like *what Smurf* and *which one of the babies*. It is worth noting the paucity of long-distance questions in the transcripts of children's spontaneous productions. For example, de Villiers et al.⁶⁸ report that they could find only 16 long-distance questions in the transcripts of Adam (from CHILDES, Brown corpus) over a period of 3 years and 6 months. In Thornton's study, however, children as young as three, and even younger, had no apparent difficulty producing a range of long-distance *wh*-questions, including medial-*wh* questions like (33) and (34).

- (33) What do you think what's in here? (Pieter 2 years and 10 months)
 (34) What do you think what babies drink to grow big? (Matthew 3 years and 3 months)

It is conceivable, as Ambridge and Lieven⁴⁷ (p. 306) remark, that *wh*-questions like (33) could have been formed simply by juxtaposing two independent questions (*What do you think? + What's in here?*). Example (34) is trickier, however, because the underlying question fragment . . . *what babies drink to grow big?* is not an acceptable independent question. The unacceptability of the *wh*-question *What babies drink to grow big?* is due to the absence of *do*-support. With *do*-support, the question is well-formed: *What do babies drink to grow big?* So, there must be another source of the question fragment . . . *what babies drink to grow big*. Notice that this sequence of words is acceptable as the complement of an indirect question (*Do you know what babies drink to grow big?*). This observation led Dabrowska, Rowland and Theakston⁶⁰ to propose that children's medial-*wh* questions like (34) may be created by the juxtaposition of an independent question and an indirect question.

Thornton³⁵ (p. 242) considered the juxtaposition analysis of children's medial-*wh* questions, but rejected it on the grounds that no child ever produced medial-*wh* questions with two occurrences of *do*-support, as in (35). According to the juxtaposition analysis, questions like (35) could be formed by juxtaposing two independent questions, as in (36), so the fact that questions like (35) were never produced argues against the juxtaposition analysis.

- (35) #What do you think what does Cookie Monster like?
 (36) What do you think? What does Cookie Monster like?

This leaves open the possibility that children's medial-*wh* questions with Object extraction, such as

(35), could have been composed using an independent question and an indirect question, as in (36).

- (37) What do you think? (Do you know) what Cookie Monster likes?

It is *ad hoc*, at best, to analyze children's medial-*wh* questions involving extraction from Subject position in one way, and ones involving extraction from Object position in a different way. Moreover, the absence of medial-*wh* questions with two instances of *do*-support is just one argument against the juxtaposition analysis.

A second argument against the juxtaposition analysis is based on the finding that children (also) never produced medial-*wh* questions that contain two full *wh*-phrases (e.g., *which Smurf*), as in (38). As example (39) indicates, two independent questions could be juxtaposed to form (38), but this never happened.

- (38) #Which Smurf do you think which Smurf is wearing roller skates.
 (39) Which Smurf do you think? Which Smurf is wearing roller skates?

Instead of medial-*wh* questions like (38), the 'extra' ingredient in children's questions with full *wh*-phrases (e.g., *which Smurf*) was a bare *wh*-word (e.g., *who*). Some examples with a bare *wh*-word in medial position are presented in (40).

- (40) Which Smurf do you think who has roller skates on? (T.D. 4 years and 9 months)
 Which animal do you think what really says 'woof woof'? (T.D. 4 years and 9 months)
 Which guy do you think who eats cookies? (T.D. 4 years and 9 months)
 Which guy did they guess who ate the green one? (D.W. 3 years and 9 months)

More often than not, there was no extra ingredient in children's long-distance *wh*-questions with full *wh*-phrases, so children produced adult-like long-distance *wh*-questions. It is not clear how this could be explained on the juxtaposition account, but it is predicted by the nativist account, as we will see in the Section 'Medial-*wh* Questions'.

A third argument against the juxtaposition analysis is based on the finding that children (also) never produced medial-*wh* questions with a nonfinite form of the verb (e.g., *to play*) in the complement clause, as in (41). Here the juxtaposition analysis anticipates

that *wh*-questions like (39) should be produced by children, since these can be derived using an independent question and an indirect question, as in (42). However, children consistently produced adult-like *wh*-questions of this kind, without even a bare *wh*-word in medial position.

- (41) #Who do you want who to play with? Cf. Who do you want to play with?
 (42) Who do you want? (Do you know) who to play with?

To summarize, three kinds of medial-*wh* questions that would be expected on the juxtaposition analysis never occurred. In view of this, nativists have proposed that children's medial-*wh* questions are formed using the basic building blocks of Universal Grammar, as attested in other human languages, but not in English. We will discuss this possibility in more detail in the following section.

CHILD LANGUAGE VERSUS ADULT LANGUAGE

Further evidence has been offered to adjudicate between the competing accounts of children's language development. The last two sections of the article are devoted to children's nonadult linguistic behavior. First, we delve further into children's medial-*wh* questions. Then we discuss some recent cross-linguistic experimental findings suggesting that children acquiring English initially speak (a fragment of) Mandarin Chinese, and that children acquiring Mandarin Chinese initially speak (a fragment of) English.

According to the usage-based account, the child's hypotheses about the local language are driven by what the child hears. By the same token, children's linguistic errors are expected to arise from less articulated grammars than those of adults in the same linguistic community. The child would not be expected to produce utterances that do not reflect the target language, but are well-formed utterances in some other human languages. As Lieven and Tomasello⁷ remark:

'The difference between young children's inventories and those of adults is one of degree: many more, initially all, of children's constructions are either lexically-specific or contain relatively low-scope slots. As well as being less schematic than many adult constructions they are also simpler with fewer parts. And, finally, children's constructions exist in a less dense network—they are more "island-like".' (p. 171)

According to the nativist perspective, children should be expected to sometimes follow

developmental paths to the adult grammar that would be surprising from a usage-based perspective. From the nativist perspective, by contrast, children are free to try out various linguistic options (as long as these options are compatible with Universal Grammar) before children set parameters so as to conform to the particular human language that is spoken to them, e.g., Mandarin Chinese or English. This line of thought is called the CONTINUITY HYPOTHESIS.^{9,39,58} According to the Continuity Hypothesis, child language can differ from the local adult language only in ways that adult languages can differ from each other. Language acquisition is viewed as a process of language change, where at any given time children are speaking a possible human language, just not the language that is being spoken around them.⁶³ However, every normal child is expected to rapidly converge on a linguistic system that is equivalent to that of adults in the local community, despite the considerable latitude in children's experience, intelligence, verbal memory capacity, attention spans, and so forth. However, children may pass through stages that make it look as if they are speaking a 'foreign' language for a while. In fact, the most interesting differences between child language and the local adult language arise when children appear to be speaking a 'foreign' language. If they occur, these differences between child and adult language would be circumstantial evidence that children are effectively ignoring the input. We will review two cases where children seem to adopt linguistic structures that are not attested in the local language, as anticipated by the Continuity Hypothesis.

Medial-*wh* Questions

In some languages, a *wh*-word is repeated in long-distance *wh*-questions. We refer to these as medial-*wh* questions. Medial *wh*-questions are attested, for example, in Romani and in dialects of German,⁶⁹ and related structures appear in Irish, and in Chamorro.⁷⁰ An example of a German medial *wh*-question is (43).

- (43) Wer_i glaubst du wer_i nach Hause geht?
 who-NOM think-2.SG you who towards
 house go-3.SG
 'Who do you think goes home?'

Recall that some English-speaking children produced medial-*wh* questions, but children never repeated a full *wh*-phrase in medial position. Interestingly, full *wh*-phrases cannot be repeated in adult languages that license medial-*wh* questions. Example (44) illustrates.

- (44) #Wessen Buch_i glaubst du wessen Buch_i Hans liest?
 who-GEN book think-2.SG you who-GEN
 book Hans read-3.SG
 'Whose book do you think whose book Hans is reading?'

In this respect, English-speaking children's medial-*wh* questions resemble those produced by adult speakers of medial-*wh* languages. The resemblance does not end there. There is a cross-linguistic prohibition against medial-*wh* questions involving complements with nonfinite verbs. And, as we saw, English-speaking children never repeated a *wh*-word in a complement clause that contained a nonfinite verb (see the deviant example (41) in the previous section).

On the basis of these striking similarities between the medial-*wh* questions produced by English-speaking children, and those produced by adult speakers of medial-*wh* languages, Thornton suggested that English-speaking children's medial-*wh* questions reveal the natural seams of human languages, whereby children have access to linguistic structures that do not characterize the local language, but do characterize other languages spoken around the globe. It should be noted also that the findings from the Thornton³⁵ experiments argue for the early emergence of complex syntax, and against the 'possibility that children learn lexically specific templates'. It is clear, at least, that children do not utilize the kinds of templates that are invoked by the juxtaposition analysis advanced by the usage-based approach. If it turns out, as claimed by Dabrowska et al.⁶⁰ (p. 571), that 'LDD questions in the input available to children ... are extremely stereotypical', then the long-distance medial-*wh* questions produced by children as young as 2-years-old are all the more remarkable.

Scope Parameters

The notion of scope is the same in human languages as in logic. At Logan Airport in Boston we spotted a sign saying *Every airplane does not carry pets*. This sentence is unquestionably ambiguous. On one reading, the universal quantifier *every* takes scope over *not*. A paraphrase of this 'strong' reading is *None of the airplanes carry pets*, meaning that every airplane is such that it does not carry pets. On a second reading *not* takes scope over *every*. A paraphrase of this 'weak' reading is *Not every airplane carries pets*. The ambiguity turns on the scope relations between the logical expressions *every* and *not*. If *every* takes scope over *not*, this yields the strong reading 'every not'. If *not* takes scope over *every*, this yields the weak reading 'not every'.

Whenever sentences have two logical operators, there is a potential scope ambiguity. However, human languages can differ in the scope relations they prefer. This section discusses the human language counterparts to the logical operators for disjunction ‘ \vee ’ and conjunction ‘ $\&$ ’, and how these operators are interpreted when they appear in sentences with another logical operator, negation ‘ \sim ’.

Suppose an English-speaking friend informs you *Ted did not order pasta or sushi*. Adult speakers of English understand this statement to entail two things, that Ted did not order pasta ($\sim P$), and that Ted did not order sushi ($\sim S$). This ‘conjunctive’ interpretation of disjunction in negative statements conforms to one of de Morgan’s laws of propositional logic. In logic, the formula for a negated disjunction is $\sim (P \vee S)$, where ‘ \vee ’ represents disjunction and ‘ \sim ’ represents negation. According to one of de Morgan’s laws, the meaning of the disjunction operator ‘ \vee ’ is inclusive-*or*. When the disjunction operator ‘ \vee ’ appears in the scope of negation, as in the formula $\sim (P \vee S)$, the formula is true if and only if the ‘conjunctive’ formula, $(\sim P \& \sim S)$, is also true. So $\sim (P \vee S)$ entails $(\sim P \& \sim S)$.

This one of de Morgan’s laws holds in human languages, including English. Consider the negated disjunction in (45). This statement generates a conjunctive entailment, just as negated disjunctions do in classical logic.

- (45) Ted didn’t order sushi or pasta. \Rightarrow Ted didn’t order pasta and Ted didn’t order sushi

Not all languages work like English, however. In Mandarin Chinese, the disjunction word is *huozhe*. When *huozhe* appears in the scope of negation (*meiyou*) as in example (45), there is no conjunctive entailment. Mandarin-speaking adults accept (46) in circumstances where Ted ordered pasta (P) but not sushi ($\sim S$), or where he ordered sushi (S) but not pasta ($\sim P$). So example (45) can be associated with the logical formula $(\sim P \vee \sim S)$ rather than the ‘conjunctive’ formula $(\sim P \& \sim S)$.

- (46) (Wo cai) Ted **meiyou** dian yidalimianshi **huozhe** shoushi.
(I guess) Ted not order pasta or sushi
‘It’s either pasta or sushi that Ted did not order’

Does this mean that the Mandarin word for disjunction, *huozhe*, has a different meaning than English *or* does? Not at all. The difference between Mandarin and English is in the scope relations between negation

and disjunction. In Mandarin, the disjunction word *huozhe* takes scope over negation, as in (46), whereas English *or* is interpreted in the scope of negation, as in (45). The meaning of disjunction remains fixed, as inclusive-*or*, across human languages.⁷¹

Expressions that must take scope over negation are called Positive Polarity Items. The Mandarin disjunction word *huozhe* is a Positive Polarity Item (PPI), but the English disjunction word *or* is not. This difference has been cast as a parameter, the Disjunction Parameter. Mandarin adopts one value of the Disjunction Parameter, because Mandarin disjunction is a Positive Polarity Item (PPI). English adopts the other value of the parameter, because English disjunction is not a PPI. In other words, Mandarin adopts the ‘plus’ value of the Disjunction Parameter ($OR = +PPI$), and English adopts the ‘minus’ value ($OR = -PPI$). In languages that take the $OR = -PPI$ ‘minus’ value of the parameter, such as English, negated disjunctions generate conjunctive entailments. In languages that take the $OR = +PPI$ ‘plus’ value, such as in Mandarin, negated disjunctions generate ‘disjunctive’ truth conditions.

We proceed to conjunction (English *and*, Mandarin *he*). In Mandarin, the conjunction word *he* is a Positive Polarity Item (PPI). It must take scope over negation ($AND = +PPI$). So, the Mandarin example (47) is interpreted as meaning ‘both not’ (*neither*).

- (47) Taide **meiyou** dian yidalimianshi **he** shoushi.
Ted not order pasta and sushi
‘As for both pasta and sushi, Ted did not order them’.

In English, by contrast, the conjunction word *and* is interpreted inside the scope of negation ($AND = -PPI$). English generates a ‘not both’ reading of negated conjunctions, as in (48).

- (48) Ted didn’t order (both) pasta and sushi.

This difference between languages has been cast as the Conjunction Parameter. In languages that adopt the $AND = +PPI$ value of the Conjunction Parameter, such as Mandarin, negated conjunctions generate the ‘both not’ (*neither*) reading. In languages that adopt the $-PPI$ value of the parameter, such as English, negated conjunctions generate the ‘not both’ reading.

These cross-linguistic differences in scope relations have led to specific predictions about the course of language development, according to the nativist account. Adopting the Continuity Hypothesis, nativists predict that children initially favor the

scope interpretation that makes sentences true in the narrowest range of circumstances.⁷² In the absence of negative evidence, children adopt the ‘subset’ value of any parameter. This ensures that children encounter positive evidence in languages where adult speakers adopt the ‘superset’ value. This proposal about children’s initial (default) parameter settings is called the **SUBSET PRINCIPLE**.⁷³

The subset value of the Disjunction Parameter is $OR = -PPI$, so English adopts the subset value of the parameter, and generates the ‘neither’ interpretation, whereas Mandarin adopts the superset value and generates the ‘not both’ interpretation. Therefore, children acquiring Mandarin are expected to initially interpret (46) to have the same meaning as the English example (45).

The subset value for the Conjunction Parameter is $AND = +PPI$, so Mandarin adopts the subset value of the parameter, and generates the ‘both not’ (i.e., *neither*) interpretation, whereas English adopts the superset value and generates the ‘not both’ interpretation. According to the Subset Principle, children acquiring English are therefore predicted to initially interpret (48) as having the same meaning as the Mandarin example (47). Exactly the opposite pattern was expected for the Disjunction Parameter.

A Truth Value Judgment task was used to investigate how children acquiring Mandarin and children acquiring English interpret both negated conjunctions and negated disjunctions, following the protocols of Goro and Akiba.⁷⁴ The findings were exactly as anticipated by the Continuity Hypothesis. In response to negated disjunctions, 20 Mandarin-speaking children (mean 4 years and 5 months) assigned the $(OR = -PPI)$ parameter value, leading them to reject statements like (46) 97% of the time in the same contexts in which English-speaking children and adults rejected (45). In contrast to children, Mandarin-speaking adults accepted negated disjunctions like (45) 95% of the time in these same contexts. The 21 English-speaking children (mean 4 years and 9 months) assigned the $(AND = +PPI)$ parameter value of the Conjunction Parameter, leading them to consistently reject statements like (48) 98% of the time in the same contexts in which child and adult Mandarin speakers rejected (47). In contrast to children, English-speaking adults accepted negated conjunctions like (48) 88% of the time in these same contexts.

In short, the Continuity Hypothesis, along with the Subset Principle, were upheld in these experiments. Across languages, children appear to initially

favor parameter values that generate scope relations that make sentences true in the narrowest range of circumstances. This ensures that children will have access to positive evidence if the local language favors the alternative scope possibilities, those that make sentences true in a broader range of circumstances.

These different interpretations assigned by children and adults are difficult to reconcile on the experience-based account of language development. For adults, the Mandarin disjunction word *huozhe* is assigned a ‘not both’ interpretation in simple negative sentences; that is, adults adopt the $(OR = +PPI)$ value of the parameter. Despite this, it was found that Mandarin-speaking children initially adopted the $(OR = -PPI)$ value of the Disjunction Parameter, and assigned the ‘neither’ interpretation according to which negation takes scope over disjunction, just as it does in English. Clearly, Mandarin-speaking children were not basing their interpretation on the adult input. In understanding negated disjunctions, Mandarin-speaking children resemble English (child and adult) speakers, rather than adult speakers of Mandarin.

Exactly the reverse pattern was found for negated conjunctions with English-speaking children. For adults, the English conjunction word *and* receives a ‘not both’ interpretation in simple negative sentences. This means that adults adopt the $(AND = -PPI)$ value of the Conjunction Parameter. However, the finding was that children acquiring English initially adopted the $(AND = +PPI)$ value of the parameter. In contrast to adults, children assigned the ‘both not’ interpretation to negated conjunctions. Clearly, English-speaking children were not basing their interpretation on the adult input. In response to negated conjunctions, English-speaking children resemble Mandarin (child and adult) speakers, rather than adult speakers of English.

CONCLUSION

The basic difference between Universal Grammar (UG) and the usage-based account is this. According to UG, not everything needs to be learned, and not everything can be learned. According to the usage-based account, everything needs to be learned and everything can be learned. This concludes our brief review of the main debate in language acquisition: whether the development of language is better explained by a usage-based account, or by an account that attributes innate linguistic knowledge to children.

REFERENCES

1. Goldberg A. Constructions: a new theoretical approach to language. *Trends Cogn Sci* 2003, 7:219–224.
2. Goldberg A. *Constructions at Work: The Nature of Generalizations in Language*. New York: Oxford University Press; 2006.
3. Pullum G, Scholtz B. Empirical assessment of the stimulus poverty argument. *Linguist Rev* 2002, 19:9–50.
4. Cowie F. *What's Within: Nativism Reconsidered*. New York: Oxford University Press; 1999.
5. Tomasello M. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press; 2003.
6. Cameron-Faulkner T, Lieven E, Theakston A. What part of *no do* children not understand? A usage-based account of multiword negation. *J Child Lang* 2007, 33:251–282.
7. Lieven E, Tomasello M. Children's first language acquisition from a usage-based perspective. In: Robinson P, Ellis NC, eds. *Handbook of Cognitive Linguistics and Second Language Acquisition*. New York: Routledge; 2008.
8. Tomasello M. Do young children have adult syntactic competence? *Cognition* 2000, 74:209–253.
9. Crain S, Pietroski P. Nature, nurture and universal grammar. *Linguist Phil* 2001, 24:139–185.
10. Crain S, Pietroski P. Why language acquisition is a snap. *Linguist Rev* 2002, 19:163–183.
11. Tomasello M. First steps toward a usage-based theory of language acquisition. *Cogn Linguist* 2000, 11:61–82.
12. Poeppel D, Wexler K. The full competence hypothesis of clause structure in early German. *Language* 1993, 69:1–33.
13. Guasti M-T. Verb syntax in Italian child grammar: finite and non-finite verbs. *Lang Acquis* 1994, 3:1–40.
14. Grinstead J. Consequences of the maturation of number morphology in Spanish and Catalan, M.A. Thesis, UCLA, 1994.
15. Kirjavainen M, Theakston A, Lieven E. Can input explain children's *me-for-I* errors? *J Child Lang* 2009, 36:1091–1114.
16. Bowerman MF. Learning the structure of causative verbs: a study in the relationship of cognitive, semantic and syntactic development. *Papers Rep Child Lang Develop* 1974, 8:142–178.
17. Pinker S. *Learnability and Cognition: The Acquisition of Verb-Argument Structure*. Cambridge, MA: Harvard University Press; 1989.
18. Maratsos M, Gudeman R, Gerard-Ngo P, DeHart G. A study in novel word learning: the productivity of the causative. In: MacWhinney B, eds. *Mechanisms of Language Acquisition*. Hillsdale, NJ: Erlbaum; 1987.
19. Hochberb J. Children's judgements of transitivity errors. *J Child Lang* 1986, 13:317–334.
20. Saffran J, Aslin R, Newport E. Statistical learning by 8-month-old infants. *Science* 1996, 274:19–26.
21. MacWhinney B. A multiple process solution to the logical problem of child language. *J Child Lang* 2004, 31:883–914.
22. Yang CD. Universal grammar, statistics or both. *Trends Cogn Sci* 2004, 8:451–456.
23. Marcus GF. Language acquisition in the absence of explicit negative evidence: can simple recurrent networks obviate the need for domain-specific learning devices? *Cognition* 1999, 73:293–296.
24. Marcus GF, Vijayan S, Bandi Rao S, Vishton PM. Rule-learning in seven-month-old infants. *Science* 1999, 283:77–80.
25. Read C, Schreiber P. Why short subjects are harder to find than long ones. In: Wanner E, Gleitman L, eds. *Language Acquisition: The State of the Art*. Cambridge: Cambridge University Press; 1982.
26. Smith N, Tsimpli I-M. *The Mind of a Savant: Language, Learning, and Modularity*. Oxford: Blackwell Publishers; 1995.
27. Chomsky N. *Aspects of the Theory of Syntax*. Cambridge, MA: The MIT Press; 1965.
28. Chomsky N. *Reflections on Language*. New York: Pantheon Books; 1975.
29. Chomsky N. *Knowledge of Language: Its Nature, Origin and Use*. New York: Praeger; 1986.
30. Crain S, Nakayama M. Structure dependence in grammar formation. *Language* 1987, 63:522–543.
31. Fodor J. *Representations—Philosophical Essays on the Foundations of Cognitive Science*. Cambridge, MA: The MIT Press; 1981.
32. Chomsky N. *Lectures on Government and Binding*. Dordrecht: Foris; 1981.
33. Merchant J. Fragments and ellipsis. *Linguist Phil* 2004, 27:661–738.
34. Crain S, Thornton R. *Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics*. Cambridge, MA: The MIT Press; 1998.
35. Thornton R. Adventures in long-distance moving: the acquisition of complex Wh-questions. Doctoral dissertation, University of Connecticut, Storrs, 1990.
36. Conroy S, Thornton R. Children's knowledge of Principle C in discourse. In: Otsu Y, ed. *Proceedings of the Sixth Tokyo Conference on Psycholinguistics*. Tokyo: Hitzu Syobo Publishing Company; 2005.
37. Pinker S. Language acquisition. In: Osherson DN, Lasnik H, eds. *An Invitation to Cognitive Science*, Vol. 1. Cambridge, MA: The MIT Press; 1990, 107–133.

38. Chierchia G. Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In: Belletti A, Rizzi L, eds. *Structures and Beyond*. Oxford: Oxford University Press; 2004.
39. Crain S. Language acquisition in the absence of experience. *Behav Brain Sci* 1991, 4:597–650.
40. Bowerman MF. The “no negative evidence” problem: how do children avoid constructing overly general grammar? In: Hawkins JA, ed. *Explaining Language Universals*. Oxford: Blackwell; 1998.
41. Brown R, Hanlon C. Derivational complexity and order of acquisition in child speech. In: Hayes J, ed. *Cognition and the Development of Language*. New York: John Wiley & Sons; 1970.
42. Marcus GF. Negative evidence in language acquisition. *Cognition* 1993, 46:53–85.
43. Morgan JL, Travis LL. Limits on negative information in language input. *J Child Lang* 1989, 16:531–552.
44. Newport EL, Gleitman H, Gleitman LR. Mother, I'd rather do it myself: some effects and non-effects of maternal speech style. In: Snow C, Ferguson C, eds. *Talking to Children: Language Input and Acquisition*. Cambridge: Cambridge University Press; 1977.
45. van Hoek K. 1997. *Anaphora and Conceptual Structure*. Chicago: University of Chicago Press.
46. Levinson S. *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge, MA: MIT Press; 2000.
47. Ambridge B, Lieven EVM. *Child Language Acquisition: Contrasting Theoretical Approaches*. Cambridge: Cambridge University Press; 2011.
48. Bybee J. Review of Goldberg (2006). *J Child Lang* 2007, 34:692–697.
49. Crain S, Gualmini A, Pietroski P. Brass tacks in linguistic theory: innate grammatical principles. In: Carruthers P, Laurence S, Stich S, eds. *The Innate Mind, Vol. 1: Structure and Contents*. New York: Oxford University Press; 2005.
50. Gualmini A, Meroni L, Crain S. *An Asymmetric Universal in Child Language. Proceedings of Sinn und Bedeutung VI*. Germany: Arbeitspapiere des Fachbereichs Sprachwissenschaft, Konstanz; 2003, 136–148.
51. Meroni L, Gualmini A, Crain S. Everybody knows. In: van Geenhoven V, ed. *Semantics in Acquisition*. Dordrecht: Springer; 89–114.
52. Crain S, Thornton R, Murasugi K. Capturing the evasive passive. *Lang Acquis* 2009, 16:123–133.
53. Fox D, Grodzinsky Y. Children's passive: A view from the by-phrase. *Ling Inq* 1998, 29:311–332.
54. Fox D, Grodzinsky Y, Crain S. An experimental study of children's passive. *MIT Working Papers in Linguistics 26: Papers on Language Processing and Acquisition*. Cambridge, MA: Department of Linguistics and Philosophy, MIT; 1995, 249–264.
55. Hamburger H, Crain S. Relative acquisition. In: Kuczaj S, ed. *Language Development*, Vol. II. Hillsdale, N.J.: Lawrence Erlbaum; 1982.
56. Crain S, McKee C, Emiliani M. Visiting relatives in Italy. In: Frazier L, de Villiers J, eds. *Language Processing and Language Acquisition*. New York: Kluwer Academic; 1990.
57. Corrêa LMS. An alternative assessment of children's comprehension of relative clauses. *J Psycholinguist Res* 1995, 24:183–203.
58. Pinker S. *Language Learnability and Language Development*. Cambridge, MA: The MIT Press; 1984.
59. Crain S, Wexler K. Methodology in the study of language acquisition: a modular approach. In: Ritchie W, Bhatia T, eds. *Handbook of Child Language Acquisition*. San Diego, CA: Academic Press; 1999.
60. Dabrowska E, Rowland C, Theakston A. The acquisition of questions with long-distance dependencies. *Cogn Linguist* 2009, 20:571–597.
61. Thornton R, Crain S. Successful cyclic movement. In: Hoekstra T, Schwartz BD, eds. *Language Acquisition Studies in Generative Grammar*. Amsterdam: John Benjamins; 1994, 215–252.
62. Crain S. The continuity assumption. In: Lasser I, ed. *The Process of Language Acquisition*. New York: Peter Lang; 2002.
63. Crain S, Goro T, Thornton R. Language acquisition is language change. *J Psycholinguist Res* 2005, 35:31–49.
64. Meroni L, Gualmini A, Crain S. Three years of continuous acquisition. In: Otsu Y, ed. *Proceedings of the Second Tokyo Conference on Psycholinguistics*. Tokyo: Hitzu Syobo Publishing Company; 2001.
65. Thornton R. Referentiality and *wh*-movement in child English: Juvenile *D-Link*ency. *Lang Acquis* 1995, 4:139–175.
66. Thornton R. Elicited production. In: McDaniel D, McKee C, Cairns HS, eds. *Methods for Assessing Children's Syntax*. Cambridge, MA: MIT Press; 1996.
67. Thornton R, Crain S, Tesan G. Principles, parameters and probability. In: Gulzow I, Gagarina N, eds. *Frequency Effects in Language Acquisition (SOLA Series)*. New York: de Gruyter; 2007.
68. de Villiers JG, Roeper T, Vainikka A. The acquisition of long distance rules. In: Frazier L, de Villiers JG, eds. *Language Processing and Acquisition*. Dordrecht: Kluwer Academic Publishers; 1990.
69. McDaniel D. Conditions on Wh-Chains. Doctoral dissertation, City University of New York, New York, 1986.
70. Chung S. *Wh*-agreement and referentiality in Chamorro. *Linguis Inq* 1994, 25:1–45.
71. Crain S, Khlentzos D. Is logic innate? *Biolinguistics* 2008, 2:24–56.

72. Goro T. Language-specific constraints on scope interpretation in first language acquisition. PhD dissertation. University of Maryland, 2007.
73. Berwick R. *The Acquisition of Syntactic Knowledge*. Cambridge, MA: MIT Press; 1985.
74. Goro T, Akiba S. The acquisition of disjunction and positive polarity in Japanese. In: Chand V, Kelleher A, Rodríguez AJ, Schmeiser B, eds. *Proceedings of the 23rd West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Press; 2004.

FURTHER READING

- Crain S. The interpretation of disjunction in universal grammar. *Lang Speech* 2008, 51:151–169.
- Crain S, Thornton R. Acquisition of syntax and semantics. In: Traxler M, Gernsbacher M, eds. *Handbook of Psycholinguistics*. San Diego, CA: Elsevier; 2005.
- Crain S, Ni W, Conway L. Learning, parsing, and modularity. In: Clifton C Jr, Frazier L, Rayner K, eds. *Perspectives on Sentence Processing*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1994.
- Crain S, Thornton R, Khlentzos D. The case of the missing generalizations. *Cogn Linguist* 2008, 20:145–155.
- Thornton R, Crain S. Levels of representation in child language. *Linguist Rev* 1999, 16:81–123. Special Issue: Processes and Representations in Grammatical Theory.