Vignette:
Minimalism for language acquisition

For The Cambridge Handbook of Minimalism

Lisa S. Pearl
University of California, Irvine
lpearl@uci.edu

December 19, 2023

Abstract

The vignette explores how Minimalism could help us better understand children's success with language acquisition, where children reliably arrive at the right language generalizations despite massive ambiguity and noise in their input. Because Minimalism is a shift in how we think language is represented in the mind, it also shifts how we think about language acquisition – in particular, what generalizations children are meant to arrive at and what their hypothesis space of possible generalizations looks like. Here, I first review some key problems children need to overcome during language acquisition and then discuss how Minimalism might help us better understand how children solve these key problems. I conclude with some thoughts on how Minimalism's goal of reducing the amount of innate, linguistic knowledge could also help us better understand language acquisition.

Keywords: ambiguity, developing linguistic processing, developing linguistic representations, domain-general learning, induction problems, language acquisition, language-specific knowledge, "Less is More" hypothesis, linguistic representations, Minimalism, noisy input, poverty of the stimulus, statistical learning, Universal Grammar, *wh*-fronting

1 Can Minimalism help us better understand how children solve language acquisition?

1.1 Solving language acquisition

This vignette focuses on whether Minimalism can help us better understand how children "solve" language acquisition. What exactly does it mean to "solve" language acquisition and why should

we care? I believe we should care because achieving native-level proficiency in one or more languages – that is, solving language acquisition – is a shockingly difficult problem. (Just ask anyone who's started learning another language as an adult.) Yet children, who seem to have more limited cognitive abilities than adults, somehow manage to achieve native-level proficiency; in contrast, adults typically fail no matter how long they try. How children succeed (especially when adults fail) is a continual source of both wonder and inquiry in the scientific study of language.

More specifically, language acquisition seems to be about solving a series of induction problems, where children need to identify the correct generalizations for their native language(s) from data that seem woefully inadequate (e.g., see Chomsky, 1981, Jackendoff, 1994, Laurence and Margolis, 2001, Crain and Pietroski, 2002). Let's consider one example generalization from English: when there's a single *wh*-word like *what* in an utterance, the default position for that *wh*-word is at the front of its clause (e.g., *What did this penguin do?* and *I saw [what this penguin did]*). One way to think about this generalization is that *wh*-fronting is required in all clauses. This generalization contrasts with the following potential generalizations, which are just some of the ones an English-learning child might consider:

- optionally allow *wh*-fronting everywhere
- require *wh*-fronting in main clauses but optionally allow it in embedded clauses
- optionally allow *wh*-fronting in main clauses but require it in embedded clauses
- require *wh*-fronting in main clauses and singly-embedded clauses, but optionally allow it in multiply-embedded clauses
- require *wh*-fronting in main clauses, optionally allow it in singly-embedded clauses, and require it in multiply-embedded clauses

Again, these are only some of the possible generalizations, which notably involve some concept of clausal structure. There are many, many other possible generalizations (many of which don't involve this concept), and part of the language acquisition problem is knowing which generalizations are worth considering. Importantly, the data children encounter are compatible with many generalizations (this has been called the *Poverty of the Stimulus* – see Pearl (2022) for a recent overview). That is, children's input is massively ambiguous. Yet, children's task is to somehow pick the correct generalization every single time, despite this massive ambiguity.

Moreover, this massive ambiguity isn't all that children have to deal with – there's also potentially misleading input, which can make the input signal "noisy". Let's consider our *wh*-fronting example again: English-learning children's input contains examples that seem to go against *wh*-fronting, such as echo questions (e.g., *This penguin did what??*) and teaching scenarios (e.g., *Now, do you see the penguin? The penguin did something. Tell me: The penguin did what?*). That is, these examples appear to be counterexamples to any generalization that requires *wh*-fronting, rather than optionally allowing it. This *wh*-fronting scenario highlights just one way that children's input can be messy. Importantly, despite this input messiness, children somehow manage to succeed at acquiring all the appropriate generalizations for their native language(s) and thereby achieve native-level proficiency.

1.2 What Minimalism could do

Minimalism is a significant shift in how we think about language knowledge. For language acquisition, this means that the linguistic representations children need to learn may be quite different than what we previously thought. If the linguistic representations differ, then the generalizations children need to make on the basis of their input also differ.

For our *wh*-fronting example, previous approaches viewed *wh*-fronting as a general property of the language (e.g., see Yang (2012), Sakas and Fodor (2012) and Sakas, Yang and Berwick (2017) for some developmental approaches in the generative tradition). In contrast, the Minimalist approach views *wh*-fronting as the result of features on individual lexical items (e.g., the *wh*-words, the heads like *do* in *What did this penguin do?* that indicate where the *wh*-word is understood), and these features need to be checked. This difference could potentially affect the difficulty of the language acquisition task.

Also, if the linguistic representations differ, then the hypothesis space of competing representations that children consider during language acquisition may also differ. For our *wh*-fronting example, previous approaches may winnow down the hypothesis space significantly; for instance, a macro-parameter approach with a single *wh*-fronting parameter for all clauses may have children only consider *wh*-fronting being required everywhere vs. being optional everywhere (e.g., Yang 2012, Sakas and Fodor 2012, Sakas et al. 2017). In contrast, a Minimalist approach based on features of individual lexical items may retain the variety of hypotheses noted before, depending on where the specific lexical features occur. (Of course, the exact hypothesis space will depend on the exact Minimalist representation being assumed – more on this in Part 2 below.) As with the representations, this difference in hypothesis spaces could also potentially affect the difficulty of the language acquisition task.

Here, I want to walk through some specific ways that the shift to Minimalist representations could impact our understanding of how children solve language acquisition. I'll first review some key problems that children need to overcome during language acquisition, relating to children's input and their ability to harness the information in their input. I'll then discuss why Minimalism might help us better understand how children solve these key problems. I'll conclude with some thoughts on how the spirit of the Minimalist approach, which is about reducing the linguistic knowledge innately built into children's minds, could also be helpful for understanding language acquisition.

2 Part 1: Children have to overcome certain problems

2.1 Problem 1: The input signal is messy

Recall that children's input can be messy for at least two reasons. First, it's massively ambiguous, because of all the potential hypotheses that children could be entertaining about the generalizations in their languages. Second, it can be noisy by containing data that are incompatible with the correct generalizations. Let's delve into each of these more concretely, using the *wh*-fronting example from before.

Ambiguity. Ambiguity in acquisition simply means that the data are compatible with more than one hypothesis under consideration. For *wh*-fronting, I mentioned several potential hypotheses that a child could be considering, involving optional vs. required *wh*-fronting in different structural environments (e.g., main clauses, singly-embedded clauses, doubly-embedded clauses).[1]

If we look at reasonable samples of American English children's input, we can see the actual data they have available to decide among these hypotheses. Since we're focusing first on ambiguity (rather than noise), let's suppose that when we look, we find that children encounter a pristine input landscape with no misleading examples (i.e., no noise). For example, from the CHILDES Treebank (Pearl and Sprouse, 2013), which has about 201,000 utterances of speech directed at children between the ages of six months to five years old, we find approximately 40,700 instances of *wh*-words, with approximately 81% of those in main clauses (as direct questions), approximately 18% of those in singly-embedded clauses, and approximately 1% in multiply-embedded clauses. Under our pristine input assumption, every single one of these instances would have the *wh*-word fronted in its clause, including instances such as ***What*** *are you doing?*, *Do you want to see* ***what*** *I have?*, and *It's* ***how*** *I wonder* ***what*** *you are*. Yet, even in this pristine sample, where children don't have to deal with noise in the input (because all the instances they encounter have the *wh*-word fronted), the ambiguity problem is real. These pristine data are compatible with all six of the hypotheses I explicitly mentioned before:

- *h1: require wh-fronting everywhere*. Every single example indeed has *wh*-fronting, whether the example is in a main or embedded clause. This is compatible with *wh*-fronting being required everywhere.

- *h2: optionally allow wh-fronting everywhere*. Every single example happens to use *wh*-fronting, whether the example is in a main or embedded clause. While this situation may be less probable if speakers are actually allowed to leave the *wh*-word in place (more on this below), these data are still perfectly compatible with the hypothesis that *wh*-fronting is optional.

- *h3: require wh-fronting in main clauses but optionally allow it in embedded clauses*. Every single main clause example indeed has *wh*-fronting, and every single embedded example happens to use *wh*-fronting. As before, while this situation is less probable if speakers are actually allowed to leave the *wh*-word in place in embedded clauses, these data are still compatible with the hypothesis that *wh*-fronting is optional in embedded clauses.

- *h4: optionally allow wh-fronting in main clauses but require it in embedded clauses*. Every single main clause example happens to use *wh*-fronting, and every single embedded example indeed has *wh*-fronting. As before, this situation is less probable if speakers are allowed to leave the *wh*-word in place in main clauses, but these data are still compatible with the hypothesis that *wh*-fronting is optional in main clauses.

[1] I note that I previously cast these hypotheses in terms of the language-wide generalizations that previous developmental approaches have used, but it's possible to cast these in terms of lexically-based features as well. The basic issue of ambiguity in this case shouldn't change.

- *h5: require wh-fronting in main clauses and singly-embedded clauses, but optionally allow it in multiply-embedded clauses*. Every single main clause and singly-embedded clause example indeed has *wh*-fronting, and every multiply-embedded clause example happens to use *wh*-fronting. As before, this situation is less probable, but still compatible with the hypothesis that *wh*-fronting is optional in multiply-embedded clauses.

- *h6: require wh-fronting in main clauses, optionally allow it in singly-embedded clauses, and require it in multiply-embedded clauses*. Every single main clause and multiply-embedded clause example indeed has *wh*-fronting, and every singly-embedded clause example happens to use *wh*-fronting. As before, this situation is less probable, but still compatible with the hypothesis that *wh*-fronting is optional in singly-embedded clauses.

The pattern is clear: if the data examples always use *wh*-fronting, these data are compatible with hypotheses that require *wh*-fronting as well as with hypotheses that optionally allow *wh*-fronting. Of course, as noted in the walkthrough above, some of these hypotheses may be more probable than others, given the data. For example, let's consider the difference between the first two hypotheses, where *wh*-fronting is required everywhere (h1) or *wh*-fronting is optional everywhere (h2). Let's also consider our 40,700 utterance sample with *wh*-words, which corresponds to about 17.4 days of input for an average three-year-old from a higher socio-economic status background.[2]

Under h1, all 40,700 instances used *wh*-fronting because they had to – that is, for any given instance, speakers have a 100% chance of using *wh*-fronting. So, the probability of using *wh*-fronting for any single instance is 1. Then, the probability of these data, given h1 = $p(data|h1)$ = $1 * 1 * ... * 1$ (40,700 times)=$1^{40,700}$=1.

Under h2, all 40,700 instances just happened to use *wh*-fronting when the *wh*-word could have been left in place. Let's say that speakers have a 50% chance of using *wh*-fronting in any given example. So, the probability of using *wh*-fronting for any single instance is 0.5. Then, the probability of using *wh*-fronting 40,700 times is really small: the probability of these data, given h2 = $p(data|h2)$=$0.5 * 0.5 * ... * 0.5$ (40,700 times)=$(0.5)^{40,700}$≈0.

So, a child capable of doing this kind of savvy statistical calculation would notice in a few weeks that h1 seems a lot more probable than h2, given these data. Happily, we have good reason to believe that children are capable of just this kind of savvy statistical learning (see Pearl 2021b for an overview). With this kind of learning ability, children could identify which hypotheses are the most probable, given the available ambiguous data. Still, the simple fact remains that the data *are* ambiguous, given the hypotheses that children could be considering – and children have to deal successfully with that ambiguity.

Noise. We idealized our input sample above to get rid of any noise – in this case, any instances where the *wh*-word remained in place. However, the real input data that children have to deal with includes misleading examples where the *wh*-word isn't fronted – these are the "noise" in the input signal. Some of these instances may be speech errors on the part of the speakers (e.g., a slip from a native speaker speaking too quickly, or a mistake from a non-native speaker whose native

[2]See Pearl and Bates (2022) for details on how to estimate how much input children hear over time.

language leaves *wh*-words in place like Japanese, Hungarian, Igbo, or Zulu). However, many of these instances may be appropriately leaving the *wh*-word in place, such as the echo questions and teaching scenarios mentioned earlier.

Let's consider again our sample of approximately 40,700 *wh*-word instances from the CHILDES Treebank, which is primarily American English caretakers speaking to children, rather than non-native speakers speaking to children. So, the *wh*-word usage is likely to reflect native speaker usage, rather than unintentional transfer of *wh*-usage by a non-native speaker. Of the approximately 40,700 *wh*-word instances, approximately 3,400 (8.3%) leave the *wh*-word in place. These include instances such as *Do **what** again?*, *She goes and gets **what**?*, and *But her Mommy tells her she has to do **what** before she goes?* These seem to be either echo questions or teaching scenarios (or both), and so aren't mistakes. Also, the fact that 8.3% of the input has the *wh*-word in place – rather than fronted – may be hard to ignore.

American English children nonetheless have to deal effectively with this apparent noise that's in their input. One reasonable solution is to notice that certain exceptions are systematic, and filter those data out when trying to identify the correct default generalization for the language. For instance, echo questions and teaching scenarios fulfill certain communicative goals and may have distinctive prosody associated with them. If children notice these predictable connections, then these seemingly-exceptional *wh*-data can be filtered out.

Another reasonable solution is to accept that there's going to be some amount of conflicting evidence (i.e., exceptions), and simply look for the hypothesis with "enough" evidence. Two reasonable ways to define "enough" are as the hypothesis with the most evidence or evidence above a certain threshold; the variational learning approach (Yang, 2002, 2004) looks for the most evidence while the approach using the Tolerance and Sufficiency Principles (Yang, 2005, 2016) looks for evidence above a certain threshold (see Pearl in pressb, Pearl 2021a, and Pearl 2021b, for recent overviews of both approaches).

Still, the point is that there is in fact noise in the input signal, given the complexity of the language system as a whole. That is, the noise in our example *wh*-fronting occurs because of the way the language system deals with *wh*-words, which is more complex than one single rule for all instances, irrespective of other factors. As with ambiguity, children have to deal successfully with the (apparent) noise.

2.2 Problem 2: Children's ability to harness the input signal is messy

Above, I discussed an external source of messiness: the ambiguity and noise in children's input signal. Another source of messiness is internal: children's developing ability to harness the information in the input signal (messy or not). In general, children are developing at least two key things simultaneously: their knowledge of their native language system, and their ability to process language data (Lidz and Gagliardi, 2015, Omaki and Lidz, 2015, Lidz and Perkins, 2018, Pearl, in pressa,b). Both of these can impact how children extract information from their input signal (see Pearl in pressb for more discussion on this point).

How developing representations can make the information messy. Recall with our *wh*-fronting example that certain types of utterances have the *wh*-word in place (echo questions and teaching scenarios). These examples seem to go against a default *wh*-fronting generalization. However, these utterances also seem to fulfill particular communicative goals and have potentially distinctive prosody. Suppose that children realize these communicative goals or recognize the distinctive prosody – that is, children's language representations include something about the communicative goal or prosodic information (or both) that marks echo questions and teaching scenarios as systematic exceptions. Then, children could filter these examples out when trying to decide if *wh*-fronting is the default. That is, their developed representations allow them to remove these examples from the set of data that has to be accounted for by a default *wh*-position generalization.

In contrast, suppose that children don't realize the impact of these communicative goals and prosodic information on *wh*-fronting. That is, children's developing representations of their language don't (yet) incorporate these communicative goals or the prosodic information for utterances with *wh*-words. Then, the echo questions and teaching scenarios are misleading counterexamples to default *wh*-fronting. The data that children learn from becomes messy because their developing representations don't yet separate out the apparent noise. At some point, children's developing representations will include enough information to mark these instances as systematic exceptions. Until then, these data are (incorrectly) included in the set of data that has to be accounted for by a default *wh*-position generalization.

How developing language processing can make the information messy. Throughout language development, children have various cognitive limitations compared to adults (e.g., limited memory, limited attention, how rapidly they can deploy their parsing abilities, among other limitations). These limitations affect how accurately children can extract information in language data. In our *wh*-fronting example, it might happen that children struggle to accurately extract information in longer utterances because of these limitations. So, for instance, a child might process *Do you want to see* ***what*** *I have?* as a truncated version, like *Do you want to see* ***what****?* In this case, an utterance with *wh*-fronting in the embedded clause is misperceived as an utterance with the *wh*-word in place. That is, an example compatible with a default *wh*-fronting generalization is misperceived as an example incompatible with this generalization, solely because the child wasn't able to process the part of the utterance occurring after the *wh*-word. In this way, the child's developing language processing abilities have skewed the information in the input signal in an unhelpful way, making it messier than it truly was.

3 Part 2: Could Minimalism help us understand how children overcome these problems?

It's possible that Minimalism could help us understand how children overcome these problems of both external and internal messiness during language acquisition. More specifically, because the representations children try to attain are different in Minimalism than previously thought, it's possible that the external messiness of the input signal is lessened; it's also possible the internal

messiness due to children's (mis)processing of the input is lessened.[3] The main point, however, is that the details of the representation matter enormously for both these acquisition problems.

3.1 Problem 1: How a different target representation could make the input signal less messy

Depending on the specifics of the Minimalist representation, it's possible that the input could be either less ambiguous or less noisy (or both). Of course, it's also possible that a Minimalist representation could in fact make the input more ambiguous and noisier, too. The key is that we have to have a specific Minimalist representation in mind – as well as the space of competing possible Minimalist representations that children consider. Once we have both of these (i.e., the specific target Minimalist representation and the Minimalist hypothesis space), we can evaluate how ambiguous children's input is by looking at realistic samples, such as those from the CHILDES Treebank (Pearl and Sprouse, 2013, 2019) or CHILDES more generally (MacWhinney, 2000). If it turns out that there are fewer potential hypotheses compatible with children's input than with previous representational approaches, then Minimalist representations would make children's input signal less ambiguous.

The same considerations hold for how Minimalist representations could affect the noisiness of children's input. That is, once we have a specific Minimalist representation in mind and the hypothesis space of competing Minimalist representations, we can evaluate these by using realistic child-input samples. In particular, we can evaluate (i) if there are misleading examples with respect to the target representation, and (ii) if so, how many there are. As with ambiguity, it could turn out that Minimalist representations cause the input to have fewer misleading examples than with previous representational approaches. If so, then the Minimalist approach will have reduced the noise in children's input signal.

3.2 Problem 2: How a different target representation could help children better harness the input signal

Again depending on the specifics of the Minimalist representation, it's possible that children's perception of the information in their input could be more helpful (though it could also be less helpful). That is, the information that children are able to extract from the input, given their developing linguistic knowledge and their developing processing abilities, could be more useful if children are trying to acquire a Minimalist representation. How might this occur? Remember that children's developing representations and developing language processing determine children's perception of two things: (i) what information is present in the input signal, and (ii) what information is relevant for learning about the representations under consideration. Because of this, children get a skewed view of their input.

[3] I'm currently unaware of specific Minimalist representation proposals that concretely explore these acquisition possibilities. However, see Goodluck and Kazanina (2020) for a review of how some aspects of Minimalism are compatible with certain empirical aspects of syntactic acquisition.

Notably, a general consensus in the acquisition community (echoed in much of my own work) is that children's skewed view of the input often seems *helpful* to acquisition (e.g., Pinker 1984, Dresher 1999, Lightfoot 1999, Pearl 2007, Pearl and Weinberg 2007, Pearl 2008, 2009, Pearl and Lidz 2009, Pearl 2011, Pearl and Sprouse 2013, Pearl 2014, Lidz and Gagliardi 2015, Pearl, Ho and Detrano 2016, Gagliardi, Feldman and Lidz 2017, Pearl 2017, Perkins, Feldman and Lidz 2017, Nguyen and Pearl 2019, Dickson, Pearl and Futrell 2022). That is, the limitations imposed by children's developing representations and developing language processing helpfully filter the information in their input, and this filtering is what allows children to achieve great feats of language acquisition.

One potential example of this kind of helpful filter is related to *wh*-words and their use (Pearl and Sprouse, 2013, Dickson et al., 2022). In particular, certain syntactic phenomena involving *wh*-words (sometimes called "syntactic islands") can be learned if children limit the information they pay attention to in the input. For instance, consider these two *wh*-questions: *What did this penguin do yesterday?* and *Who has that penguin seen this morning?* The information available in the input includes the lexical items (e.g., *did, has, do, see*), tense information (e.g., PAST), and the syntactic structure of the *wh*-question (e.g., structures corresponding to inflectional phrases (IPs) and verbs phrases (VPs)), among other information. My work with my colleagues (Pearl and Sprouse, 2013, Dickson et al., 2022) proposes that children only pay attention to certain information; here, these two *wh*-questions would both be perceived as something like IP_{PAST}-*VP*, despite having many differences in the actual input signal. With this filter on their input, children can then learn quite a lot about syntactic islands.

This idea of helpful filtering is in line with one well-known hypothesis for why children are better able than adults to acquire native-level proficiency in a language, known as the "Less is More" hypothesis (Newport, 1990). Put simply, this hypothesis describes why having "less" cognitive resources to use when acquiring language (like children) nonetheless leads to "more" language acquisition success than adults typically manage: children's limitations lead to helpfully skewed views of the information in their input. In the example above for syntactic island acquisition, the potential cause of the specific helpful filter on *wh*-utterances could be due to children's "lesser" cognitive resources. That is, the reason children only extract certain information from a *wh*-utterance (like IP_{PAST}-VP) is because they actually *can't* extract more information out in real time. This limited information extraction leads to a skewed view of the information in the input, which turns out to be potentially helpful.

Several computational modeling studies have demonstrated how this helpful skewing could occur in a variety of domains (e.g., Pearl, Goldwater and Steyvers 2011, Pearl and Sprouse 2013, Phillips and Pearl 2015, Pearl and Phillips 2018, Yang 2020, Dickson et al. 2022, Pearl and Bates 2022). Several behavioral studies have demonstrated that children and adults make different inferences on the basis of language acquisition input (e.g., Hudson Kam and Newport 2005, 2009, Hudson Kam 2015, Hendricks, Miller and Jackson 2018); one explanation for these different inferences is that children have different views of the information in the input than adults do.

So, if children have a helpfully skewed view of the information in their input, it could be that the skewing is even more helpful when the target representation is a Minimalist representation, compared with previous representational approaches. That is, the filtered subset of information in

the input that children use could enable them to more easily identify the target Minimalist representation within a defined hypothesis space of competing Minimalist representations. In this happy situation, adopting Minimalism would be helpful for understanding children's acquisition success. However, as before with issues of input ambiguity and noise, we can only evaluate this possibility when we have in mind both a specific Minimalist representation and a specific hypothesis space of competing Minimalist representations.

4 Concluding thoughts: Another potential benefit

While many questions remain to be explored for Minimalist representations and acquisition, I think the spirit of the Minimalist approach may offer an additional benefit for understanding language acquisition. In particular, the Minimalist approach aims to reduce the amount of language-specific knowledge that's biologically built into children (that is, to shrink the amount of material in Universal Grammar). Why is this potentially good for helping us understand language acquisition?

One answer is that reducing what's in Universal Grammar makes Minimalist explanations of language acquisition more amenable to including domain-general learning components. For example, there are domain-general learning mechanisms like statistical learning and probabilistic reasoning (e.g., see Pearl 2021a,b, in pressb for recent overviews of mechanisms like reinforcement learning and Bayesian inference in language acquisition); there are also domain-general learning preferences like preferring more compact representations (e.g., see Pearl (2022) for an overview of the Minimal Description Length framework in language acquisition). The benefit of domain-general components is that these components can solve some parts of the acquisition process (e.g., efficiently navigating a pre-defined hypothesis space: see Pearl 2021a for discussion). If these domain-general components solve those acquisition parts, then Universal Grammar doesn't have to have language-specific components to do so. So, Universal Grammar can be smaller, exactly as Minimalism aims to make it.

We also have a fair amount of evidence that (i) very young children can use certain domain-general components to learn, and (ii) these domain-general components can allow children to leverage their available input more effectively (e.g., see Pearl 2021a for an overview). That is, these domain-general components are plausible and useful cognitive components in a child's acquisition process. So, I think that by more easily incorporating plausible, useful domain-general components into explanations of language acquisition, a Minimalist approach could help provide more plausible, complete explanations about how children solve language acquisition. To me, this is a very helpful aspect of Minimalism indeed.

References

CHOMSKY, NOAM. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

CRAIN, STEPHEN; and PAUL PIETROSKI. 2002. Why language acquisition is a snap. *The Linguistic Review* 19.163–183.

DICKSON, NIELS; LISA PEARL; and RICHARD FUTRELL. 2022. Learning constraints on *wh*-dependencies by learning how to efficiently represent *wh*-dependencies: A developmental modeling investigation with Fragment Grammars. *Proceedings of the Society for Computation in Linguistics* 5(1).220–224.

DRESHER, B. ELAN. 1999. Charting the learning path: Cues to parameter setting. *Linguistic Inquiry* 30(1).27–67.

GAGLIARDI, ANNIE; NAOMI H FELDMAN; and JEFFREY LIDZ. 2017. Modeling statistical insensitivity: Sources of suboptimal behavior. *Cognitive Science* 41(1).188–217.

GOODLUCK, HELEN; and NINA KAZANINA. 2020. Fragments along the way: Minimalism as an account of some stages in first language acquisition. *Frontiers in Psychology* 11.584–591.

HENDRICKS, ALISON EISEL; KAREN MILLER; and CARRIE N JACKSON. 2018. Regularizing unpredictable variation: Evidence from a natural language setting. *Language Learning and Development* 14(1).42–60.

HUDSON KAM, CARLA L. 2015. The impact of conditioning variables on the acquisition of variation in adult and child learners. *Language* 91(4).906–937.

HUDSON KAM, CARLA L; and ELISSA 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, CARLA L; and ELISSA L NEWPORT. 2009. Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology* 59(1).30–66.

JACKENDOFF, RAY S. 1994. *Patterns in the mind: Language and human nature*. Basic Books.

LAURENCE, STEPHEN; and ERIC MARGOLIS. 2001. The poverty of the stimulus argument. *The British Journal for the Philosophy of Science* 52(2).217–276.

LIDZ, JEFFREY; and ANNIE GAGLIARDI. 2015. How Nature Meets Nurture: Universal Grammar and Statistical Learning. *Annual Review of Linguistics* 1(1).333–352.

LIDZ, JEFFREY; and LAUREL PERKINS. 2018. The importance of input representations. *Linguistic Approaches to Bilingualism* 8(6).743–748.

LIGHTFOOT, DAVID. 1999. *The development of language: Acquisition, change, and evolution*. Blackwell Oxford.

MACWHINNEY, BRIAN. 2000. *The CHILDES Project: Tools for Analyzing Talk*. Mahwah, NJ: Lawrence Erlbaum Associates.

NEWPORT, ELISSA L. 1990. Maturational constraints on language learning. *Cognitive Science* 14(1).11–28.

NGUYEN, EMMA; and LISA PEARL. 2019. Using Developmental Modeling to Specify Learning and Representation of the Passive in English Children. *Proceedings of the 43rd annual Boston University Conference on Language Development*, ed. by Megan M. Brown and Brady Dailey, 469–482. Somerville, MA: Cascadilla Press.

OMAKI, AKIRA; and JEFFREY LIDZ. 2015. Linking parser development to acquisition of syntactic knowledge. *Language Acquisition* 22(2).158–192.

PEARL, LISA. 2007. *Necessary Bias in Natural Language Learning*. College Park, MD: University of Maryland, College Park dissertation.

PEARL, LISA. 2008. Putting the Emphasis on Unambiguous: The Feasibility of Data Filtering for Learning English Metrical Phonology. *Bucld 32: Proceedings of the 32nd annual Boston University Conference on Child Language Development*, ed. by Harvey Chan, Heather Jacob and Enkeleida Kapia, 390–401. Somerville, MA: Cascadilla Press.

PEARL, LISA. 2009. Learning English Metrical Phonology: When Probability Distributions Are Not Enough. *Proceedings of the 3rd Conference on Generative Approaches to Language Acquisition, North America (GALANA 2008)*, ed. by Jean Crawford, Koichi Otaki and Masahiko Takahashi, 200–211. Somerville, MA: Cascadilla Press.

PEARL, LISA. 2011. When unbiased probabilistic learning is not enough: Acquiring a parametric system of metrical phonology. *Language Acquisition* 18(2).87–120.

PEARL, LISA. 2014. Evaluating learning strategy components: Being fair. *Language* 90(3).e107–e114.

PEARL, LISA. 2017. Evaluation, use, and refinement of knowledge representations through acquisition modeling. *Language Acquisition* 24(2).126–147.

PEARL, LISA. 2021a. How statistical learning can play well with Universal Grammar. *A Companion to Chomsky*, ed. by Nicholas Allott, Terje Lohndal and Georges Rey, 267–286. Wiley Online Library.

PEARL, LISA. 2021b. Theory and predictions for the development of morphology and syntax: A Universal Grammar+ statistics approach. *Journal of Child Language* 48(5).907–936.

PEARL, LISA. 2022. Poverty of the Stimulus Without Tears. *Language Learning and Development* 18(4).415–454.

PEARL, LISA. in pressa. Computational cognitive modeling for syntactic acquisition: Approaches that integrate information from multiple places. *Journal of Child Language* .

PEARL, LISA. in pressb. Modeling syntactic acquisition. *Oxford handbook of experimental syntax*, ed. by Jon Sprouse, Oxford University Press.

PEARL, LISA; and ALANDI BATES. 2022. A new way to identify if variation in children's input could be developmentally meaningful: Using computational cognitive modeling to assess input across socio-economic status for syntactic islands. *Journal of Child Language* 1–34.

PEARL, LISA; SHARON GOLDWATER; and MARK STEYVERS. 2011. Online Learning Mechanisms for Bayesian Models of Word Segmentation. *Research on Language and Computation* 8(2).107–132.

PEARL, LISA; TIMOTHY HO; and ZEPHYR DETRANO. 2016. An argument from acquisition: Comparing English metrical stress representations by how learnable they are from child-directed speech. *Language Acquisition* 1–36.

PEARL, LISA; and JEFFREY LIDZ. 2009. When domain-general learning fails and when it succeeds: Identifying the contribution of domain-specificity. *Language Learning and Development* 5(4).235–265.

PEARL, LISA; and LAWRENCE PHILLIPS. 2018. Evaluating language acquisition models: A utility-based look at Bayesian segmentation. *Language, Cognition and Computational Models*, ed. by A. Villavicencio and T. Poibeau, 185–224. Cambridge: Cambridge University Press.

PEARL, LISA; and JON SPROUSE. 2013. Syntactic islands and learning biases: Combining experimental syntax and computational modeling to investigate the language acquisition problem. *Language Acquisition* 20.19–64.

PEARL, LISA; and JON SPROUSE. 2019. Comparing solutions to the linking problem using an integrated quantitative framework of language acquisition. *Language* 95(4).583–611.

PEARL, LISA; and AMY WEINBERG. 2007. Input filtering in syntactic acquisition: Answers from language change modeling. *Language Learning and Development* 3(1).43–72.

PERKINS, LAUREL; NAOMI FELDMAN; and JEFFREY LIDZ. 2017. Learning an input filter for argument structure acquisition. *Proceedings of the 7th Workshop on Cognitive Modeling and Computational Linguistics (CMCL 2017)*, 11–19.

PHILLIPS, LAWRENCE; and LISA PEARL. 2015. Utility-based evaluation metrics for models of language acquisition: A look at speech segmentation. *Proceedings of CMCL* 68–78.

PINKER, STEVEN. 1984. *Language learnability and language development.* Cambridge, MA: Harvard University Press.

SAKAS, WILLIAM; and JANET FODOR. 2012. Disambiguating Syntactic Triggers. *Language Acquisition* 19(2).83–143.

SAKAS, WILLIAM GREGORY; CHARLES YANG; and ROBERT BERWICK. 2017. Parameter setting is feasible. *Linguistic Analysis* 41.391–408.

YANG, CHARLES. 2002. *Knowledge and Learning in Natural Language*. Oxford, UK: Oxford University Press.

YANG, CHARLES. 2004. Universal grammar, statistics or both? *Trends in Cognitive Science* 8(10).451–456.

YANG, CHARLES. 2005. On productivity. *Yearbook of Language Variation* 5.333–370.

YANG, CHARLES. 2012. Computational models of syntactic acquisition. *WIREs Cognitive Science* 3.205–213.

YANG, CHARLES. 2016. *The price of linguistic productivity: How children learn to break the rules of language*. MIT Press.

YANG, CHARLES. 2020. How to make the most out of very little. *Topics in Cognitive Science* 12(1).136–152.