

THE IMPACT OF CONDITIONING VARIABLES ON THE ACQUISITION OF VARIATION IN
ADULT AND CHILD LEARNERS

Author(s): Carla L. Hudson Kam

Source: *Language*, Vol. 91, No. 4 (DECEMBER 2015), pp. 906-937

Published by: Linguistic Society of America

Stable URL: <https://www.jstor.org/stable/24672252>

Accessed: 19-02-2021 13:23 UTC

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

Linguistic Society of America is collaborating with JSTOR to digitize, preserve and extend access to *Language*

THE IMPACT OF CONDITIONING VARIABLES ON THE ACQUISITION OF VARIATION IN ADULT AND CHILD LEARNERS

CARLA L. HUDSON KAM

University of British Columbia

Natural human languages often contain variation (sociolinguistic or Labovian variation) that is passed from one generation of speakers to the next, but studies of acquisition have largely ignored this, instead focusing on aspects of language that are more deterministic. Theories of acquisition, however, must be able to account for both. This article examines variation from the perspective of the statistical learning framework and explores features of variation that contribute to learnability. In particular, it explores whether conditioning variables (i.e. where the pattern of variation is slightly different in different contexts) lead to better learning of variation as compared to when there are no conditioning variables, despite the former being conceptually more difficult. Data from two experiments show that adult learners are fairly good at learning patterns of both conditioned and unconditioned variation, the latter result replicating earlier studies. Five-to-seven-year-old children, in contrast, had different learning outcomes for conditioned versus unconditioned variation, with fewer children regularizing or imposing deterministic patterns on the conditioned variation. However, the children who did not impose deterministic patterns did not necessarily acquire the variation patterns the adults did.*

Keywords: acquisition, sociolinguistic variation, language change, artificial languages

1. INTRODUCTION. Research on first language acquisition is generally concerned with understanding what children know about the language in their environment and when (in development) they know it, with the goal of understanding something about the mechanisms involved. This kind of research has been quite fruitful: we now know a tremendous amount about children learning a variety of languages, and there has been a great deal of theorizing, if not agreement, about the underlying mechanisms leading to the patterns we see. A case in point is the literature surrounding English-acquiring children's learning of past-tense forms: it has yielded a great deal of data as well as work on mechanisms (e.g. Marchman 1997, Marcus et al. 1992, Ramscar 2002, Rumelhart & McClelland 1986). However, while research in child language acquisition has traditionally focused on aspects of language that are more or less deterministic, there is by now a great deal of evidence that many of the patterns children are exposed to and eventually learn are not deterministic; rather, they are variable. Indeed, an entire field of research, namely sociolinguistics, is focused on describing and explaining linguistic variation. (The case of the English past tense is no exception to this generalization about the typical target of work in acquisition; even though the way past is encoded is variable, it is consistent at the level of individual verbs.) Yet variation has received much less attention in the acquisition literature, which is unfortunate; understanding how human learners cope with deterministic and variable aspects of the language in their input is crucial if we are to understand the totality of humans' abilities to acquire language. This project represents a first step at remedying this situation. It examines the acquisition of variation from within a theoretical framework that is currently influential in the field of language acquisition, namely statistical learning, making explicit links with existing work on the acquisition of nondeterministic aspects of language. Moreover, it employs a

* This research was supported by National Institutes of Health Grant HD 048572, and a Discovery Grant (Individual) from NSERC, both to Carla L. Hudson Kam. The author thanks Sarah Wilson and Whitney Goodrich for their assistance. Thanks also to the participants of ViLA 2012 for comments on a presentation based on these data, and to Susannah Kirby, Marie Coppola, Shana Poplack, and two anonymous referees for comments on an earlier version of the manuscript.

methodology, miniature artificial language learning, that facilitates the investigation of mechanisms. In so doing, it brings the study of the acquisition of variation more squarely into the domain of things that are relevant for those interested in acquisition more broadly, while highlighting the need for more research on this topic.

1.1. THE ACQUISITION OF VARIATION. When thinking about language acquisition, researchers tend to think about things such as how children learn that their language is SVO versus SOV, that the unaspirated /p/ following an /s/ and the aspirated /p/ that comes at the beginning of a word are allophones of the same phoneme (and so in the same category), or that *-ing* indicates an ongoing action and cooccurs with some variant of the copula. Learning these kinds of consistent regularities certainly comes with challenges for the learner, so they are very worthy of study. However, consistent regularities of this sort are not the totality of what children have to learn. Much of what children end up knowing is not regularities of the deterministic sort mentioned above, but rather probabilities of occurrence. Probabilistic knowledge of the language ranges from things like the likelihood that a direct object versus a sentential complement follows a particular verb (e.g. Garnsey et al. 1997; cf. Newmeyer 2003) to more complicated Labovian or sociolinguistic variation (see e.g. Chambers et al. 2002), and how children come to have this kind of knowledge has received very little attention to date.

Sociolinguistic variation is especially interesting from the perspective of acquisition because of the apparent challenge it presents to children. Take the well-studied example of *-t/-d* deletion in English. Word-final /t/s and /d/s occurring as the second sound in a consonant cluster are often deleted. This deletion is not random, however; it is probabilistically affected by a variety of things, including the following phonological context and the morphemic status of the *-t/-d* with respect to the word it is in (Labov 1989). In particular, deletion is more likely if the following segment is an obstruent than if it is a liquid, which in turn leads to more deletion than a following glide, which is itself more likely to lead to deletion than a following vowel (Labov 1989). The rates of deletion are not just different from one another; they also have their own likelihoods. That is, while it is correct to say that deletion is more likely before an obstruent than a liquid, it is also the case that each of those environments has its own approximate rate of deletion. While the preceding cline is shared across dialects of English, there are also differences between dialects. In particular, following pauses are sometimes related to high rates of deletion (as in New York City), and sometimes to low rates of deletion (as in Philadelphia) (Guy 1980).

Thus, it is clear that speakers have to learn the particulars of how deletion works. They have to learn which factors are involved (e.g. the following phonological context) and what the specific likelihoods associated with each level or category of the factor are. Add in the effects of social variables that can change or modify the likelihoods, and it is clear that the patterns a speaker must acquire are very complicated. It is important to note that sociolinguistic variation is variable in two ways. First, it is the variation between two or more potential forms (in this case, *-t/-d* and \emptyset), and second, the occurrence of one form or the other is not deterministic. That is, one cannot say with 100% certainty that a particular *-t/-d* will be deleted in a particular utterance, because the process itself is variable, or probabilistic. Although the variation is not free, it is not completely determined, either.

The growing literature on children's acquisition of variable rules in natural languages, while sparse, shows that children can learn variation.¹ While this might seem

¹ I use the term *variable rules* here following the Labovian tradition. While there are other ways of characterizing variation, such as exemplar theory (e.g. Foulkes 2010), the particular characterization does not affect the discussion in the present article, so I have chosen to use the more historically common terminology.

self-evident, it is not a given; it would be entirely possible, for instance, for variation to be so complicated that it takes many years of language exposure to be acquired and so is only acquired in adolescence or later. Indeed, the literature shows that there is no age at which variation can be expected to be acquired across the board. Roberts (1997), for instance, examined children's knowledge of *-t/-d* deletion in Philadelphia. In her sample of sixteen three- and four-year-old children, she found that even the youngest children generally showed the same patterns of deletion as adult speakers of the same dialect, at least in terms of gradations of deletion occurrence by phonological environment. The children differed from adults when the morphemic status of the *-t/-d* was considered, however, demonstrating that the children were not simply mimicking deletion rates in individual words they had heard but rather were in the course of learning a system—they just had part of it wrong. Interestingly, the children she studied had not yet acquired the social constraints on deletion; this contrasts with the results of Labov (1989), who found that stylistic constraints were learned quite early and suggested that stylistic constraints on variation may actually be learned earlier than linguistic ones.

Roberts (1997) and Labov (1989) both focused on the speech produced by children as compared to what was already known about norms for the adult speech community. That is, children's targets and inputs were assumed to be the community adult norms. (Although some limited data from eight mothers of children in the study were collected and analyzed, they were collected in adult-adult interactions.) More recent work by Smith, Durham, and Fortune (2007) on variation in a Scottish dialect demonstrates that the language actually heard by children may differ rather dramatically from the community adult norms, a finding consistent with those of Roberts (2002). Smith and colleagues compared children's productions to their actual input, and found that children match their input quite closely. Thus, when children deviate from community norms, they may be doing so precisely because their input did, rather than because one kind or aspect of variation is harder to learn than others.

Of particular relevance for the issue at hand, Smith and colleagues (2007) compared the productions of individual children and their caregivers and found very tight correspondences between the two, thereby showing not only that children can generally learn variation, but also that they can learn the very particular probabilities present in their caregivers' speech. This is consistent with several studies showing that children from families with high socioeconomic status use more standard variants than those from families with lower socioeconomic status from an early age (Díaz-Campos 2005). That is, they learn what they are hearing. This does not mean, however, that young children ascribe social meaning to variation or control stylistic variation. That seems to come in much later (see Nardy et al. 2013). The increasing use of nonstandard variants over early childhood found in some studies (see Nardy et al. 2013) also has a potential source in the input: Smith, Durham, and Richards (2013) found that the speech of older children's caregivers contained more nonstandard variants than that of the younger children's caregivers, and this was reflected in the children's speech. (This was true only for nonstandard variants that were somewhat socially marked. Less stigmatized nonstandard variants were highly frequent even in the youngest children's input.) In apparent contradiction to Smith, Durham, & Fortune 2007, Smith, Durham, and Richards (2013) found that younger preschoolers matched the production probabilities of their caregivers' speech much less than older preschoolers did. They suggest, however, that this is due to changes in the caregivers' speech over the younger ages: as children gain linguistic competence the caregivers shift their speech to be more adult-like (i.e. less standard), and it takes some time for the children to catch up to the new patterns. Even

differences in variation in children's speech by gender can be linked to input differences (Foulkes et al. 2005). Thus, the data suggest that as long as children are hearing variation they will learn it, and can do so from the very beginning of language acquisition, despite the apparent complexity of variation.

1.2. STATISTICAL LEARNING STUDIES. At first glance, children's success at learning sociolinguistic variation is not surprising, given the by now well-known work on statistical language learning. It has been repeatedly demonstrated in this line of work that humans of all ages are quite sensitive to probabilities, and that they can use this ability to acquire aspects of language. For instance, Saffran, Aslin, and Newport (1996) showed that infants (and adults; Saffran, Newport, & Aslin 1996) can use transitional probabilities (TPs) between syllables—that is, the likelihood that a particular syllable follows another individual syllable—to segment words out of running speech. This ability is not limited to syllables, however. Finn and Hudson Kam (2008) showed that adults can compute TPs over segments, and various studies have shown that TPs can also be computed over nonlinguistic sounds (Saffran et al. 1999), visual objects (Fiser & Aslin 2002), and even movements (Hunt & Aslin 2001). Moreover, statistical learning abilities have also been shown in nonhuman animals (Hauser et al. 2001, Toro & Trobalón 2005). Thus, the abilities demonstrated by Saffran, Aslin, and Newport appear to be quite general in that they are neither species- nor domain-specific. Nevertheless, recent studies show a relationship between individuals' performance on statistical learning tasks and their language-learning abilities, demonstrating that these abilities are indeed relevant for language acquisition (Evans et al. 2009, Kidd 2012).

However, the link between the kinds of computational abilities demonstrated in statistical learning tasks and the acquisition of sociolinguistic variation is not as clear as it may seem. Sociolinguistic variation involves probabilities of particular forms in particular contexts, but most of the work on statistical learning examines the ability to use probabilities to extract larger forms from a stimulus—in the case of Saffran, Aslin, & Newport 1996, using probabilities between syllables to extract multisyllabic words, and in other work using probabilities between word classes to find phrases (Thompson & Newport 2007, Wilson & Hudson Kam 2014)—which is quite different. And while there is ample evidence that the extracted chunks are stored as units (Graf Estes et al. 2007, Mirman et al. 2008), there is much less evidence that the probabilities themselves are part of those representations (Perruchet & Poulin-Charronnat 2012). Indeed, there is some disagreement as to whether TPs are actually being computed in such tasks and, even if they are, whether they are computed when the learning task becomes more realistic (Arnon & Ramscar 2012, Perruchet & Pacton 2006; cf. Aslin et al. 1998).

There is an older literature on probability learning that might be more relevant for thinking about the acquisition of sociolinguistic variation: the probability-learning literature. In these studies participants were exposed to a stimulus containing probabilistically occurring events, for example, a set of lights each of which flashes according to some underlying probability. Participants were asked to make predictions about each event prior to its occurrence, for example, to predict which light in the set will flash next. The general finding in such studies is that adults quite quickly come to predict event types according to their underlying probabilities of occurrence. For example, if there are two lights, one that flashes 70% of the time (light A) and the other 30% of the time (light B), 70% of the time participants will predict that the next light to flash will be light A (see Estes 1964, 1976). The relevance of this line of work is not lost on sociolinguists. Labov (1994), for instance, refers to this literature when discussing the acquisition of sociolin-

guistic variation. However, this exquisite sensitivity to probabilities has only been demonstrated in adults in such studies. Studies of probability learning in children typically show a very different pattern. In particular, children do not learn the probabilities; rather, they impose patterns on the event types, things like A, A, B, A, A, B (to continue using the previous example) (Stevenson & Weir 1959, Stevenson & Zigler 1958, Weir 1964; see Hudson 2002 for a review). Thus, the probability-learning literature suggests that children, but not adults, should have difficulty learning sociolinguistic variation.

1.3. REGULARIZATION OF VARIATION IN LANGUAGE THROUGH LEARNING. Interestingly, there is some artificial language work that mirrors the findings of the probability-learning literature. In the first of these studies, Hudson Kam and Newport (2005) exposed adults and children five-to-seven years of age to an artificial language containing variation modeled on that found in the errors of late learners (Adamson 1988, Johnson et al. 1996, Kanno 1998, Meisel et al. 1981, Newport 1988, 1990, Sorace 2000, Wolfram 1985).² They found that while adults learned the variation veridically, children imposed regularity on the language. In the input languages used in the study, the determiners were variable and inconsistent; they sometimes occurred with nouns but not always, and their presence (or absence) was not in any way predictable in any individual noun phrase. However, it was also not completely random, as the probabilities of determiner occurrence were the same in each of the exposure sessions, as well as for nouns occurring in different sentence types or syntactic positions. But it was not conditioned on anything, as sociolinguistic variation is. After six days of exposure to the language, participants' knowledge of determiner usage was tested. Hudson Kam and Newport found that adult learners replicated their input probability of determiner occurrence when they produced novel utterances, but that most children failed to learn the same information. Instead, the children imposed a systematicity on the determiners that was not present in the input. Put another way, they regularized. In a subsequent study adults and children were exposed to a slightly different kind of variation—competition between several different determiner forms occurring in the same context, with one being more frequent than the others (Hudson Kam & Newport 2009). Again, which determiner form would occur in any sentence or noun phrase was not predictable, but the probabilities of occurrence were consistent across input sessions, sentence types, and syntactic types, meaning that the variation was not random. This more complicated variation produced somewhat different results: children again regularized the language, but now the adult learners did also, though to a lesser degree than the child learners.

Numerous other studies since have found similar results. Austin and colleagues (2006), for instance, found widespread regularization of determiners by five-to-six-year-old children, but not adults, in a variety of different language conditions, and Hudson Kam (2009) found that adults could also learn the probabilities of different basic word orders in an artificial language. Wonnacott and Newport (2005), however, showed that adults are prone to regularize variation in both determiner usage and word order when the test sentences contain vocabulary items that were not in the training vocabulary and so are novel at test.³ And Smith and Wonnacott (2010) found occasional regu-

² As pointed out by a referee, sometimes variation is conditioned in the speech of late learners. There is, for instance, evidence that second language learners can acquire sociolinguistic variation to some degree (Mougeon et al. 2010; see Regan 2013 for a summary). Additionally, there can be patterns in their errors that depend on the same kinds of factors seen in sociolinguistic variation and that are shared across speakers (see, for example, the papers in Bayley & Preston 1996).

³ Interestingly, this is also true for test sentences containing all old (i.e. training set) vocabulary items if they are mixed in with sentences containing novel vocabulary test items (Wonnacott & Newport 2005).

larization of variable plural marking by adult learners.⁴ Thus, it seems that while children are very likely to regularize variation in language, adults will as well under certain circumstances. Indeed, in related literature on iterated learning, where learners get a previous learner's output as their input, chains of adult learners almost always end up regularizing over variation (Reali & Griffiths 2009, Smith & Wonnacott 2010).

The experimental work just discussed was not intended as a linguistic analogue to the probability-matching work; it was motivated by real-life examples of children regularizing variation in their language input. The most comprehensive discussion of regularization of probabilistic patterns in naturalistic acquisition is found in work by Newport and colleagues, who studied a child (they called Simon) who learned American Sign Language (ASL) exclusively from his late-learner parents (Newport 1999, Ross & Newport 1996, Singleton & Newport 2004). As is often the case with late learners, the parents' signing contained errors. Their errors comprised omissions as well as the production of incorrect morphological forms given the context. Importantly, the errors evinced unsystematic variation (Singleton & Newport 2004), even at the level of individual morphemes, something common to late learners (Johnson et al. 1996, Kanno 1998, Lardiere 2007, Meisel et al. 1981, Wolfram 1985). (This is not to say that all variation in L2 speech/signing is unsystematic; see n. 2.) Instead, the parents sometimes produced the right form but other times produced an incorrect form, where 'correct' simply means the form commonly used by native signers. These errors introduced variability into Simon's input: sometimes the form produced by his parents was the correct form, but other times it was not, and the variation between the correct form and incorrect alternatives was not predictable. When Simon's signing was compared to that of his parents, it did not contain the same vacillation between forms present in his parents' productions. Instead, his signing looked much like that of children who learned from native-signing parents. He had imposed regularity onto the variation by using the most frequent form all the time.⁵

This regularization (or FREQUENCY BOOSTING, as Singleton and Newport (2004) call it) was not apparent for all morphemes, however: it did not happen with classifiers (Ross & Newport 1996, Singleton & Newport 2004). Classifiers, it turns out, are hard even for children learning from native-signer input and are the morpheme type over which Simon's parents were the least consistent, producing the correct forms less than half of the time, in contrast to at least 65% of the time for the other morpheme types examined (Singleton & Newport 2004). Exactly why Simon did not regularize the classifiers is not entirely clear; however, follow-up work by Ross and Newport (1996) suggests that the failure to regularize the classifiers is not a result of the low frequency of the correct ASL forms in Simon's parents' signing. They studied several other children who were in similar, but not identical, situations. These children were also deaf,

⁴ This was actually a study of iterated learning or chains of individual learners, investigating the possibility for adult learners to regularize languages. Smith and Wonnacott (2010) found regularization in several of their five-learner chains. Notably, it was always the result of a single learner regularizing the language and this more consistent language being passed along to the next learner and being maintained as consistent by that and all subsequent learners. (It is worth noting that Smith and Wonnacott removed additional variation introduced by the learners, for example, blends of the variable forms, which artificially removed one learning outcome at the level of the chains, namely, a final system with increased variation as compared to the initial seed language.) Although there are critiques of learner chains as models of how change occurs in natural languages (Niyogi & Berwick 2009), for the purposes of the present study, all that is relevant is the fact that some learners do impose regularity on the variation they experience.

⁵ In one instance the parents' most frequent form was not the correct ASL morpheme, and this incorrect morpheme was the one Simon regularized (Singleton & Newport 2004).

but were born to hearing parents who learned a sign language in response to having a deaf child. The parents were much less proficient than Simon's parents and provided much worse input than Simon's parents (who had used ASL as their primary language for many years prior to the birth of their son). Despite receiving even less consistent input, Ross and Newport found that these children also regularized the (often faint) probabilistic patterns present in their input. Thus, the regularization witnessed by Singleton and Newport (2004) was not an isolated idiosyncratic learning outcome, nor does it only result from observing fairly frequent probabilistic patterns.

1.4. SUMMARY AND RESULTING RESEARCH QUESTION. From the perspective of a learner, sociolinguistic variation is complicated; it involves learning numerous probabilistic relationships between different aspects of the linguistic signal—for example, the occurrence of a segment and the phonological or morphological context—that can be modulated by things outside of the linguistic signal, such as the communicative and social context in which the speech is uttered. Despite the apparent complexity of sociolinguistic variation, the real-world facts indicate that such learning is possible; children do learn sociolinguistic variation. This is more or less consistent with most of the experimental literature on statistical learning, but it conflicts with what we would expect based on other literature on learning, in particular, the probability-learning literature. Additionally, it conflicts with data from other naturalistic and experimental investigations of learners acquiring language containing variation. This latter conflict is especially interesting, because the variation examined in the miniature artificial language experiments is descriptively simpler than most if not all variation discussed by sociolinguists, since the former involves simple probabilities, which should be easier to learn, not harder. And yet the descriptively more complex variation is learned, while the seemingly simpler variation is not. From a learning perspective, then, the learning outcomes are quite perplexing.

Why do we find learners, especially children, failing to acquire variation in some studies but not others? There are several possible explanations that are easily dismissed. One is that participants regularize in experiments but not in real life. Although not unreasonable, since there are many differences between natural language acquisition and artificial language experiments, this possibility is easily dismissed by the evidence: the experiments were inspired by cases of regularization in real language learning 'in the wild' (e.g. Singleton & Newport 2004). Another possibility is that probabilistic phonological variation is learnable but probabilistic variation that is morphosyntactic is not—it is regularized. This is consistent with Payne (1980), who suggested that phonetically conditioned phenomena (i.e. variation) are easier to acquire than morphologically conditioned ones in child second dialect acquisition. However, although it is true that much of the stable (and so by definition learnable) variation studied by sociolinguists is phonological, there is also evidence for stable morphosyntactic variation (Auger 1998; cf. Buchstaller 2009). Indeed, some of the variables examined in studies of the acquisition of sociolinguistic variation are morphosyntactic (Smith et al. 2007, Smith et al. 2013, and see Levey 2006a,b for work on the acquisition of discourse-level variation). Another possibility that might occur to some readers is the lack of any social conditioning in the variation that is regularized by learners, suggesting that social conditioning might be a necessary feature of learnable variation. However, if that were the crucial distinction we would expect social aspects of variation to be learned earlier than linguistic factors in typical acquisition situations, which is not the case; as discussed previously, although social conditioning can be learned earlier than was once thought, it is typically learned after linguistic conditioning (see Nardy et al. 2013 for a review).

Another possibility is the very presence of conditioning variables. Stable variation is conditioned on numerous factors that all affect the likelihood of one form occurring versus the other. In the case of *-t/-d* deletion, for instance, the following phonological environment and the morphemic status of the *-t/-d* both change the likelihood of the segment being produced by a speaker (Labov 1989). Moreover, the conditioning is shared by all speakers within a speech community. Note, however, that the conditioning factors do not determine whether the sound will be deleted in any particular utterance; that is, one can never say with 100% certainty what will happen. In contrast, the variation encountered by Simon and the learners in the experimental studies can be described as unconditioned; there is an overall likelihood language-wide of one form occurring versus another that is unaffected by any intra-utterance or social factors in any consistent way.

On the one hand, this hypothesis seems self-evidently true; this is a true characterization of the difference between the two kinds of variation. However, that does not mean that conditioning, or lack thereof, is what leads to the various learning outcomes. First, this is not the only difference. In circumstances where variation is learned veridically, the learner's input is provided by multiple speakers who share the same variable system. In cases where learners have been shown to change the language as they learn it, either a single speaker provided the input (as in Hudson Kam & Newport 2005, 2009 and Ross 2001) or multiple speakers who produced different variable patterns provided the input (Kotsinas 1988, Singleton & Newport 2004). This is true even of cases where the between-speaker differences emerge from contact between speakers of different dialects rather than nonnative speaker idiolects (Kerswill & Williams 2000). Thus, there is at least one other plausible cause of regularization or, rather, one plausible trigger for learning variation, namely shared patterns of variation, which is missing in documented cases of regularization.⁶ Second, as hinted above, this description of the difference (conditioning leading to learning) would appear to predict that stable sociolinguistic variation should be HARDER to learn, not easier: in the case of unconditioned variation a learner has to acquire only one probability, but in the case of conditioned variation several probabilities (and possibly their interactions) must be learned. However, the very presence of shared conditioning factors may signal to the learner that the probabilities are important and should be learned and followed. Whatever the mechanism, it is at least possible that it is consistent conditioning per se that leads to some variation being learned rather than regularized.

If so, we would expect to see consistent conditioned variation being learned even in experimental settings. This possibility is investigated in two experiments. In the first, adult learners were exposed to a miniature artificial language based on the ones used by Hudson Kam and Newport (2005, 2009) but containing conditioned morphosyntactic variation. In the second experiment, child learners were exposed to a similar language. In both studies the question is whether this kind of variation proves to be more learnable than unconditioned variation.

2. EXPERIMENT 1. This first experiment serves as an important check on the method. We know that adult second language learners can acquire sociolinguistic variation, although their command of it is usually not quite native-like (Mougeon et al. 2010), mirroring what is seen for other aspects of language (namely, few adults acquire totally

⁶ This fits in with a growing literature showing that children seem to rather naturally infer that language is shared knowledge (see e.g. Sabbagh & Henderson 2007). If they see language as something that is shared by multiple speakers, they might need evidence for shared variation in order to learn it.

native-like competence). Given this, we would expect adult learners to be able to acquire conditioned variation in a miniature artificial language if this kind of experiment taps into the same abilities as the learning of natural languages.

To assess whether learners can acquire conditioned variation in an experimental setting, adult learners were exposed to one of three versions of a miniature artificial language containing variation. In all three languages the determiners occurred with nouns probabilistically; what differed was the presence of an additional conditioning variable, namely syntactic role, in two of the languages but not the third. Specifically, in two languages the proportion of nouns occurring with determiners was different for subject and object nouns (object here always means direct object). Importantly, the conditioning variable used is one that we know is available to our participants in their native language, and previous work has shown that adult learners can acquire probabilities involving just these syntactic categories (Hudson Kam 2009). Thus, the categories involved in the conditioning should not, by themselves, present any problems for the learners. Note that although the nature of the conditioning is somewhat different from most sociolinguistic variation, if it is conditioning *per se* that facilitates the learning of variable patterns, then we should see learning independent of the particulars of the conditioning.

2.1. METHOD.

PARTICIPANTS. Of the fifty-six native English speakers who completed the study, eight did not pass the vocabulary criterion test (described in more detail below) and so did not complete the tests. The forty-eight participants whose data are included comprised thirty-seven females and eleven males, mean age 22.5 years (range: eighteen to forty-five). All were students at the University of California, Berkeley, or members of the Berkeley community at the time of the study. They were recruited via flyers posted around campus and were paid for their participation.

STIMULI. The artificial language used in this study consisted of forty-one words: twenty-eight nouns divided into two noun classes, six intransitive verbs or predicates, four transitive verbs or predicates, one negative, and two determiners, one for each of the two noun classes. The noun classes are much like masculine and feminine in French or Spanish (except that nouns were assigned to classes arbitrarily, which is how adult learners apprehend the less arbitrary classes in real languages; see Cain et al. 1987). The only consequence of noun class is determiner selection: nouns in each class occur with a different determiner. Note that as in natural languages, the distribution of nouns in the two classes is not equal; there were thirteen nouns in one class and fifteen in the other. This language is a smaller version of the ones used in Hudson Kam & Newport 2005, 2009, facilitating comparison across studies.⁷

The nouns included animate and inanimate nominals, and the inanimates included both count and mass nouns (according to English classifications). The language had both intransitive and transitive verbs, but no ditransitives. The meanings of the words in the verb category included meanings encoded by verbs in English as well as meanings encoded by adjectives and prepositions (e.g. ‘be big’, ‘be red’ (intransitive), and ‘be inside (of)’ (transitive)). Although having these sorts of meanings be expressed by verbs may be unusual, it is not unheard of in actual languages (see e.g. Dryer 2005, Stassen 1997, 2005). The negative is a general negative word that turns a positive statement into a negative, as in 1.

⁷ The lab was finding a very high rate of failure on the vocabulary criterion among participants exposed to the larger language, so the language was adjusted slightly to ensure better learning.

- (1) sig filk genta poe
 NEG be.red plastic.donut DET2
 'The donut isn't red.'

As exemplified in 1, the word order was VS(O), and determiners followed nouns, as is common in VSO languages (Greenberg 1963). Importantly, the differences in word order and in the mapping of meaning to grammatical category between English and the miniature language ensured that participants had to learn the grammar of the new language; they could not simply transfer their L1 knowledge onto a novel vocabulary.

One hundred and twenty video scenes with meanings expressible in the artificial language were created. Half corresponded to intransitive sentences, the other half to transitive sentences. These scenes and their accompanying sentences in the artificial language were the exposure set. Each noun occurred one to three times in the intransitive exposure set (mode = 2), and one to six times as the subject (mode = 2) and one to four times as the object (mode = 2) in the transitive sentences. Overall, each noun occurred four to twelve times in the entire exposure set, distributed across the three grammatical roles. Each intransitive verb occurred eight to twelve times in the exposure set, and each transitive verb ten to seventeen times. There were relatively few negative sentences in the exposure: two transitive and fifteen intransitive sentences. Negative sentences were primarily included to help participants learn the meaning of the verbs, especially the intransitives. No word was presented only in negative sentences; individual nouns and verbs were presented either in both negative and positive sentences or only in positive sentences.

PRESENTATION. Participants were told that they were in a language-learning experiment and that we were interested in differences between child and adult language learners. They were instructed to pretend that they were on a desert island and that they had to learn the local language by listening to the inhabitants speak it. They were told that they would be tested at the end of the experiment to see what they had learned about the language. They were aware that they would be asked to speak the language at test time.

Exposure occurred via videotapes showing scenes or events that were accompanied by a spoken sentence describing the scene or event in the miniature language. The entire exposure set of 120 scenes and their accompanying sentences was presented to participants six times over the course of the experiment, during eight exposure sessions each lasting twenty-five to thirty minutes. Each session contained approximately ninety of the exposure sentences. Sentences, recorded by a female researcher, were spoken at a normal rate with English prosody and phonology and sounded very natural and fluent.

There was no explicit instruction in either vocabulary or grammar; whatever participants learned, they learned solely from the auditory exposure to the sentences along with their accompanying visual scenes. Participants were asked to repeat each sentence after hearing it, although there was no monitoring of this. They were informed that this pronunciation practice would be helpful later when they had to produce their own sentences during testing at the end of the experiment. Participants were run individually in a quiet room. The experimenter was present at the beginning of the session, but left the room as the video started to play.

The entire experiment took nine sessions (eight exposure sessions and one test session). Participants completed the experiment over eleven to thirteen days.

EXPERIMENTAL MANIPULATION. Participants were divided into three different input conditions. In all three conditions the determiners occurred with nouns probabilistically. In two conditions the overall rate of occurrence was 60%. That is, in the entire ex-

posure set, 60% of the time nouns occurred with determiners and 40% of the time they did not. (This mirrors exposure in the studies reported in Hudson Kam & Newport 2005, 2009.) In the third condition, the overall rate of determiner occurrence was reversed, such that 40% of the time nouns were accompanied by determiners and 60% of the time they were not. (The reason for this difference is explained below.) Recall that the 120-sentence exposure set was presented to participants six times. The rate of determiner occurrence (60% or 40%, depending on condition) was true of the experiment overall, as well as for each individual run through the exposure set. It was also true for each of the two noun classes, both overall and within a single run through the exposure set. There was variation over nouns, such that some nouns occurred with the determiner more or less often, and this varied by exposure set. That is, it was not the case that the likelihood of determiner occurrence for each noun was consistent across sessions. It also varied for individual sentences, which were presented five times over the course of the exposure. Since each individual presentation of a sentence was independent of the other five presentations, it could be different each time (within limits, since there were only two possibilities with respect to determiner presence in intransitive sentences and four for transitives).

Two things differed between the input conditions: whether the variable determiner occurrence was linguistically conditioned and the particulars of the conditioning, if present. In one input condition, unconditioned variation ($-CV\ S=O$), the variation was not conditioned. As in previous studies employing a similar methodology, determiners occurred equally often with subject and object nouns and, therefore, in transitive and intransitive sentences. In the other two input conditions the likelihood differed according to the syntactic role of the noun. In the $+CV\ S>O$ condition, determiners occurred more often with nouns in subject position than with nouns in object position. As with sociolinguistic variation in natural languages, it was not simply a more versus less contrast; there was also a specific probability of determiner occurrence associated with each syntactic position. In this condition determiners occurred with nouns in subject position 80% of the time. This was true for both transitive and intransitive subjects. Object nouns, in contrast, occurred with determiners only 20% of the time. In the other input condition with conditioned variation ($+CV\ O<S$), the relation between syntactic position and determiner occurrence was flipped, such that determiners occurred with subject nouns 20% of the time and with object nouns 80% of the time. Because half of the sentences in the input set were transitive and half were intransitive, there were more subject nouns than object nouns in the set. Thus, the overall rate of determiner occurrence is different for the two conditioned-variation languages. In the first, the overall rate is 60%, just like the condition with unconditioned variation ($-CV\ S=O$). In the second, the overall rate is 40%. Importantly, although the percentages differed for subject and object nouns in both versions of conditioned variation, determiner occurrence was still probabilistic, both overall and when considered for any specific position. Thus, it is conceptually very similar to linguistically conditioned sociolinguistic variation. A summary of the conditions is presented in Table 1. As with the overall percentages, these syntactic-position-specific percentages were true of the participants' language experience overall and within each of the six exposure sets, although the particular instantiation of the percentages differed by sets, since each exposure set was created independently so as to conform to the intended patterns.

TESTING. The final (ninth) session consisted of testing. Participants were first given a vocabulary test in which they were tested on a subset of the nouns in the language (the

CONDITION	OVERALL	NOUN TYPE	
		SUBJECTS	OBJECTS
-CV (S=O)	60%	60%	60%
+CV (S>O)	60%	80%	20%
+CV (O>S)	40%	20%	80%

TABLE 1. Summary of the input conditions showing the percentage of nouns occurring with determiners overall and by syntactic role.

twelve nouns needed in the subsequent sentence-production task). Participants were asked to provide a name for each object as it appeared on a video monitor and were given as much time as they needed to respond. All responses were live scored and videotaped to enable checking by another coder.⁸

Participants who knew at least five of the vocabulary items progressed to the test of primary interest, the sentence-completion task. (As mentioned in the Participants section, forty-eight of the fifty-six participants passed this test. From here onward only their data are discussed and analyzed.) In this task, participants had to produce their own novel sentences, allowing us to examine their production of determiners. Participants saw a novel scene and heard the first word (i.e. the verb) of what would be the corresponding sentence. The verb was provided to constrain the relationship between the noun and its grammatical role in each sentence and to ensure that each noun produced would surface in all three roles in participants' productions. (Some scenes could be described in multiple ways, such that the subject and object could be reversed. Providing the verb forced participants to put the nouns into the intended roles.) Participants were asked to produce the complete sentence and were given as much time as they needed to provide an answer. For example, a participant sees a large stuffed frog toy and hears the word for 'big'. She should then say the sentence corresponding to 'big frog (DET)'.

There were twenty-four test sentences (twelve transitive, twelve intransitive), resulting in thirty-six possible NPs, and so thirty-six possible determiners. The test was designed such that each of the twelve nouns tested in the vocabulary test appeared once in each of the three possible grammatical roles (intransitive subject, transitive subject, and transitive object). Participants were tested on the transitive sentences first, followed by the intransitives. Participants were asked to indicate where a noun they could not remember went in the sentence (for instance, by saying 'X'), enabling the inclusion of incomplete responses in the analyses. Responses were videotaped and later transcribed and analyzed.

2.2 RESULTS.

VOCABULARY TEST. Mean performance on the vocabulary test was generally high at 9.6/12 (range 6–12, $SD = 1.88$). The means for the three conditions (-CV S=O = 9.00, $SD = 0.62$; +CV S>O = 9.65, $SD = 0.49$; +CV O>S = 10.11, $SD = 0.37$) were not significantly different from each other ($F(2,45) = 1.25$, $p = 0.296$). Differences between the input languages do not appear to have affected participants' abilities to learn the vocab-

⁸ There was actually a prior vocabulary test administered after participants watched the videotape in the fourth session. Participants were told that the test was designed to give them some idea of how they were doing up to this point—that it was for their own benefit and would not be analyzed. Participants were asked to provide a name for each of twelve objects as it appeared on a video monitor and were given as much time as they needed to respond. All responses were videotaped, but in accord with the instructions the results were not analyzed. The same twelve nouns were tested in both vocabulary tests, but in different orders.

ulary of the language. Thus, any differences in performance on the sentence-completion test are unlikely to be due to differences in knowledge of the languages.

SENTENCE-COMPLETION TEST. Figure 1 shows the percentage of nouns accompanied by the correct determiner overall and separately for subjects and objects, by condition. Looking first at the overall percentages, the effect of input language is significant ($F(2,45) = 14.65, p < 0.001$). Planned comparisons between input conditions with different overall mean determiner occurrence (60% vs. 40%) show that participants in the +CV O>S input condition produced significantly fewer determiners overall than participants in either of the other input conditions (–CV S=O: one-tailed $t(45) = -2.45, p = 0.036$; +CV S>O: one-tailed $t(45) = 5.41, p < 0.001$). Thus, participants appear to be sensitive to the overall probabilities of occurrence. However, this could simply be a by-product of sensitivity to the syntactic category-specific probabilities. Indeed, the difference between productions of the two groups exposed to determiners at an overall rate of 60% suggests that the category-specific probabilities affected determiner production such that overall rates differed. This difference is seen both in a significant t -test comparing mean overall production rates (two-tailed $t(45) = 2.4, p = 0.021$) and in the CIs around the means for the two groups, shown in Table 2, which overlap very little.⁹

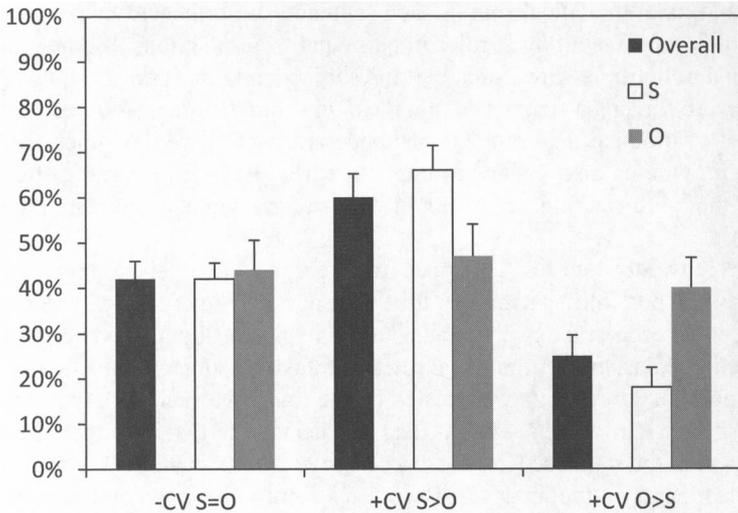


FIGURE 1. Percentage of nouns produced with correct determiners overall and separately for subject and object nouns, for the three input conditions. Error bars represent standard error.

The separate means for subject and object nouns (Fig. 1) demonstrate a great deal of sensitivity to the by-position probabilities. Although participants do not reproduce the input percentages associated with the two types precisely (see also Table 1), the overall patterns in all three groups match the patterns in the input. That is, participants produced more versus fewer determiners in line with their input. As we might expect if participants did learn the category-specific probabilities associated with their particular input language, an ANOVA with condition as a between-subjects factor and position as

⁹ This t -test is two tailed because I had no a priori predictions about direction of difference (since the difference was not anticipated). This is in contrast to the contrasts reported immediately above, where differences in a particular direction were anticipated and so tests were conducted accordingly.

a within-subjects factor revealed a significant main effect of condition ($F(2,45) = 8.36$, $p < 0.001$) but not syntactic position ($F(1,45) = 0.27$, $p = 0.61$), and a significant interaction between the two factors ($F(2,45) = 13.39$, $p < 0.001$). When syntactic position is analyzed separately for each condition, it is only significant (that is, the production of determiners was only significantly different for the two syntactic positions) for the two conditions in which participants heard different amounts of determiners in the two positions ($-CV S=O$: $F(1,11) = 0.19$, $p = 0.67$; $+CV S>O$: $F(1,16) = 8.82$, $p = 0.009$; $+CV O>S$: $F(1,18) = 16.27$, $p = 0.001$).

CORRECT DETERMINERS ONLY			
CONDITION	OVERALL	NOUN TYPE	
		SUBJECT	OBJECT
$-CV S=O$	33.76–51.14	34.32–49.72	30.20–58.66
$+CV S>O$	48.79–70.67	54.64–77.58	32.34–62.08
$+CV O>S$	15.56–34.80	8.66–27.17	26.00–53.61

ALL DETERMINERS			
CONDITION	OVERALL	NOUN TYPE	
		SUBJECT	OBJECT
$-CV S=O$	40.22–60.48	37.90–57.47	41.50–73.86
$+CV S>O$	50.82–72.89	56.21–80.60	33.97–63.73
$+CV O>S$	19.30–41.32	11.05–31.74	32.39–65.24

TABLE 2. 95% confidence intervals for the means: determiner production by noun type and input language condition.

Although the majority of the remainder of the productions involved the production of no determiner with the noun (mean = 53.4%, $SD = 24.48$), which was the other option contained in the input, occasionally participants produced the wrong determiner or some other phonological form (e.g. a blend of the two determiners, or adding a consonant to the end of a determiner) after the noun (mean = 4.81%, $SD = 8.64$).¹⁰ Importantly, the production of incorrect or other determiners did not differ by condition ($F(2,45) = 1.61$, $p = 0.211$), while the production of nouns unaccompanied by determiners did ($F(2,45) = 10.55$, $p < 0.001$), the latter result being exactly as we would expect given the previous analyses. Nevertheless, it is worthwhile to at least inspect the data that include any and all determiner-like forms produced by participants, that is, to look at the percentage of nouns produced with determiners of any kind, not just correct ones. As Figure 2 shows, the patterns are the same in this analysis. Although the specific percentages change somewhat, generally getting even closer to the input percentages, participants are still very much reproducing the patterns (i.e. relative frequencies) present in their input, if not the actual probabilities precisely. (Confidence intervals for the means for these data are also shown in Table 2.)

2.3. DISCUSSION. The results of experiment 1 show that adult learners can acquire patterns of variation conditioned on syntactic roles in a miniature artificial language. Participants who were exposed to a language in which determiners occurred with different frequencies for subject and object nouns used determiners accordingly in their own novel productions, in contrast to participants who heard, and used, determiners with equal frequencies across syntactic categories. Although participants did not match the input probabilities precisely, they did produce determiners according to the general patterns present in their input. This is very much in line with what we know about adult

¹⁰ These percentages are based on all productions; that is, the denominator is the number of correct nouns produced.

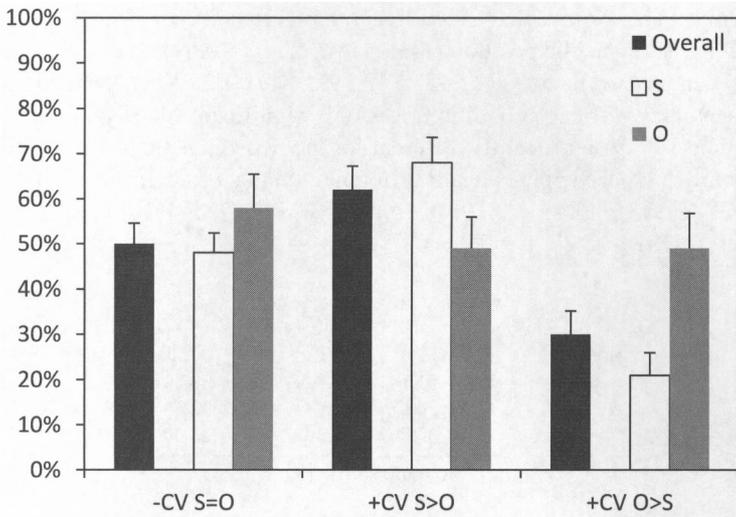


FIGURE 2. Percentage of nouns produced with any determiner overall and separately for subject and object nouns, for the three input conditions.

L2 learners' abilities to acquire sociolinguistic variation in natural languages (Regan 2013). It also replicates findings from previous experimental studies on unconditioned variation using a slightly different language (Hudson Kam & Newport 2005, 2009). Thus, this first experiment serves as an important (double-)check on the methodology.

There is also a potentially interesting pattern in the data. Participants' productions did not deviate from the input probabilities in a uniform way. In the condition where determiners were equally likely for subject and object nouns, learners produced determiners less often than they had heard them, importantly equally so for both subjects and objects, consistent with the results from previous studies (Hudson Kam & Newport 2005, 2009). In the other two input conditions, however, participants sometimes underproduced and at other times overproduced determiners. Participants in the +CV S>O condition underproduced determiners with subject nouns, but overproduced determiners with object nouns. Participants in the +CV O>S condition matched the probability of determiner occurrence for subject nouns almost perfectly, but underproduced determiners with object nouns. It is not entirely clear what caused these patterns of over- and underproduction. Underproduction has been seen before in similar studies and so was not unexpected in the present data. Overproduction has also been seen in previous studies with adults, but only when the variation in the input was quite complex, involving numerous possible determiner forms (Hudson Kam & Newport 2009). It is possible that learners in the current study found the patterns of variation complex enough to induce overproduction of the determiners. However, if complexity is the source of the overproduction, we would expect it to occur equally in both +CV conditions and for both subject and object nouns, but this is not the case. Another possibility is that participants were better able to learn the probabilities associated with subject nouns since subjects were twice as frequent as object nouns (due to their occurrence in both transitive and intransitive sentences) and that participants were not yet at asymptotic performance for object determiners. Whatever the case, the overall finding stands: participants' productions reflected the overall relative patterns of determiner occurrence in their input.

3. EXPERIMENT 2. Having established that adult learners can acquire conditioned probabilistic morphosyntactic variation in an experiment, and so validated the method,

I now turn to the question of whether child learners can as well. This is not a given; previous research using miniature artificial languages has shown that adult learners are often much better able to learn variation than children, who have a tendency to impose systematicity on variation in their input (Austin et al. 2006, Hudson Kam & Newport 2005, 2009; cf. Boyd & Goldberg 2012). But if conditioning is crucial for child learners to acquire variation, then we would expect learning outcomes to look different for children exposed to conditioned versus unconditioned variation. This was examined in a second experiment.

3.1 METHOD.

PARTICIPANTS. Participants were twenty-seven children aged five to seven (mean = 6;2).¹¹ An additional three children started the study but did not complete it. Participants were recruited and run at local preschool and child care facilities. Parental consent was acquired first, and then each child personally assented to participate. Children were given small stickers or toys daily for participating when the site allowed it.

STIMULI. The artificial language used in this study was exactly the same as that used with children in Hudson Kam & Newport 2005, 2009. It consisted of seventeen words: four verbs, twelve nouns, and one determiner. The words are a subset of those in experiment 1. Unlike the language used with adult learners in experiment 1, there is only one noun class and therefore one determiner. As in the first experiment, all aspects of the language other than determiner usage were consistent and regular. The overall percentage of NPs with determiners was 60% for both the conditioned- and unconditioned-variation languages. The lexicon was used in conjunction with a set of real objects and actions whose semantically possible combinations produce more sentences than were actually used in the exposure set, allowing us to reserve a variety of sentences for testing. As in experiment 1, word order was VSO, and determiners followed nouns.

PRESENTATION. The methods used with child participants differ from those used with adults in three major respects (but are the same as those used in Hudson Kam & Newport 2005, 2009). First, we used a live presentation method for the child participants. Second, we taught the children the nouns and verbs explicitly. We did not teach them anything about determiner usage, however. Finally, presentation (but not testing) typically occurred in groups of two or three. This allowed us to run numerous children at the same site within as short a time as possible. Importantly, previous research shows that differences between child and adult learners cannot be ascribed to these differences in methodology (see Hudson Kam & Newport 2005, 2009).

The exposure set consisted of twenty-four sentences, twelve intransitive and twelve transitive. Each of the twelve nouns appeared once in each syntactic position (intransitive subject, transitive subject, transitive object) in the set of exposure sentences. The intransitive sentences were split equally between the two intransitive verbs: six were 'fall' events, and six were 'move' events. The transitive sentences consisted of three 'inside-of' events and nine 'hit' events. This imbalance reflects the fact that there are more possible 'hit' events in the reference world than there are 'inside-of' events. As mentioned, we directly taught participants the vocabulary items, excluding the determiner. However, there was no explicit teaching of the grammatical aspects of the language.

There were six exposure sessions and a final test session, each of which lasted from ten to twenty minutes. The sessions were completed over nine days by all participants. On the first day the experimenter began by explaining to the child/children that she was

¹¹ The means for the two experimental conditions were both also 6;2.

going to teach them a new language called Sillyspeak. She said that they would learn some new words for things and some new ways to say things. Usually, the experimenter proceeded to chat for a few moments after this, explaining what it means to speak another language (see also McDaniel & Cairns 1996). To prevent the children from asking the experimenter questions about the language, she told them that she herself did not know the language and that she was going to learn it with them. Then the first actual exposure, consisting only of vocabulary training, began. This consisted of the experimenter running through the nouns and verbs on the vocabulary list four times as follows. Each run began with the four verbs. The experimenter said ‘if you want to say “hit” in Sillyspeak you say “flimm”’, then the same thing for the other three verbs. Participants were asked to repeat each new word after they heard it. After running through the four verbs, participants were then taught the nouns. On the first run through the nouns, participants were shown a toy and asked to name it in English. They were corrected if required, ensuring that they were encoding the intended meaning. The experimenter then told the participant(s) how to say the word in Sillyspeak. Importantly, the nouns were presented without determiners during vocabulary learning. This sequence was then repeated three times, each run having a different ordering within the verbs and nouns. That is, the verbs always came first, but the order in which the verbs/nouns occurred with respect to each other differed on each run through the vocabulary. The ordering was the same for all participants, regardless of condition.

The second session began with another pass through the vocabulary. The experimenter then said she was going to demonstrate how to ‘put words together’ (say sentences), which, at this point in the session, consisted of presenting the twelve intransitive sentences in the exposure set to the children. She showed the participants a scene and then said the corresponding sentence out loud. The sentence was read from a piece of paper on the experimenter’s lap. This served two purposes. First, it ensured that the experimenter would say the sentence correctly, especially with respect to the determiner usage. Second, it made the story about the experimenter not being knowledgeable about the language more plausible (why would she need to read the sentences if she actually knew the language?). As with the vocabulary, participants were asked to repeat each sentence after hearing it. The session continued with a second run through the vocabulary, followed by the twelve transitive sentences in the exposure set.

The third and fourth exposure sessions proceeded in exactly the same way as the second (but with different internal orderings of vocabulary). Sessions 5 and 6 were slightly different. Session 5 comprised one pass through the vocabulary, one pass through the intransitive sentences, one pass through the transitive sentences, and a second pass through the intransitive sentences. Session 6 was one pass through the vocabulary, one pass through the transitive sentences, one through the intransitive sentences, and a final pass through the transitive sentences. This design resulted in twelve passes through the vocabulary and six passes through each kind of sentence.

Occasionally words or sentences were repeated by the experimenter. This occurred only if it was very clear that the child had not been paying attention to the initial presentation. To encourage continued engagement, the children were allowed to help enact the sentences after the initial sentence exposure (that is, in sessions 3–6).

EXPERIMENTAL MANIPULATION. There were two conditions, and so two languages, in this experiment, conditioned and unconditioned variation. As in the previous experiment, the variation was in the determiners. In the unconditioned variation (–CV) language, 60% of the nouns occurred with a determiner, and 40% occurred with no determiner. These percentages were the same for subjects and objects (and for intransitive and tran-

sitive subjects), as well as for each session. In the conditioned variation language (+CV), 75% of the nouns occurring as subjects (transitive or intransitive) and 25% of the nouns occurring as objects cooccurred with the determiner. These percentages were the same for all input sessions. As in experiment 1, all other aspects of the language were consistent. All other aspects of exposure were the same across conditions.

TESTING. The final (seventh) session consisted of testing. In contrast to exposure, testing was conducted separately for each child. All responses were written down live by the observer (a second experimenter) and videotaped. (Five participants' production tasks were also transcribed from the video for reliability assessment by a researcher who was blind to the child's input condition. There were no disagreements between the two transcriptions.)

First, a set of objects was selected for use in the production task. The objects corresponding to the nouns used in the exposure sentences were shown to participants, and participants were asked to name as many of these objects as they could until at least five objects had been named. If the child could not name five objects, recognition was tested until at least five objects were named and/or recognized.¹² Recognition was tested by saying a noun in Sillyspeak and asking the child to select the appropriate toy. Children also often offered verbs that they knew at this stage, although they were not asked to do so.

The child's set of named or recognized nouns was then used in the production task. Using the objects the participant had named and/or recognized, the experimenter prompted the participants to produce novel sentences in which the test nouns occurred as both subjects and objects, in the following way. First, the relevant object or objects were shown to the child, who was reminded of the Sillyspeak label, usually by having the child him/herself repeat the name. Next, the child was shown or asked to enact a novel event or scene, and was told what the sentence should mean in English and what the first word of the corresponding Sillyspeak sentence was. For example, if the participant had known the word for 'bear', the experimenter would show the child the toy bear, ask the child to say the name, and then wind the bear up (making it move), put it down in front of the participant, and say 'OK, I want you to tell me how to say "the bear moves" in Sillyspeak. The first word would be "mirt", right?'. If participants had difficulty, they were reminded that they had learned how to say similar things in Sillyspeak, but they were not reminded how to say any exposure set sentences in Sillyspeak. Testing ended when all nouns had been used at least once in all three positions, or when the child refused to participate further.¹³

Sentences were not tested in any particular order, although the first few test sentences were often intransitive, allowing the children to gain confidence with the task before attempting the longer (transitive) sentences.

3.2 RESULTS. Because vocabulary elicitation was typically stopped when the child had named enough nouns (i.e. five), reporting performance on any kind of vocabulary

¹² This method, slightly different from that used in Hudson Kam & Newport 2005, 2009, meant that no children who participated were excluded from the production test and that all children provided a similar amount of production data, provided they could actually produce sentences given the additional help. The former was the primary motivation, given the difficulty of finding child participants for this kind of study.

¹³ Occasionally a child would ask to stop before all nouns had been used in all three positions, usually with the result that one noun had not been used in one syntactic position, still leaving us with adequate data in all three syntactic positions. Only one child insisted on stopping before any transitive sentences had been produced. Her data are not included in any of the analyses of transitive sentences or comparisons of subjects and objects.

measure is misleading for this second study. (I say ‘typically’ because sometimes a child would spontaneously continue to say more nouns.) Thus, results are discussed only for the sentence-production task.

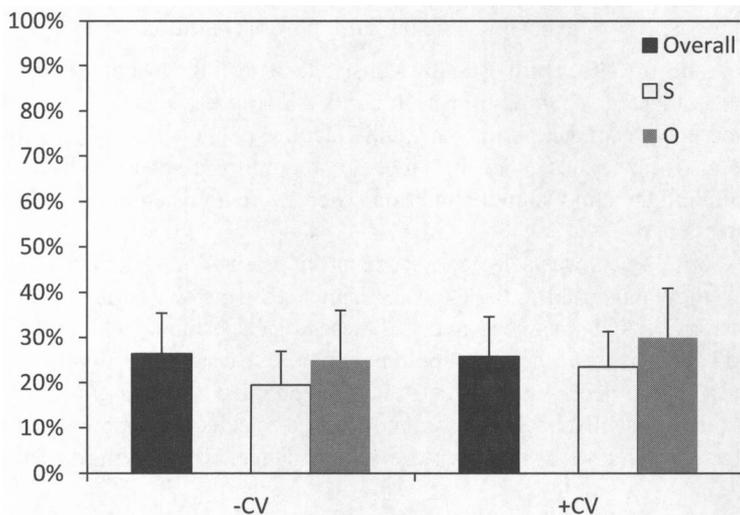


FIGURE 3. Percentage of nouns produced with a determiner overall and separately for subject and object nouns, for children exposed to unconditioned (–CV) and conditioned (+CV) variation.

Figure 3 shows the percentage of nouns produced with determiners overall, and separately for subjects and objects, by condition.¹⁴ Unlike the adult data presented earlier, the children in both input conditions look remarkably similar, and children in neither input condition appear to be reproducing the probabilities of determiner occurrence found in their input. An ANOVA with condition and role (subject vs. object) as factors shows no significant effect of condition ($F(1,24) = 0.109, p = 0.744$) nor interaction between condition and role ($F(1,24) = 0.016, p = 0.9$), either or both of which would be expected to be significant if the children were matching their input languages. Instead, both groups appear to be producing slightly more determiners with objects than with subjects, something we did not anticipate, and this difference approaches the standard criterion for significance ($F(1,24) = 4.08, p = 0.055$), with a modest effect size (partial eta-squared = 0.145). Moreover, all of the children appear to be radically underproducing determiners: recall that the overall percentage of determiners heard by children in both conditions was 60%. Indeed, only one of the bars—object nouns for the +CV condition—is even within 10% of the input percentage (it is actually within 5%), and given that it is the only one it looks suspiciously like a coincidence.¹⁵

However, this data pattern is not all that surprising if we consider that children are very likely to impose rules on inconsistent patterns in a language and that different children might impose different patterns (Hudson Kam & Newport 2005, 2009). Hudson Kam and Newport (2005) examined individual speakers’ productions for evidence of three different kinds of systematic determiner usage: consistent use, consistent nonuse, and consistent other (an idiosyncratic but consistent pattern involving both use and

¹⁴ Unlike the adults, whenever a child produced a determiner it was the correct one. This makes some sense, as the children only heard one possible determiner.

¹⁵ Note that it is much less a deviation from the input if we consider the fact that the children did not hear determiners with the nouns during the vocabulary training.

nonuse). The data from the present study were examined using the same classification system.¹⁶ Table 3 shows the number of children who used determiners systematically and unsystematically in their productions by input condition. While many children exposed to conditioned variation did still produce determiners in a systematic way, the proportion who do so is quite different as compared to children exposed to unconditioned variation. A chi-squared test comparing the distribution in the +CV condition (observed values) to that in the -CV condition (expected values) is significant ($\chi^2(1, N = 15) = 9.72, p = 0.002$). There are also differences in the ways that the two groups of children were systematic. Three of the ten systematic producers in the -CV condition either used determiners all the time ($N = 2$) or imposed their own idiosyncratic pattern (one, who used determiners only and always with one particular noun), with the rest being consistent nonusers. No children in the +CV condition used determiners in a systematic way; the eight children classified as systematic were all nonusers.¹⁷ Although this is a qualitative assessment, it does again suggest that something different is going on in the two conditions.

CONDITION	CLASSIFICATION	
	SYSTEMATIC	UNSYSTEMATIC
-CV	83.3% (10)	16.7% (2)
+CV	53.3% (8)	46.7% (7)

TABLE 3. Percentage of child participants who are systematic and unsystematic in their determiner production by input condition (raw numbers in parentheses).

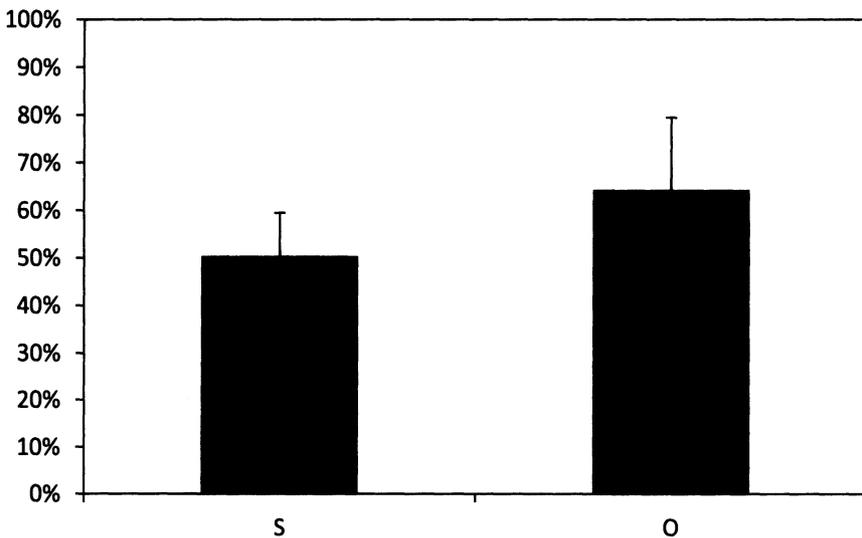


FIGURE 4. Percentage of subject and object nouns produced with a determiner by unsystematic producers in the +CV condition.

But what of the children who were unsystematic in their productions? The hypothesis that conditioned variation leads to learning predicts not only that fewer children will impose regularity on variation, but crucially also that children who do not impose regularity will actually learn the variation, or at least the general tendencies or relative fre-

¹⁶ As in Hudson Kam & Newport 2005, each speaker was allowed one deviation from the pattern. For instance, a child who produced fifteen noun tokens but only one determiner would be classified as a systematic nonuser.

¹⁷ The numbers are too small for a statistical analysis, especially given the presence of zeros in some cells.

quencies, as was the case with the adults in the previous experiment. Figure 4 shows the percentage of nouns produced with determiners for subject and object nouns for children in the +CV condition who were classified as unsystematic ($N = 7$). It is clear from the figure that the children are not reproducing the input percentages. Moreover, they are not reproducing the general pattern of more determiners with subjects than objects; they deviate so greatly from the input probabilities that the confidence intervals do not include the target percentages (S input percentage = 75%, CI = 28.23–72.58%; O input percentage = 25%, CI = 27.13–101.14%).

Although the +CV language was created with probabilities conditioned by the syntactic positions of the nouns, there are other categories over which learners could potentially track probabilities and still be reproducing their input. Indeed, when the productions of these seven children were reexamined, the pattern that emerged suggests that they were tracking the probabilities over sentence types, that is, transitive and intransitive, as shown in Figure 5. This pattern holds for six of the seven unsystematic producers. Although the specific probabilities differed for each child, they all produced determiners at least thirty percentage points more often in intransitive sentences than in transitive sentences. Only one child showed a different pattern in her speech: she produced more determiners in transitive than intransitive sentences. Importantly, the transitive < intransitive pattern is not apparent in the unsystematic producers who were exposed to the –CV language (in fact, the opposite is true for those children), so it does not seem to be something generally true of children who are classified as unsystematic producers. This same pattern is apparent in the adults in the +CV S>O condition in experiment 1. However, this is a consequence of their having matched the patterns with respect to subjects and objects present in their input, something not true of the children.

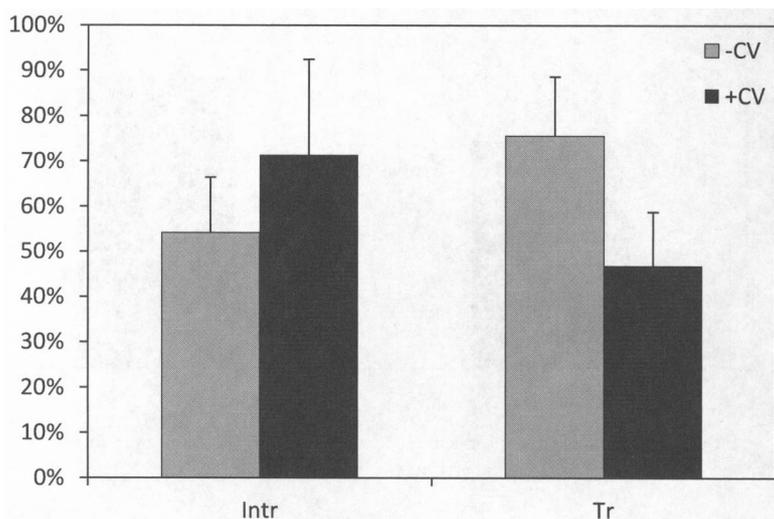


FIGURE 5. Percentage of nouns produced with a determiner in transitive and intransitive sentences by input condition (unsystematic producers only).

4. DISCUSSION AND CONCLUSION. This study was designed to assess the hypothesis that the difference between sociolinguistic variation, which is learnable, and variation that gets changed by learners, particularly child learners, is conditioning. By conditioning I mean that the chance of something occurring differs in different linguistic contexts, or put another way, that contexts condition the variation, not that the occurrence

is deterministic. This hypothesis was explored with respect to variable determiner occurrence using noun role or position, namely subject versus object, as the relevant linguistic contexts.

Experiment 1 found that adult learners were fairly good at learning the patterns of variation. This was the case both when the determiners occurred more frequently with subjects than with objects and when they occurred more frequently with objects than with subjects. Although the adults did not learn the exact probabilities associated with determiner occurrence in all cases, this is not terribly surprising. Compared to real language-learning situations, any artificial language experiment is going to be short. Thus, we might expect learning to be less than perfect. Moreover, even in cases of variation in real languages there is individual variation in the use of the variable forms (see e.g. the individual-level data in Smith et al. 2007). What is consistent within a language community are the patterns, more specifically the degree of impact of any contextual variable on production, not the exact rates. For example, a certain context could make *-t/-d* deletion twice as likely for all speakers, while at the same time the specific rates of production (or deletion) may vary between speakers.¹⁸ Most important in the data presented here is the lack of regularization by the adult learners. Hudson Kam and Newport (2009) found that adult learners began to regularize over similar inconsistently occurring determiners as the number of probabilities they had to acquire increased. Specifically, when the language included two probabilistically occurring determiners, each associated with a different noun class, adults could learn the probabilities associated with each, but when the language included two additional probabilistically occurring determiners that were not restricted to one or the other noun class, adult learners began to regularize the language. In that study, adding complexity to the variation led to regularization. Importantly, in a language with even more determiners (sixteen), which occurred in a completely predictable fashion, there was no regularization. Thus, even small increases in unpredictable variation pushed the learners in their study to regularize, while much larger increases in predictable variation did not. A priori, it seems as if conditioned but still unpredictable (i.e. not deterministic) variation might count as increasing the complexity of the variation. Yet, it did not lead to regularization by the adult learners in experiment 1 of this article.

The results for the children in experiment 2 were somewhat different. Group data showed that learning was no different between the children exposed to conditioned and unconditioned variation. However, although many children exposed to the language with conditioned variation were quite consistent in their productions of determiners, the proportion of children who were unsystematic was substantially greater in the +CV condition than in the -CV condition, suggesting that the nature of the variation present in the language did have an effect on learning. When the data from just those unsystematic +CV children were examined more closely, however, it was clear that they were not reproducing the subject/object asymmetry present in their input language. Instead, it appears that they were approaching the determiner asymmetries from the perspective of sentence types, in particular, transitive versus intransitive sentences.

This last result, while tentative given the small number of participants, raises the question of why transitive/intransitive and not subject/object, the answer to which has potentially broad implications for our theories of acquisition. One possibility is that it is not transitive versus intransitive, but something else. As can be inferred from a comparison of Figs. 3 and 5, the children were actually more likely to produce the determiner with

¹⁸ There are, of course, limits to the specific rates that could occur for this to be true.

intransitive subjects and transitive objects than with transitive subjects. For the seven children who were unsystematic producers, the means (and standard deviations) for the three positions were as follows: intransitive subject = 71.4% (32.4), transitive subject = 28.33% (35.9), transitive object = 64.29% (40.2). The standard deviations are large, due to differences between the children in their absolute rates combined with a small sample size. The point here is the pattern, which seems quite like an ergative-absolutive pattern. Interestingly, Goldin-Meadow and her colleagues have found evidence that the ergative-absolutive pattern is what emerges in young homesigners (children inventing their own gestural communication systems) world-wide (Goldin-Meadow & Mylander 1998, Goldin-Meadow et al. 2000, Schulman et al. 2001). Thus, it is possible that children with no input (the homesigners) or with input that contains no consistent evidence of a pattern will impose an ergative-absolutive system on the language. However, this explanation for the behavior of the children in experiment 2 is inconsistent with the behavior of the children in the -CV condition and with the children in Hudson Kam & Newport 2005, 2009. If the lack of a consistent pattern leads to the imposition of an ergative pattern, then we should see all children who are exposed to inconsistent variation imposing ergative patterns, which is not the case. Moreover, it ignores the fact that there is a consistent way of marking subject and object in the languages, namely position in the sentence. (Although one could argue that children only impose ergative-absolutive patterns on the part of the grammar that is inconsistent, and in so doing create a mixed system, something that is very common crosslinguistically; see e.g. Croft 1990.)

Another possibility, which is less interesting from the perspective of theories of acquisition, is that it is a production issue: when faced with something difficult to produce (such as the variable determiners in this study), retrieval is facilitated by (i) facing less competition from other to-be-produced words, so occurrence in shorter sentences, and (ii) having more time to retrieve the item, so when the thing to be produced is located at the end of a sentence.¹⁹ (The latter depends on speakers starting to speak before an utterance is fully planned, something for which there is evidence; see Brown-Schmidt & Konopka 2008.) Both of these factors converge in the intransitive sentences to facilitate determiner production. In the transitive sentences, these conspire to produce different likelihoods of production in the subject and object noun phrases; specifically, the determiners will be less likely in the subject than the object noun phrases.²⁰ On this story, determiner production in the intransitive sentences and with the objects of the transitives evinces the learners' true knowledge of the underlying probabilities, while production with the subjects of the transitive underestimates it. Note that on this story there is not necessarily any underlying asymmetry. This is an intriguing possibility for how ergative languages might arise, but only in the case of languages with VS ordering, since for SV ordering the subject comes earlier.²¹ However, this still does not explain why this production asymmetry occurs only in the +CV condition.

¹⁹ These two factors have not been independently demonstrated, to my knowledge. But they follow from what we know about speech errors (e.g. word exchange errors where words from the same grammatical category switch places (Garrett 1980) and higher rates of disfluency early in the sentence as compared to later, reflecting greater difficulty early as opposed to late (Clark & Fox Tree 2002)).

²⁰ I am being intentionally vague about exactly how this might happen, in part because there are multiple routes to the same outcome. First, the determiners could be directly selected on their own. Second, they could be selected by the nouns (see Myers-Scotton & Jake 2000). Either could produce the effect discussed in the text.

²¹ Although see Jaeger & Norcliffe 2009 for a discussion of whether S and V form a unitary constituent from the perspective of production such that they must be fully planned before production can begin.

A third possibility is that subject and object are not linguistic primitives that children can track probabilities over. The idea that subject and object are not primitives is not a new one; in fact, it has been debated for many years, with different theories (and theorists) taking different positions (e.g. Chomsky 1965, Pullum 1980, Van Valin & LaPolla 1997). However, if we take children's lack of probability matching over the categories subject and object as evidence for the lack of a primitive, then we must also say that noun is not a primitive, since children seem incapable of tracking probabilities over the category as a whole, evidenced by performance in the -CV condition, as well as in previous studies (Hudson Kam & Newport 2005, 2009). This is consistent with the views of emergentists such as Tomasello (see e.g. Tomasello 2003), but it then raises the question of how something like intransitive sentence could be a primitive while noun is not. Moreover, it is almost certainly the case that children of the age tested in this study do have the category noun in English (if it is emergent), so why is it not transferred into the new language as they learn it, as presumably happens for adults? It is possible that for children, whose linguistic knowledge is more fragile and less explicit than that of adult language learners (the adults in this study at least, since all had some explicit knowledge of language categories acquired in language classes), the category noun must be constructed anew each time, in contrast to the transitive-intransitive distinction, which appears to be quite basic for child language learners (see Fisher et al. 2010 for a review; cf. Tomasello 2000). Alternatively, noun might simply be a less salient category than transitivity in this particular learning context. All of the sentences clearly involve one or two arguments in the presence of a vignette with one or two things doing or being something. There is not a lot of extraneous auditory or visual material to complicate interpretation. Thus, the distinction between transitive and intransitive sentences is quite salient and could be based on a conceptual rather than primitively linguistic distinction. (This raises the question of what could lead to noun being the more salient category, but that is beyond the scope of this article.) Whatever the source of this asymmetry, it is very interesting from the perspective of understanding what child language learners bring to the table. Interestingly, there is a precedent in the sociolinguistic literature for children to be initially tracking variation over the wrong contexts/variables. Labov (1989) reports that children initially treat derivational *-t/-ds* as if they were monomorphemic, with the result being that the probabilities of deletion are incorrect when compared to those of adults but make sense within the child's own linguistic system. Thus, this mislearning fits with what we already know about possible patterns in the acquisition of sociolinguistic variation.

To this point, variation has been discussed, at least implicitly, as either stable (learnable) or unstable (immediately regularized by learners during the process of acquisition), but there is a complication with these definitions, namely variation that is indicative of a change in progress. Such variation has conditioning factors and so by the account outlined in the current article should be learned, yet it is not—at least, not perfectly. The only speculation I can offer that is consistent with the current hypothesis and data is that this imperfect learning may be the norm in all learning of variation, as it was in both experiments presented here: learners always only approximate their input; sometimes the deviation is large, other times quite small. Interestingly, much of the work in variationist sociolinguistics is concerned not with establishing the bounds on stable variation, but rather with understanding ongoing change in languages, using apparent time as a window into real time. Indeed, variation and change are tightly coupled conceptually, as exemplified by titles such as *The handbook of language variation and*

change (Chambers et al. 2002).²² Thus, it may be the case that the type of learning outcomes observed in the present work are actually quite indicative of the learning of conditioned variation, that is, small changes imposed on the system of variation that are consistent with the existing patterns, in contrast to the outcomes that occur in the rare instances when learners are faced with unconditioned variation, namely, major changes that impose different characteristics on the language. I would not want to claim that the DEGREE of difference between the input and output in the +CV conditions in this experiment is exactly characteristic of real language learning. In cases of real language learning, learners have much more exposure to the system to be acquired. Although children do take time to master it (Labov 1989, Roberts 1997, Smith et al. 2007) and might often (or always) pass through a phase of finding the wrong generalizations about conditioning before arriving at the right ones, as was case in the present study (see also Labov 1989), they eventually must at least get very close in cases of stable variation. However, I am suggesting that some small degree of difference might be typical in situations where learners are faced with such complexity.

Another possibility that is consistent with recent work showing that children initially match their parental input quite well early on for variables involved in a change (Rutter 2014), just as they do when faced with stable variation (Smith et al. 2007, Smith et al. 2013), is that conditioning must be consistent across speakers and real time at the individual level (i.e. a learner's lifetime) for it to be stable. On this story, conditioned variation that is indicative of a change in progress will initially be learned veridically, but as children enter the wider speech community and hear different variation (when they enter school, for instance, and are exposed to other speakers with different patterns), they end up changing it. Just how much they adjust their speech patterns might depend on how variable the variation is (i.e. how different the conditioning is in different speakers), as well as how integrated they are in social networks outside the home (see Kerswill & Williams 2000). This is different from the situation where the home variety is distinct from a shared community variety, in which case, if children are young enough upon exposure to the new variety, they will learn it, not change it (Chambers 1992, Payne 1980). Again, this suggests that as long as conditioning is consistent within a context, it will be learned.

This leads into an important point. It might seem that the present work provides little insight beyond what we already know about the acquisition of sociolinguistic variation on the basis of studies of naturalistic acquisition. That is, we already know that it is learnable, and the present study merely confirms this, while raising several interesting questions. However, the point of the experiments reported here is not to show that sociolinguistic variation is learnable, but rather to test one hypothesis about WHY it is learnable in comparison to sometimes seemingly simpler variation that is NOT learned. This is not something that can be discerned by examining naturalistic situations. To really tease apart possibilities, which, crucially, often come out of an understanding of the real-world data, more controlled studies are needed. Artificial language experiments afford researchers that level of control. Importantly, artificial languages have been shown to rely on the same neural structures as real language (see e.g. Finn et al. 2013), providing confidence that we are investigating language learning and not some other cognitive

²² Interestingly, several recent authors discussing variation and change cite the work by Hudson Kam and Newport in their discussions of how and why change might happen (e.g. Lightfoot 2010, Troutman et al. 2008). But Hudson Kam and Newport's findings do not necessarily predict the small incremental changes seen in cases of normal language change, as opposed to the kinds of changes (namely creole emergence) that their work was designed to investigate.

ability and therefore that findings from artificial language experiments are relevant for understanding real language acquisition. Thus, artificial languages are an important tool available for answering just the kinds of questions these data raise. Ideally, data from the real world and from artificial language experiments complement each other, with the latter first confirming what we see 'in the wild' and then allowing us to go beyond. That is the case with the present work; the learning seen in the adults and children mirrored what we see in natural language acquisition, while suggesting why learners exposed to different kinds of variation sometimes learn it but other times do not. Moreover, it establishes the validity of the methodology for testing ideas about the acquisition of variation.

Two variables suggest themselves for further study on the basis of the present results and how they compare to other evidence, both of which have been hinted at in this discussion. The first, which was discussed only a few paragraphs above, is consistency through time, and how it might be related to how slow, incremental change progresses through a speech community. We know that where there is a change in progress, individuals continue to change as they age, although how much they change appears to depend on how advanced the change is (Tagliamonte & D'Arcy 2009). Exactly why we see this 'change in the rate of change' is, however, unknown.²³ The second is consistency across speakers. In the introduction I mentioned that conditioning variables (or the lack thereof) are not the only thing that varies between situations where variation had previously been shown to lead to change versus learning; multiple speakers who produced variation in the same way was also a feature of situations where learning had been observed. The results of this study demonstrate that a learner does not require multiple speakers using variation in the same way to learn it. However, that does not mean that interspeaker variation has no effect on learning. Indeed, the naturalistic data suggest that it may be very important. There is evidence that the variation in adult L2 speakers' productions, even their ostensible errors, can often be described in terms of (probabilistic) conditioning contextual variables, much like sociolinguistic variation (e.g. Wolfram 1985). However, Simon's case suggests that this kind of L2 variation is not learned by children. Examining this question at a more micro level, albeit using a very limited sample, Hudson (2000) showed that Simon deviated most from his parents' productions when they were most different from each other. Work by Ganuza (2008) also bears on this. She examined the production of syntactic variation (in particular, inversion and noninversion) in adolescents in a multilingual and multiethnic setting in Sweden who presumably learned from their L2-speaking parents (and other community members). She found that the adolescents' productions were sensitive to only some of the factors known to affect adult L2 speech. That is, they did not acquire the variation as it had occurred in their input. Thus, it seems that conditioning does not always lead to veridical acquisition, even when the factors involved in the conditioning are the same across speakers. For the variation to be learned when there are multiple speakers providing the input, the nature and degree of influence of the factors may also need to be shared, or consistent, across speakers. While difficult to investigate in naturalistic data, both of these are amenable to investigation using experimental techniques such as those employed here.

²³ That is not to say that this topic has not received attention. Tagliamonte and D'Arcy (2009), for instance, provide a mathematical model that accommodates both the apparent- and real-time data such that the 'change in change' is explained, in contrast to previous models of change. However, they do not provide any cognitive or learning-based explanation for it. They talk about 'acquiring increments' and 'accruing more increments' as leading to more change, but it is not clear what this means from a cognitive or learning perspective.

In sum, the evidence presented here suggests that conditioning factors do make variation more learnable for children, in contrast to unconditioned variation, which has been shown in several studies to be changed by learners (Hudson Kam & Newport 2005, 2009, Singleton & Newport 2004). Notably, the conditioned variation in the present studies was more complex than the unconditioned variation, in that there were multiple probabilities or rates of determiner occurrence to be learned rather than just one, but this did not lead to regularization by the adult learners as had other forms of complexity in previous studies (Hudson Kam & Newport 2009) and actually led to fewer children producing the language in a systematic way. Moreover, those children who did regularize were ALL nonusers, something not true of the regularizers in the –CV language condition or in previous studies showing children’s propensity to regularize variation, where there were a variety of different kinds of systematic patterns imposed, nonuse being just one of them (Hudson Kam & Newport 2005, 2009). Although the specific probabilities produced by the adults and the children in the +CV conditions were not exactly the same as those present in their input, for the adults the patterns were consistent with their input, and for the children the patterns in the productions seem to reveal something quite interesting about the categories that they attached probabilities to. Presumably, with more input the children would have learned the appropriate factors associated with the variation, but in the short term, the categories they conditioned over were not those used to construct the input stimuli. Note, however, that these children produced a different conditioned system, not a systematic way of using determiners, as has been common in previous experimental work. This is interesting because it supports the idea common in sociolinguistics that conditioned variation is a normal, indeed fundamental, part of language. Thus, the present work not only provides an explanation for the apparent contradiction between the existence of stable sociolinguistic variation and the regularization of variation that has been documented in some circumstances, but, importantly, it also provides some insights into the nature of language learning.

REFERENCES

- ADAMSON, HUGH D. 1988. *Variation theory and second language acquisition*. Washington, DC: Georgetown University Press.
- ARNON, INBAL, and MICHAEL RAMSCAR. 2012. Granularity and the acquisition of grammatical gender: How order-of-acquisition affects what gets learned. *Cognition* 122.292–305. DOI: 10.1016/j.cognition.2011.10.009.
- ASLIN, RICHARD N.; JENNY R. SAFFRAN; and ELISSA L. NEWPORT. 1998. Computation of conditional probability statistics by 8-month-old infants. *Psychological Science* 9.321–24. DOI: 10.1111/1467-9280.00063.
- AUGER, JULIE. 1998. Le redoublement des sujets en français informel québécois: Une approche variationniste. *Canadian Journal of Linguistics* 43.37–63.
- AUSTIN, ALISON C.; ELISSA L. NEWPORT; and ELIZABETH WONNACOTT. 2006. Predictable versus unpredictable variation: Regularization in adult and child learners. Paper presented at the Boston University Conference on Child Language Development (BUCLD), Boston, MA.
- BAYLEY, ROBERT, and DENNIS R. PRESTON (eds.) 1996. *Second language acquisition and linguistic variation*. Amsterdam: John Benjamins.
- BOYD, JEREMY K., and ADELE E. GOLDBERG. 2012. Young children fail to fully generalize a novel argument structure construction when exposed to the same input as older learners. *Journal of Child Language* 39.457–81. DOI: 10.1017/S0305000911000079.
- BROWN-SCHMIDT, SARAH, and AGNIESZKA E. KONOPKA. 2008. Little houses and casas pequeñas: Message formulation and syntactic form in unscripted speech with speakers of English and Spanish. *Cognition* 109.274–80. DOI: 10.1016/j.cognition.2008.07.011.
- BUCHSTALLER, ISABELLE. 2009. The quantitative analysis of morphosyntactic variation: Constructing and quantifying the denominator. *Language and Linguistics Compass* 3. 1010–33. DOI: 10.1111/j.1749-818X.2009.00142.x.

- CAIN, JACQUELIN; MARCIA WEBER-OLSEN; and ROSSLYN SMITH. 1987. Acquisition strategies in a first and second language: Are they the same? *Journal of Child Language* 14.333–52. DOI: 10.1017/S0305000900012964.
- CHAMBERS, JACK K. 1992. Dialect acquisition. *Language* 68.673–705. DOI: 10.2307/416850.
- CHAMBERS, JACK K.; PETER TRUDGILL; and NATALIE SCHILLING-ESTES (eds.) 2002. *The handbook of language variation and change*. Malden, MA: Blackwell.
- CHOMSKY, NOAM. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- CLARK, HERBERT H., and JEAN E. FOX TREE. 2002. Using *uh* and *um* in spontaneous speaking. *Cognition* 84.73–111. DOI: 10.1016/S0010-0277(02)00017-3.
- CROFT, WILLIAM. 1990. *Typology and universals*. Cambridge: Cambridge University Press.
- DÍAZ-CAMPOS, MANUEL. 2005. The emergence of adult-like command of sociolinguistic variables: A study of consonant weakening in Spanish-speaking children. *Selected proceedings of the 6th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*, ed. by David Eddington, 56–65. Somerville, MA: Cascadilla.
- DRYER, MATTHEW S. 2005. Order of adposition and noun phrase. *The world atlas of language structures*, ed. by Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie, 346–47. Oxford: Oxford University Press.
- ESTES, WILLIAM K. 1964. Probability learning. *Categories of human learning*, ed. by Arthur W. Melton, 89–128. New York: Academic Press.
- ESTES, WILLIAM K. 1976. The cognitive side of probability learning. *Psychological Review* 83.37–64. DOI: 10.1037/0033-295X.83.1.37.
- EVANS, JULIA L.; JENNY R. SAFFRAN; and KATHRYN ROBE-TORRES. 2009. Statistical learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research* 52.321–35. DOI: 10.1044/1092-4388(2009/07-0189).
- FINN, AMY S., and CARLA L. HUDSON KAM. 2008. The curse of knowledge: First language knowledge impairs adult learners' use of novel statistics for word segmentation. *Cognition* 108.477–99. DOI: 10.1016/j.cognition.2008.04.002.
- FINN, AMY S.; CARLA L. HUDSON KAM; MARC ETTLINGER; JASON VYTLACIL; and MARK D'ESPOSITO. 2013. Learning language with the wrong neural scaffolding: The cost of neural commitment to sounds. *Frontiers in Systems Neuroscience* 7:85. DOI: 10.3389/fnsys.2013.00085.
- FISER, JÓZSEF, and RICHARD N. ASLIN. 2002. Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences* 99.15822–26. DOI: 10.1073/pnas.232472899.
- FISHER, CYNTHIA; YAEL GERTNER; ROSE M. SCOTT; and SYLVIA YUAN. 2010. Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science* 1.143–49. DOI: 10.1002/wcs.17.
- FOULKES, PAUL. 2010. Exploring social-indexical variation: A long past but a short history. *Laboratory Phonology* 1.5–39. DOI: 10.1515/labphon.2010.003.
- FOULKES, PAUL; GERARD DOCHERTY; and DOMINIC WATT. 2005. Phonological variation in child-directed speech. *Language* 81.177–206. Online: <http://www.jstor.org/stable/4489857>.
- GANUZA, NATALIA. 2008. *Syntactic variation in the Swedish of adolescents in multilingual urban settings: Subject-verb order in declaratives, questions and subordinate clauses*. Stockholm: Centrum för tvåspråkighetsforskning dissertation.
- GARNSEY, SUSAN M.; NEAL J. PEARLMUTTER; ELIZABETH MYERS; and MELANIE A. LOTOCKY. 1997. The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language* 37.58–93. DOI: 10.1006/jmla.1997.2512.
- GARRETT, MERRILL F. 1980. Levels of processing in sentence production. *Language production, vol. 1: Speech and talk*, ed. by Brian L. Butterworth, 177–220. London: Academic Press.
- GOLDIN-MEADOW, SUSAN, and CAROLYN MYLANDER. 1998. Spontaneous sign systems created by deaf children in two cultures. *Nature* 391.279–81. DOI: 10.1038/34646.
- GOLDIN-MEADOW, SUSAN; ELIF YALABIK; and LISA GERSHKOFF-STOWE. 2000. The resilience of ergative structure in language created by children and by adults. *Proceedings of the Boston University Conference on Language Development (BUCLD)* 24.1. 343–53.

- GRAF ESTES, KATHERINE; JULIA L. EVANS; MARTHA W. ALIBALI; and JENNY R. SAFFRAN. 2007. Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science* 18.254–60. DOI: 10.1111/j.1467-9280.2007.01885.x.
- GREENBERG, JOSEPH H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. *Universals of language*, ed. by Joseph H. Greenberg, 92–96. Cambridge, MA: MIT Press.
- GUY, GREGORY R. 1980. Variation in the group and the individual: The case of final stop deletion. *Locating language in time and space*, ed. by William Labov, 1–36. New York: Academic Press.
- HAUSER, MARK D.; ELISSA L. NEWPORT; and RICHARD N. ASLIN. 2001. Segmentation of the speech stream in a nonhuman primate: Statistical learning in cotton-top tamarins. *Cognition* 78.B53–B64. DOI: 10.1016/S0010-0277(00)00132-3.
- HUDSON, CARLA L. 2000. Variation: When can it be learned and when does it lead to language change? Paper presented at the meeting of the Society for Pidgin and Creole Linguistics, Chicago, IL.
- HUDSON, CARLA L. 2002. *Pidgin, creoles, and learners: How adults and children contribute to the formation of creole languages*. Rochester: University of Rochester dissertation.
- HUDSON KAM, CARLA L. 2009. More than words: Adults learn probabilities over categories and relationships between them. *Language Learning and Development* 5.115–45. DOI: 10.1080/15475440902739962.
- 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* 1.151–95. DOI: 10.1080/15475441.2005.9684215.
- HUDSON KAM, CARLA L., and ELISSA L. NEWPORT. 2009. Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology* 59.30–66. DOI: 10.1016/j.cogpsych.2009.01.001.
- HUNT, RUSKIN H., and RICHARD N. ASLIN. 2001. Statistical learning in a serial reaction time task: Access to separable statistical cues by individual learners. *Journal of Experimental Psychology: General* 130.658–80. DOI: 10.1037/0096-3445.130.4.658.
- JAEGER, T. FLORIAN, and ELISABETH J. NORCLIFFE. 2009. The cross-linguistic study of sentence production. *Language and Linguistics Compass* 3.866–87. DOI: 10.1111/j.1749-818X.2009.00147.x.
- JOHNSON, JACQUELINE S.; KENNETH D. SHENKMAN; ELISSA L. NEWPORT; and DOUGLAS L. MEDIN. 1996. Indeterminacy in the grammar of adult language learners. *Journal of Memory and Language* 35.335–52. DOI: 10.1006/jmla.1996.0019.
- KANNO, KAZUE. 1998. Consistency and variation in second language acquisition. *Second Language Research* 14.376–88. DOI: 10.1191/026765898667204482.
- KERSWILL, PAUL, and ANN WILLIAMS. 2000. Creating a New Town koine: Children and language change in Milton Keynes. *Language in Society* 29.65–115.
- KIDD, EVAN. 2012. Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology* 48.171–84. DOI: 10.1037/a0025405.
- KOTSINAS, ULLA-BRITT. 1988. Immigrant children's Swedish: A new variety? *Journal of Multilingual and Multicultural Development* 9.129–40. DOI: 10.1080/01434632.1988.9994324.
- LABOV, WILLIAM. 1989. The child as linguistic historian. *Language Variation and Change* 1.85–97. DOI: 10.1017/S0954394500000120.
- LABOV, WILLIAM. 1994. *Principles of linguistic change, vol. 1: Internal factors*. Oxford: Blackwell.
- LARDIERE, DONNA. 2007. *Ultimate attainment in second language acquisition: A case study*. Mahwah, NJ: Lawrence Erlbaum.
- LEVEY, STEPHEN. 2006a. Tense variation in preadolescent narratives. *Journal of English Linguistics* 34.126–52. DOI: 10.1177/0075424206290256.
- LEVEY, STEPHEN. 2006b. The sociolinguistic distribution of discourse marker *like* in preadolescent speech. *Multilingua* 25.413–41. DOI: 10.1515/MULTI.2006.022.
- LIGHTFOOT, DAVID. 2010. Language acquisition and language change. *Wiley Interdisciplinary Reviews: Cognitive Science* 1.677–84. DOI: 10.1002/wcs.39.
- MARCHMAN, VIRGINIA A. 1997. Children's productivity in the English past tense: The role of frequency, phonology, and neighborhood structure. *Cognitive Science* 21.283–304. DOI: 10.1207/s15516709cog2103_2.

- MARCUS, GARY F.; STEVEN PINKER; MICHAEL ULLMAN; MICHELLE HOLLANDER; T. JOHN ROSEN; and FEI XU. 1992. Overregularization in language acquisition. *Monographs for the Society for Research in Child Development* 57.1–182. DOI: 10.2307/1166115.
- MCDANIEL, DANA, and HELEN S. CAIRNS. 1996. Eliciting judgments of grammaticality and reference. *Methods for assessing children's syntax*, ed. by Dana McDaniel, Cecile McKee, and Helen Smith Cairns, 233–54. Cambridge, MA: MIT Press.
- MEISEL, JÜRGEN M.; HARALD CLAHSSEN; and MANFRED PIENEMANN. 1981. On determining developmental stages in natural second language acquisition. *Studies in Second Language Acquisition* 3.109–35. DOI: 10.1017/S0272263100004137.
- MIRMAN, DANIEL; JAMES S. MAGNUSON; KATHARINE GRAF ESTES; and JAMES A. DIXON. 2008. The link between statistical segmentation and word learning in adults. *Cognition* 108.271–80. DOI: 10.1016/j.cognition.2008.02.003.
- MOUGEON, RAYMOND; TERRY NADASDI; and KATHERINE REHNER. 2010. *The sociolinguistic competence of immersion students*. Bristol: Multilingual Matters.
- MYERS-SCOTTON, CAROL, and JANICE L. JAKE. 2000. Four types of morpheme: Evidence from aphasia, code switching, and second-language acquisition. *Linguistics* 38.1053–1100. DOI: 10.1515/ling.2000.021.
- NARDY, AURÉLIE; JEAN-PIERRE CHEVROT; and STÉPHANIE BARBU. 2013. The acquisition of sociolinguistic variation: Looking back and thinking ahead. *Linguistics* 51.255–84. DOI: 10.1515/ling-2013-0011.
- NEWMAYER, FREDRICK J. 2003. Grammar is grammar and usage is usage. *Language* 79.682–707. DOI: 10.1353/lan.2003.0260.
- NEWPORT, ELISSA L. 1988. Constraints on learning and their role in language acquisition: Studies of the acquisition of American Sign Language. *Language Sciences* 10.147–72. DOI: 10.1016/0388-0001(88)90010-1.
- NEWPORT, ELISSA L. 1990. maturational constraints on language learning. *Cognitive Science* 14.11–28. DOI: 10.1207/s15516709cog1401_2.
- NEWPORT, ELISSA L. 1999. Reduced input in the acquisition of signed languages: Contributions to the study of creolization. *Creolization, diachrony, and language acquisition*, ed. by Michel DeGraff, 161–78. Cambridge, MA: MIT Press.
- NIYOGI, PARTHA, and ROBERT C. BERWICK. 2009. The proper treatment of language acquisition and change in a population setting. *Proceedings of the National Academy of Sciences* 106.10124–29. DOI: 10.1073/pnas.0903993106.
- PAYNE, ARVILLA C. 1980. Factors controlling the acquisition of the Philadelphia dialect by out-of-state children. *Locating language in time and space*, ed. by William Labov, 143–78. New York: Academic Press.
- PERRUCHET, PIERRE, and SEBASTIEN PACTON. 2006. Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences* 10.233–38. DOI: 10.1016/j.tics.2006.03.006.
- PERRUCHET, PIERRE, and BÉNÉDICTE POULIN-CHARRONNAT. 2012. Beyond transitional probability computations: Extracting word-like units when only statistical information is available. *Journal of Memory and Language* 66.807–18. DOI: 10.1016/j.jml.2012.02.010.
- PULLUM, GEOFFREY K. 1980. Syntactic relations and linguistic universals. *Transactions of the Philological Society* 78.1–39. DOI: 10.1111/j.1467-968X.1980.tb01031.x.
- RAMSCAR, MICHAEL. 2002. The role of meaning in inflection: Why the past tense does not require a rule. *Cognitive Psychology* 45.45–94. DOI: 10.1016/S0010-0285(02)00001-4.
- REALI, FLORENCIA, and THOMAS L. GRIFFITHS. 2009. The evolution of frequency distributions: Relating regularization to inductive biases through iterated learning. *Cognition* 111.317–28. DOI: 10.1016/j.cognition.2009.02.012.
- REGAN, VERA. 2013. Variation. *The Cambridge handbook of second language acquisition*, ed. by Julia Herschensohn and Martha Young-Scholten, 272–91. Cambridge: Cambridge University Press.
- ROBERTS, JULIE. 1997. Acquisition of variable rules: A study of (-t, d) deletion in preschool children. *Journal of Child Language* 24.351–72.
- ROBERTS, JULIE. 2002. Child language variation. In Chambers et al., 333–48.
- ROSS, DANIELLE S. 2001. *Disentangling the nature–nurture interaction in the language acquisition process: Evidence from deaf children of hearing parents exposed to non-native input*. Rochester: University of Rochester dissertation. [Dissertation Abstracts International, 62-07B, 3402.]

- ROSS, DANIELLE S., and ELISSA L. NEWPORT. 1996. The development of language from non-native linguistic input. *Proceedings of the Boston University Conference on Language Development (BUCLD)* 20.623–45.
- RUMELHART, DAVID E., and JAY L. MCCLELLAND. 1986. On learning the past tenses of English verbs. *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 2: Psychological and biological models*, ed. by Jay L. McClelland, David E. Rumelhart, and the PDP Research Group, 16–71. Cambridge, MA: MIT Press.
- RUTTER, BEN. 2014. The acquisition of newly emerging sociophonetic variation: /str-/ in American English. *Journal of Child Language* 41.1166–78. DOI: 10.1017/S0305000913000342.
- SABBAGH, MARK A., and ANNETTE M. E. HENDERSON. 2007. How an appreciation of conventionality shapes early word learning. *New Directions for Child and Adolescent Development* 115.25–37. DOI: 10.1002/cd.180.
- SAFFRAN, JENNY R.; RICHARD N. ASLIN; and ELISSA L. NEWPORT. 1996. Statistical learning by 8-month-old infants. *Science* 274.1926–28. DOI: 10.1126/science.274.5294.1926.
- SAFFRAN, JENNY R.; ELIZABETH K. JOHNSON; RICHARD N. ASLIN; and ELISSA L. NEWPORT. 1999. Statistical learning of tone sequences by human infants and adults. *Cognition* 70.27–52. DOI: 10.1016/S0010-0277(98)00075-4.
- SAFFRAN, JENNY R.; ELISSA L. NEWPORT; and RICHARD N. ASLIN. 1996. Word segmentation: The role of distributional cues. *Journal of Memory and Language* 35.606–21. DOI: 10.1006/jmla.1996.0032.
- SCHULMAN, BARI W.; CAROLYN MYLANDER; and SUSAN GOLDIN-MEADOW. 2001. Ergative structure at sentence and discourse levels in a self-generated communication system. *Proceedings of the Boston University Conference on Language Development (BUCLD)* 25.2.815–24.
- SINGLETON, JENNY L., and ELISSA L. NEWPORT. 2004. When learners surpass their models: The acquisition of American Sign Language from inconsistent input. *Cognitive Psychology* 49.370–407. DOI: 10.1016/j.cogpsych.2004.05.001.
- SMITH, JENNIFER; MERCEDES DURHAM; and LIANE FORTUNE. 2007. ‘Mam, my trousers is fa’in doon!’: Community, caregiver and child in the acquisition of variation in a Scottish dialect. *Language Variation and Change* 19.63–99. DOI: 10.1017/S0954394507070044.
- SMITH, JENNIFER; MERCEDES DURHAM; and HAZEL RICHARDS. 2013. The social and linguistic in the acquisition of sociolinguistic norms: Caregivers, children, and variation. *Linguistics* 51.285–324. DOI: 10.1515/ling-2013-0012.
- SMITH, KENNY, and ELIZABETH WONNACOTT. 2010. Eliminating unpredictable variation through iterated learning. *Cognition* 116.444–49. DOI: 10.1016/j.cognition.2010.06.004.
- SORACE, ANTONELLA. 2000. Syntactic optionality in non-native grammars. *Second Language Research* 16.93–102. DOI: 10.1191/026765800670666032.
- STASSEN, LEON. 1997. *Intransitive predication*. Oxford: Clarendon.
- STASSEN, LEON. 2005. Predicative adjectives. *The world atlas of language structures*, ed. by Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie, 478–81. Oxford: Oxford University Press.
- STEVENSON, HAROLD W., and MORTON W. WEIR. 1959. Variables affecting children’s performance in a probability learning task. *Journal of Experimental Psychology* 57.403–12. DOI: 10.1037/h0039258.
- STEVENSON, HAROLD W., and EDWARD F. ZIGLER. 1958. Probability learning in children. *Journal of Experimental Psychology* 56.185–92. DOI: 10.1037/h0044902.
- TAGLIAMONTE, SALI A., and ALEXANDRA D’ARCY. 2009. Peaks beyond phonology: Adolescence, incrementation, and language change. *Language* 85.58–108. DOI: 10.1353/lan.0.0084.
- THOMPSON, SUSAN P., and ELISSA L. NEWPORT. 2007. Statistical learning of syntax: The role of transitional probability. *Language Learning and Development* 3.1–42. DOI: 10.1080/15475440709336999.
- TOMASELLO, MICHAEL. 2000. Do children have adult syntactic competence? *Cognition* 74.209–53. DOI: 10.1016/S0010-0277(99)00069-4.
- TOMASELLO, MICHAEL. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

- TORO, JUAN M., and JOSEP B. TROBALÓN. 2005. Statistical computations over a speech stream in a rodent. *Perception and Psychophysics* 67.867–75. DOI: 10.3758/BF03193539.
- TROUTMAN, CELINA; BRADY CLARK; and MATTHEW GOLDRICK. 2008. Social networks and intraspeaker variation during periods of language change. *University of Pennsylvania Working Papers in Linguistics* 14.2:18. Online: <http://repository.upenn.edu/pwpl/vol14/iss2/18>.
- VAN VALIN, ROBERT D., JR., and RANDY J. LAPOLLA. 1997. *Syntax: Structure, meaning and function*. Cambridge: Cambridge University Press.
- WEIR, MORTON W. 1964. Developmental changes in problem-solving strategies. *Psychological Review* 71.473–90. DOI: 10.1037/h0041785.
- WILSON, SARAH T., and CARLA L. HUDSON KAM. 2014. Learning syntax through statistics: When do transitional probabilities need a boost? Atlanta: Georgia Institute of Technology, MS.
- WOLFRAM, WALT. 1985. Variability in tense marking: A case for the obvious. *Language Learning* 35.229–53. DOI: 10.1111/j.1467-1770.1985.tb01026.x.
- WONNACOTT, ELIZABETH, and ELISSA L. NEWPORT. 2005. Novelty and regularization: The effect of novel instances on rule formation. *Proceedings of the Boston University Conference on Language Development (BUCLD)* 29.2.663–73.

Department of Linguistics
Totem Field Studios
2613 West Mall
Vancouver, BC, Canada
V6T 1Z4
[Carla.HudsonKam@ubc.ca]

[Received 17 October 2012;
revision invited 13 May 2013;
revision received 11 July 2014;
accepted 12 November 2014]