Learning a Language from Inconsistent Input: Regularization in Child and Adult Learners

Alison C. Austin, Kathryn D. Schuler, Sarah Furlong & Elissa L. Newport

Published online: 22 Sep 2021.

Submit your article to this journal ↗

Article views: 4301

View related articles ↗

View Crossmark data ↗

Citing articles: 9 View citing articles ↗

Routledge
Taylor & Francis Group

Learning a Language from Inconsistent Input: Regularization in Child and Adult Learners

Alison C. Austin*, Kathryn D. Schuler ⓘ*,**, Sarah Furlong ⓘ#, and Elissa L. Newport ⓘ

Georgetown University, Center for Brain Plasticity and Recovery, Washington, DC, USA

ABSTRACT

When linguistic input contains inconsistent use of grammatical forms, children produce these forms more consistently, a process called "*regularization*." Deaf children learning American Sign Language from parents who are non-native users of the language regularize their parents' inconsistent usages. In studies of artificial languages containing inconsistently used morphemes, children, but not adults, regularized these forms. However, little is known about the precise circumstances in which such regularization occurs. In three experiments we investigate how the type of input variation and the age of learners affects regularization. Overall our results suggest that while adults tend to reproduce the inconsistencies found in their input, young children introduce regularity: they learn varying forms whose occurrence is conditioned and systematic, but they alter inconsistent variation to be more regular. Older children perform more like adults, suggesting that regularization changes with maturation and cognitive capacities.

Learning a language is a daunting task. With little to no explicit instruction, young children must extract the relevant parts of the speech stream and learn the allowable combinations of sounds, words, and sentences. A central question is what biases and abilities allow children to learn a language so successfully. One mechanism hypothesized as part of this task is a sensitivity to the distributional statistics of the language environment. By tracking various statistics concerning the frequencies and co-occurrences of linguistic elements, young language learners can extract regularities at many levels, including phonology (Maye et al., 2002), phonotactics (Chambers et al., 2003), word segmentation (Aslin et al., 1998; Newport & Aslin, 2004; Saffran et al., 1996), morphology (Schuler, Yang & Newport, 2016), and syntax (Gerken, 2004, 2006; Gomez, 2002; Gomez & Gerken, 1999; Koulaguina & Shi, 2013; Mintz et al., 2002; Reeder, Newport & Aslin, 2013; Thompson & Newport, 2007; Wonnacott, Tanenhaus & Newport, 2008).

Although this evidence demonstrates that children are sensitive to the statistical distributions in their linguistic environment, the process by which such statistics become linguistic knowledge is still unclear. In many cases children grow up to speak a language with much the same statistics as the one they grew up with; but in some circumstances children acquire a language that is quite different than their input, suggesting that the input statistics alone do not always govern learning. Previous research has found that children may produce a typologically common linguistic structure, even when that pattern is not the most frequent or the most consistent pattern in their input (Culbertson & Newport, 2015, 2017). This provides evidence that children may be biased toward certain grammatical patterns

and that these biases may play a role in language change over time. Other evidence shows that children contribute to sound change, also suggesting their role in historical language change (Labov, 2007; Roberts, 1999).

Children may also acquire a language different from their input when they grow up in a highly variable, unstable, or emerging linguistic community. *Pidgin languages* develop in communities of adults who speak mutually distinct languages but need to communicate with one another. In the hands of adults, pidgin languages often have a simple grammatical structure, with more complex syntax and functional morphemes absent or used inconsistently within and across speakers. Over time, some pidgins will become more stable and more structurally complex, a process called *creolization*.

Many researchers have suggested that an important part of this process involves changes in the language introduced by the way in which children acquire, regularize, and expand its structures as they learn it as their native language (see DeGraff, 1999, for various views on this question). Though investigators differ regarding whether they believe children are the driving force behind this structural change (Bickerton, 1984; Hall, 1966; Kegl et al., 1999; Kocab et al., 2016; McWhorter, 2005; Mufwene, 2007; Sankoff, 1979; Sankoff & Laberge, 1973; Senghas & Coppola, 2001), there are a number of circumstances in which children clearly do play an important role in reorganizing their linguistic input.

Unfortunately, there is little clear data about how such processes occur and precisely how children learn and change the language in these circumstances. However, a somewhat similar situation – in which children learn their native language from non-native, inconsistent input – can also be found in the Deaf community, where studies of this learning have provided an additional empirical base to these questions. Most parents of Deaf children are hearing and do not know how to sign before their children are born; and even most Deaf parents of Deaf children are themselves not native signers. Deaf children thus often learn American Sign Language (ASL) from parents who are late learners of ASL, and they may not have exposure to other, more fluent users of the language. Although this situation is in many ways quite different from that of creolization, it is similar in one important way: many Deaf children learn their native language entirely from models who do not have native proficiency and whose language contains many inconsistencies as compared with native ASL.

Newport and colleagues (Newport, 1999; Ross, 2001; Ross & Newport, 1996; Singleton & Newport, 2004) have studied several such cases, each involving a Deaf child learning ASL from parents who did not themselves learn to sign until their teens or later. Non-native speakers of a language typically do not attain the same level of language proficiency as native speakers do (Boudreault & Mayberry, 2006; Johnson & Newport, 1989; Mayberry & Fischer, 1989; Newport, 1990; Singleton & Newport, 2004), and their utterances contain inconsistent grammatical errors that form part of the linguistic input of their children. Strikingly, however, the children studied by Newport and colleagues used ASL with much greater consistency than their parents – and often a level of consistency roughly equal to that of children learning from parents who were native signers. Singleton and Newport (1999, 2004) suggest that the children attained this outcome by producing only the most frequent forms of their parents' language (and not their parents' errors). Though the parents made many errors, for a given grammatical context they usually used one morphemic form (typically, but not always, the correct form) more often than any other. Children, then, *regularized* these forms in their own language productions, using them *more* consistently than their parents did. Newport (1999) has suggested that this regularization process may also underlie many of the changes that occur in creolization.

To understand this phenomenon further, Hudson Kam & Newport (2005, 2009) brought learning from inconsistent input into the lab by using miniature artificial languages. They created artificial languages in which most constructions were regular and consistent, but "determiners" (short, meaningless words "*ka*" and "*po*" that followed the nouns) had the inconsistent, probabilistic usage that is characteristic of many function words and other closed class elements in non-native speakers and in pidgins. In one experiment (Hudson Kam & Newport, 2005), determiners were probabilistically present or absent throughout the language. Adult participants were exposed to one of four languages

that differed in how frequently determiners were present. After learning, the adult learners mirrored their frequency of occurrence in their own productions, which is noticeably unlike the children described above.

Crucially Hudson Kam & Newport (2005, 2009) also compared adults with children learning the same artificial languages. In one language the main determiner appeared, probabilistically, in 60% of all noun phrases, with other determiners absent or occurring at lower frequencies. The adults again learned the inconsistencies present in the input: their probability of producing each determiner matched those in the input. In contrast, the children in these experiments looked like children learning natural languages from non-native input: they regularized one of the forms, using it in virtually all of their noun phrases. This suggests that children in artificial language studies can and do regularize inconsistent input.

This pattern of results from Hudson Kam & Newport (2005, 2009) opened many questions about the circumstances in which learners will regularize and why. One question – why adults do not regularize like children – has received considerable attention in the literature. Hudson Kam & Newport (2009) suggested that children may regularize due to memory or other cognitive limitations (Newport, 1990). However, experiments attempting to limit adult memory have produced mixed results: sometimes adults regularize more under these circumstances (Hudson Kam & Chang, 2009; Hudson Kam & Newport, 2009), but sometimes their behavior is unchanged (Perfors, 2012). Other research using adult learners or computational models has asked whether domain-general or language-specific biases can explain regularization patterns (Ferdinand, Kirby, & Smith, 2019; Perfors, 2016; Reali & Griffiths, 2009; Smith & Wonnacott, 2010). Smith et al. (2017) summarize this work and caution that the relationship between learning biases, language structure, and regularization is a complex one: weak biases can lead to strong regularization over multiple generations of learners, but strong biases can have little to no effect when individuals differ in how they regularize.

Unfortunately, most of this work has focused on whether adults will ever regularize inconsistent input; considerably less research has focused on children – who are, after all, the learners of central interest on this hypothesis. Hudson Kam (2015) and Samara et al. (2017) have used artificial language paradigms to ask how children learn conditioned variation, in which variation of forms occurs conditionally dependent on specific grammatical or social contexts. In Hudson Kam (2015), children were exposed to one of two languages with an optional determiner. In the inconsistent language, the determiner occurred more often across the language as a whole (60% of nouns), but presence of the determiner was not predicted by any particular context. In the conditioned variation language, the determiner occurred more often on subject nouns (75%) and less often on object nouns (25%). Hudson Kam (2015) found that, while children made the language more regular in both conditions, they regularized significantly less when determiner use was conditioned on grammatical role. Similarly, Samara et al. (2017) exposed children to one of three languages in which determiner use was either unconditioned, deterministically conditioned on speaker identity (boy vs. girl), or probabilistically conditioned on speaker identity. Children learned the conditioning in the deterministically conditioned language, and – as in Hudson Kam (2015) – while children made the language more regular in both the unconditioned and probabilistically conditioned languages, regularization was much stronger when determiner use was unconditioned.

While conditioned variation is common in languages of the world (e.g Labov, 1989; Aissen, 2003; Chappell & Verstraete, 2019; D'Arcy & Tagliamonte, 2015; Kroch & Small, 1978), there are also many circumstances in which variation is not conditioned but is inconsistent, as studied by Singleton and Newport (2004) and Hudson Kam & Newport (2005, 2009). However, questions about the nature of regularization in children learning from inconsistent input have not been further explored. There are a number of questions that are still open and of interest.

First, while children in Hudson Kam & Newport (2005, 2009) were more likely than adults to produce systematic usages, their way of doing so varied somewhat from child to child: while most children formed a "rule" in their determiner usage – a regular pattern that could be identified by its consistency across their productions – they did not all form the same rule and did not all regularize the

majority determiner. This may have been due to the complexity of the language and the difficulty children had in learning it. There are also questions of interest from their results about the mechanisms of regularization. Is regularization limited to production, or would the same phenomenon appear (as it did in some of Hudson Kam & Newport's results) in recognition and rating tasks as well? Were minority determiners not used by children because they were used inconsistently, or simply because they occurred infrequently? Evidence from computational models (Morgan & Levy, 2016) and from children learning an inconsistent gender marker in natural language (Miller, 2013) suggest that frequency, or amount of exposure to a form, may play an important role in whether it is regularized. Finally, how much inconsistency can this regularization process bear? Is it the simply the frequency of a given form that contributes to whether it will be changed, or is it the conditional probability of forms that matters for regularization?

The current set of studies seeks to address several of these questions. First, in order to make the languages easier to learn (and hopefully to obtain less idiosyncratic results) in young children, we made our artificial languages smaller and the learning procedure shorter and more engaging. In Experiment 1 we use this new paradigm to compare children and adults' learning of two very similar artificial languages: one with inconsistent determiners and the other with determiners that were equally frequent but used in a systematic way. This comparison allows us to assess the effects of inconsistency versus frequency per se. In Experiments 2 and 3 we vary the ages of the children tested, asking whether all children regularize or whether regularization changes over age, and we ask whether a very high degree of inconsistency will still result in regularization. We also assess children's recognition of the forms as well as their productions. We will then return to considering the types of learning mechanisms that can explain our results and the potential reasons for differences in learning by children and adults.

Experiment 1

One purpose of this experiment was to compare how children and adults learn languages that contain inconsistent input, using simpler materials and a more enjoyable learning procedure than we used in Hudson Kam & Newport. Participants from both age groups were taught an artificial language in which most constructions were consistent, but inconsistency was introduced in the language's "determiners." As in Hudson Kam & Newport (2005, 2009), the language had a verb-subject-object (VSO) word order, with spoken sentences that accompanied short, filmed vignettes of the referent events. The "determiners" were short CV syllables that followed the nouns (as would be the case in real VSO languages). In the *inconsistent* language, the determiners varied between one of two forms (*ka* and *po*), with one of the determiners appearing in 67% of the noun phrases and the other appearing in the remaining 33% of the noun phrases.[1] Which determiner was used in any given noun phrase was unpredictable beyond these probabilities.

We compared the learning of this inconsistent language with the learning of a *lexically consistent* language, in which determiners were used with the same 67%-33% split between forms used in the inconsistent language, but where the forms were used deterministically. In the lexically consistent language, determiner forms were perfectly dependent on the specific noun with which they co-occur; 67% of the nouns in the language always occurred with *ka* and 33% of the nouns always occurred with *po*. While we expect that in the inconsistent language children will respond quite differently than adults, the lexically consistent language will help to shed light on when and why they do so.

One possibility is that children have trouble learning probabilistically occurring forms and regularize whenever their input is probabilistic in nature. A second possibility is that children fail to learn

---

[1]Most of the languages in, Hudson Kam & Newport (2005, 2009) contained two classes of nouns: those that took *ka* (or no determiner) and those that took *po* (or no determiner). This subcategorization of nouns might have been difficult to learn, and the variation between presence and absence of determiners was hard to score (e.g., when participants produced no determiner, was this a failure to learn the determiner system or a regularization of the "determiner absent" condition?). In the simpler language of the present experiments, all nouns in the inconsistent languages took *ka* at one probability and *po* at another.

the minority determiner simply because it is infrequent: perhaps children have poorer memories, or perhaps they need a greater amount of evidence before they begin producing a given form of a determiner. If the cause of regularization is low frequency, children should regularize whenever the minority determiner is used infrequently. A third possibility is that it is inconsistency per se that children avoid, fail to learn, or regularize in their own productions. The lexically consistent language will allow us to differentiate some of these possibilities, since here the frequencies of *ka* and *po* were precisely matched to those in the inconsistent language, but their usage was perfectly predictable.

Several recent studies suggest that children are indeed able to learn, and do not regularize, this kind of lexically consistent conditioned variation. When patterns of use are deterministically conditioned on lexical items (e.g., Thothathiri and Rattinger, 2016; Wonnacott, Newport, and Tanenhaus, 2008; Wonnacott, 2011), speaker identity (Samara et al., 2017), animacy or grammatical gender (Culbertson et al, 2017, 2019), both children and adults are able to learn the consistent pattern, using the correct form for a given item rather than regularizing. However, some studies suggest that forms encountered infrequently may be regularized, even if they are lexically consistent. Schwab et al. (2018) found that, when children receive only limited exposure to a lexically consistent pattern, they may regularize in production tasks, only revealing they have learned the underlying conditioning in judgment tasks.

Methods

Participants

Participants were 28 5 and 6 year-old children and 17 adult controls. Fifteen children (mean = 5.91, range = 4.94– 6.78) and nine adults (mean = 20.52, range = 18.84– 23.06) were exposed to the inconsistent language; the remaining thirteen children (mean = 6.12, range = 5.43– 6.76) and eight adults (mean = 22.29, range = 20.04– 25.62) were exposed to the lexically consistent language. Seven additional children and three additional adults began the experiment but stopped participating before producing any data (6 children and 2 adults ended participation before the production test; 1 child and 1 adult ended due to experimenter error). Adult participants were undergraduate students at the University of Rochester, former students, or age mates who attended another college. Children were recruited from local after-school programs. All participants were monolingual native English speakers. Adults received 5 USD per session for their participation and a 15 USD bonus for completing the experiment. Children received stickers each session and a small set of toys on the final day.

Description of the language

Both languages had the following characteristics:

Lexicon. The artificial language contained six nouns, eight verbs, and two determiners. Nouns and verbs were one or two syllables long, with half of each type in each class. Each noun began with a different initial consonant and referred to one of six puppet animals: *clydum* (giraffe), *daffin* (dog), *flugit* (bee), *geed* (ladybug), *mawg* (lion), and *spad* (rabbit). Verbs were either transitive: *flairb* (drag), *glim* (ram into), *lepal* (jump on), *tombur* (rock), or intransitive: *bleggen* (move forward), *frag* (fall over), *jentaf* (sit), *zub* (clap). There were two determiners, *ka* and *po,* with no corresponding semantics, distributed as described below.

Grammar. The word order was Verb – Subject – Object (VSO), with determiners following the noun. Because the language included both transitive and intransitive verbs, there were two possible sentence structures:

Verb-Noun-Determiner (e.g., *frag mawg ka (the bee falls))*

Verb-Noun$_{Subj}$-Determiner-Noun$_{Obj}$-Determiner (e.g., *glim mawg po flugit ka* (*the lion rams the bee*)).

Determiner manipulation. The two languages differed only in the distribution of the determiner forms. In the inconsistent language, the *dominant* determiner was used 67% of the time and the *minority* determiner was used the remaining 33% of the time. Aside from this ratio, determiner usage was completely probabilistic and unpredictable. To ensure that there were no unintended patterns underlying the determiner distribution, this 67%-33% split held approximately[2] true not only for the overall distribution of determiners, but also for each noun, for each syntactic position, and for the combination with each verb. The determiners within each utterance were also independent of each other: a dominant determiner occurred in the object position 67% of the time, regardless of whether the dominant determiner was used in the subject position in that utterance, and vice versa. In contrast, determiners in the lexically consistent language were distributed in a lexically consistent manner: 4 of the 6 nouns (67%) were <u>always</u> used with the dominant determiner, while 2 of the 6 nouns (33%) were <u>always</u> used with the minority determiner. (This is similar to gender subcategories in natural languages.) Because of this lexical division, the overall frequencies and probabilities of *ka* and *po* were identical to those in the inconsistent language, but their distribution was deterministic rather than inconsistent. As in the inconsistent language, which of the two determiners (*ka* or *po*) was the dominant determiner was counterbalanced across subjects.

Stimuli

The stimuli consisted of video clips of simple events acted out by puppets and spoken sentences (audio clips) in the artificial language that played along with each video. The sentences were natural speech spoken in the artificial language used for that experiment. Each video lasted no more than 5 seconds and showed one puppet performing an action, either by itself (e.g., falling) or on another puppet (e.g., head-butting or dragging). Across the videos presented during the experiment, the direction of the action (from right to left or left to right) was randomized.

Exposure materials

Thirty videos were selected to make the exposure corpus, three different events for each intransitive verb and four or five different events for each transitive verb. Each noun occurred three times as the subject and three times as the object. Four of the nouns were used in the intransitive sentences; each occurred three times as an intransitive subject. This set of intransitive scenes was divided into two lists of six scenes each. Each half was combined with the full set of transitive scenes. This created two different sets of 24 scenes each, which were used on alternating days of the experiments. The sentences were combined with the scenes, as described above, in order to produce the target determiner probabilities in each exposure set for each day. That is, the sentence paired with a scene would have either *ka* or *po* as the subject and either *ka* or *po* as the object determiner. For the inconsistent language, the determiners used in a given scene-sentence pair varied by day, in order to produce the target 67%-33% determiner probabilities on each day and for each noun and noun position. For the lexically consistent language, whether the determiner was *ka* or *po* depended entirely on the noun – 4 of the 6 nouns (67%) always occurred with the dominant determiner (regardless of syntactic position or day) while 2 of the 6 (33%) always occurred with the minority determiner.

Testing materials

An additional 20 novel scenes – 16 transitive (4 per verb) and four intransitives (1 per verb) – were used to elicit productions during testing. All of the test scenes were novel and were not drawn from the exposure set.

The set of test sentences was divided into two lists, used on alternating days of testing. Each list contained 12 of the 20 scenes: all four intransitives and 8 of the transitive scenes. In both lists, each

---

[2]The actual probabilities ranged from 63%-68% for occurrence with each noun, 64%-70% for occurrence with each transitive verb, and 63%-71% for occurrence with each intransitive verb.

noun appeared as a subject twice (0–1 times in intransitive and 1–2 times in transitive sentences) and 1–2 times as an object.

Procedure

The experiment took place over 5 sessions. All participants followed identical procedures. Each session lasted no more than 30 minutes. Most subjects completed the study over five consecutive days. Each day consisted of 4 types of games and activities: vocabulary learning, vocabulary testing, sentence exposure, and production testing.

Vocabulary learning. Game 1 introduced the puppets to the participants. Images of the puppets appeared randomly and quickly on a computer screen. The participants' task was to click on each puppet as quickly as possible, thus "catching" it and making it disappear from the screen. The corresponding noun in the artificial language for that puppet was heard each time it was successfully caught. At the start and end of the game, the puppets were displayed statically on the screen. Participants heard the nouns for the puppets in the artificial language and were asked to repeat them. As they became more familiar with the language, participants were asked to name the puppets rather than repeat the noun. Game 2 was used only when a participant had trouble learning a specific noun. It was identical to game 1, except that the task was to catch only a single puppet. Participants heard the noun for this puppet when it was clicked on, but a buzzing noise if any other puppet was caught. Game 3 consisted of a matching task, in which participants clicked a button to hear one of the nouns and were asked to click on the picture of the corresponding puppet. When the correct puppet was chosen, the noun for the puppet was repeated. When an incorrect puppet was chosen, a buzzing noise was heard, and they were instructed to choose another puppet until the correct one was selected.

Vocabulary testing. To test for vocabulary knowledge, the puppets were displayed on the screen one at a time; participants were asked to name the puppet in the artificial language. If incorrect, the experimenter would help by giving the initial sound of the word. If the participant still could not name the puppet, they were told the correct answer. All responses were recorded on the computer.

Sentence exposure. Participants viewed the videos one at a time; each video was followed immediately by a sentence in the artificial language describing what happened. Participants were asked to repeat each sentence aloud. This was done to ensure that participants were paying attention to the sentences and that they had heard all the words in the sentence correctly. During the first two sessions, participants were allowed to watch the video and hear the sentence as often as needed in order to help them repeat the sentence. In difficult cases, the experimenter would repeat the sentence in order to help. During the final three sessions, the video and sentence played only once. Sentences from the exposure set were presented in random order.

Production test. On each test trial, participants viewed novel scenes of the puppets. These scenes were of the same actions and the same puppets the participants had seen before, but in novel combinations. Participants then heard the verb corresponding to the action in the scene as a prompt, with the pronunciation of this verb ending with a rising intonation (as though the sentence were incomplete). Their task was to repeat the verb and complete the sentence. Productions were recorded via computer.

Sequence of components. On days 1–2, participants did vocabulary learning, vocabulary testing, and sentence exposure. On days 3–5, participants did vocabulary learning, vocabulary testing, sentence exposure, and production test.

Results

Vocabulary testing

Performance on the vocabulary test was used as an indicator of how well the language was learned. Both children and adults performed well on the vocabulary test. Adults in both languages knew all six nouns (mean = 6.0, sd = 0.0), children in the inconsistent language knew an average of 5.4 (sd = 0.83), and children in the lexically consistent language knew an average of 5.15 (sd = 1.41).

Production data

All sentences produced during testing were later transcribed and coded for whether or not the participant used the dominant determiner.

To determine whether children used the dominant form more than adults and whether this behavior differed across conditions, we ran a mixed-effects logistic regression model predicting use of the dominant form by Age Group (adult[3] vs child), Language (inconsistent vs. lexically consistent), and Day.[4] For this and all future models, we conducted model comparison in order to select the best-fitting random effect structure. Selecting random-effects in this way allows us to generate models that are only as complex as justified by the data, avoiding failures to converge and adding complexity (e.g., by item effects) only when doing so makes a significant contribution to model fit (Bates et al., 2015). Following this, we included random intercepts for participants. There was no main effect of Day and no interaction with Day (all p's > 0.76). Figure 1 therefore shows participants' mean and individual use of the dominant determiner by language condition over all testing days.

Overall, participants used the dominant form significantly more than chance (where chance = 50%, β = 1.23, SE = 0.16, z = 7.85, p < .001). There was a main effect of Age Group, with young children significantly more likely to use the dominant form than adults (β = 0.83, SE = 0.31, $\chi^2(1)$ = 6.25, p = .01). There was also a main effect of condition, with participants in the Lexically consistent language using the dominant form significantly less than participants in the Inconsistent condition (β = −0.67, SE = 0.31, $\chi^2(1)$ = 5.62, p = .02), and a significant interaction between Age Group and Condition, showing that the effect of age depends on condition ($\chi^2(1)$ = 5.68, p = .02). These results show that children are much more likely to use the dominant form than adults, but this difference lies in the inconsistent language condition; their use of the dominant form is significantly and substantially lower in the lexically consistent language (β = −1.50, SE = 0.62), where adults and children perform similarly.

As can be seen in Figure 1, these effects appear to be due to the fact that children in the inconsistent language regularize their use of their dominant determiner, whereas children in the lexically consistent language and adults in both conditions match their input. To test where use of the dominant determiner was significantly different from the input level, we conducted t-tests against the level of the dominant determiner in the input (67% in both languages). In the inconsistent language, adults used the dominant determiner in 67.47% of all determiners produced (SE = 3.65%) – not significantly different from the input level (t(8) = 0.13, p = .90). Children used the dominant form substantially more often than adults (mean = 85.82%, SE = 4.44%). This is significantly and dramatically different from the probability with which it was present in the input (t(14) = 4.24, p < .001). In the lexically consistent language, adult productions mirrored those found in their input (mean dominant determiner = 69.86%, SE = 1.30%, t(7) = 2.22, p = .06), as was the case for the inconsistent language. However, in stark contrast to the inconsistent language, children in the lexically consistent language also mirrored the input probabilities (mean = 67.31%, SE = 5.07%, t(12) = −0.06, p = .95). This suggests that the children did not regularize the inconsistent language merely because of the probabilistic nature of their input or because of the low frequency of the minority form, but rather because of their inconsistent distribution. When the determiner forms were distributed with the same probabilities but in a lexically determined way, children did not regularize their usage.

---

[3]We used simple coding for all fixed effects in our models. Throughout the paper, the first level listed in the parenthetical corresponds to the reference level for that fixed effect in the model.
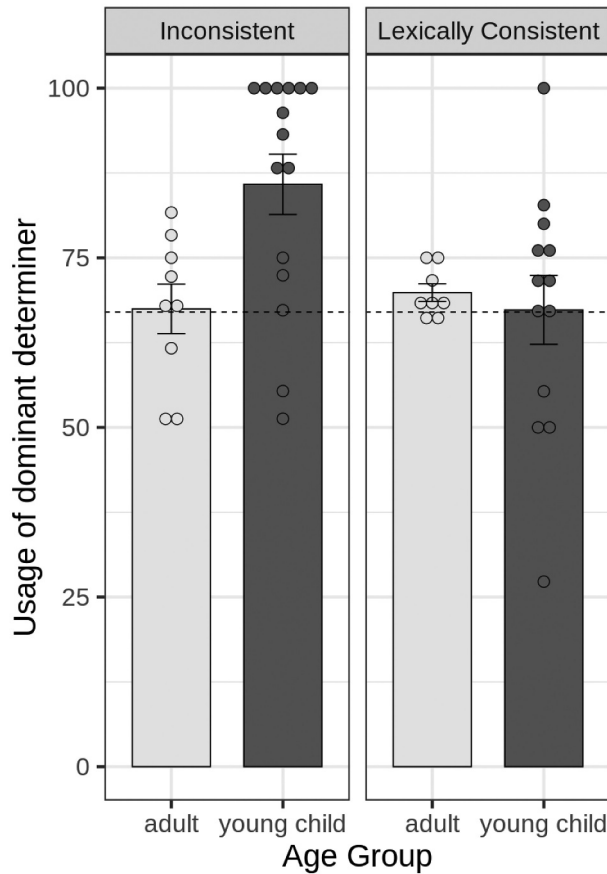[4]In all models, Day is a continuous predictor, centered on Day 4.

Figure 1. Mean production of dominant determiner in Experiment 1 by age group and language condition. Dots indicate individual participant production of the dominant form; error bars indicate standard error of the mean. Dashed line indicates level of the dominant determiner present in the input.

Of particular interest in the lexically consistent language is whether the participants learned the lexical pattern behind the 67%-33% distribution. Indeed, both age groups learned the correct determiner for each group of nouns. By the final day of testing, adults used the dominant determiner 98.21% of the time (SE = 1.79%) with nouns that were heard exclusively with the dominant determiner and used the minority determiner 92.86% of the time (SE = 5.40%) with nouns that were heard exclusively with the minority determiner. Child participants showed the same pattern, using the dominant determiner with dominant-only nouns (mean = 86.26%, SE = 8.96%) and the minority determiner with minority-only nouns (mean = 80.60%, SE = 6.40%).

To assess this learning, the data were entered into a mixed-effect logistic regression predicting use of the correct determiner – that is, matching the lexically consistent pattern – by Age Group (adult vs. child), Noun Type (minority, dominant), and Day. Overall, participants used the correct determiner significantly greater than chance (where chance = 50%, β = 0.96, SE = 0.31, z = 3.08, p < .01). There was a significant main effect of Noun Type, indicating that participants were significantly more likely to be correct for nouns taking the more dominant determiner compared with the minority one (β = 5.82, SE = 0.72, χ^2 (1) = 33.16, p < .001). There was no main effect of Age Group, indicating that children learned the lexical conditioning as well as adults (β = −0.17, SE = 0.55, χ^2 (1) < 0.001, p = 1.00). Here we found no main effect of day (χ^2 (1) = 0.11, p = .74). However, there was an interaction between Day and Noun Type, with dominant nouns differing from minority nouns (β = 1.10, SE = 0.36, χ^2 (1) = 9.08 p < .01). This indicates that learning increased as exposure increased, particularly for the minority (less

frequent) nouns. We also found a three-way interaction between Day, Age Group, and Noun Type, with children differing significantly from adults (β = −3.34, SE = 0.71, χ^2 (1) = 26.23, p < .001). This indicated that the effect of noun type over days is only true for adults; children matched the noun type through all three days.

We can also ask whether participants' choice of determiner in the inconsistent language follows a rule-like pattern (even though the input in this language was carefully constructed not to contain any such rule). For example, participants in the inconsistent language might have used the dominant determiner with all of their nouns; or they might have conditioned their determiner use lexically, always producing the same determiner with a given noun, even though lexical conditioning was not present in the input in this language. To assess this, we examined the productions of each participant to determine whether a consistent response pattern appeared for that participant. Participants were coded as *following a rule* (for example, using the dominant form consistently with all nouns, or using one determiner form consistently for some nouns and another determiner form consistently for other nouns) *if all or all but one of their productions* on a given day were consistent with the rule. By the final day of testing, 11 of 15 children in the inconsistent language regularized the dominant form on this very strict criterion. In the lexically conditioned language, 7 of 13 children conditioned the determiner form on the lexical item, 1 regularized the dominant form, and 1 regularized the minority form on this very strict criterion (> 91% of test trials). Eleven out of thirteen children matched the lexically conditioned pattern on more than 80% of test trials.

These results are consistent with other metrics of regularization, such as entropy and mutual information, which provide a value corresponding to how regular or conditioned a participant's productions are. For example, in our artificial language, an entropy of 0 would indicate a participant used one determiner consistently on every trial (either the dominant or the minority form) and an entropy of 1 would indicate a participant used both determiners equally often. We calculated the entropy of each participant's productions on each day of testing in the Inconsistent language.[5] Then we ran a mixed-effects regression predicting entropy by Age Group and Day. Analyzed in this way, we found a main effect of Age Group, with children's productions showing significantly lower entropy than adults (β = −0.35 SE = 0.09, χ^2 (1) = 12.43, p < .001) – that is, children's productions were significantly more regular – with a mean entropy on the final day of testing of 0.20 (SE = 0.08) for children and 0.58 (SE = 0.03) for adults. There were no significant effects of Day or interactions with Age Group and Day on entropy (all p's > 0.55).

Taken together, these results suggest that children are very likely to form highly consistent rules: they regularize the dominant form (or occasionally form another rule) when provided with inconsistent input, or learn the determiner form that accompanies specific lexical items when provided with lexically consistent input. In contrast, adults readily learn lexical rules that are presented in their input, but they are much less likely to form a rule when the input is inconsistent (though a few adults do condition their productions on the noun to a greater extent than the input).

Discussion

Experiment 1 demonstrated the viability of using this new procedure to investigate how children learn language from inconsistent input. In accord with Hudson Kam & Newport (2005, 2009), the adults in Experiment 1 were quite adept at mirroring their input. In the 67%-33% inconsistent condition, adults used determiners with almost precisely the same probabilities as they were used in the input. In contrast, the children's productions in the inconsistent condition did not look like their input. Children were much more likely to behave in a systematic, rule-like way than adults were. Additionally, the type of rules the two age groups made were quite different. Children favored a rule that maximized the use of the dominant determiner, while adult productions were more likely to reflect probabilistic variation between the two forms.

---

[5]In this and all future experiments, entropy was calculated using the *infotheo* package in R.

In the lexically consistent condition, both children and adults were able to learn a deterministic rule regarding how determiners were distributed in the language, even though the overall determiner frequencies and probabilities were matched to those in the inconsistent language. This demonstrates that young children's regularization of inconsistency was not a response to variation or probabilistic usage in general and was not simply due to the low frequency of the minority determiner, but rather was a response to the inconsistent nature of that variation. The frequency of the minority determiner in the lexically consistent language was exactly the same as in the 67%-33% inconsistent language, but only in the lexically consistent language did they use the minority determiner in the same way in their own productions.

Experiment 2

In order to probe further how determiner production changes with age, in Experiment 2 we included three age groups: 5–6 year old children, 7–8 year old children, and college-age adults. While 7–8 year olds are only slightly older than the group of children used in Experiment 1, previous studies of language acquisition suggest that the ability of learning a language to native proficiency begins to decline at about this age (Beaudrault & Mayberry, (2006): Johnson & Newport, 1989; Mayberry & Fischer, 1989; Newport, 1990; Weber-Fox & Neville, 1996).

Experiment 2 also included an additional measure of language learning. From the first experiment it is unclear whether regularization occurs as part of *learning* or rather is strictly a *production* effect. Hudson Kam & Newport (2005, 2009) included not only a production task but also a grammaticality rating task; however, in their studies, ratings sometimes were regularized in line with productions and sometimes were not. In Hudson Kam & Newport (2005), children judged as highly acceptable only those utterances with the determiners present, in accord with their productions; they rated sentences with determiners absent (which they heard 40% of the time) as low as sentences with misplaced determiners (a form that they had never heard). In Hudson Kam & Newport (2009), children rated most highly the form they produced most frequently; but they rated the minority determiners next best and significantly better than forms they had not heard in their input. In a miniature language study of children learning two novel gender-related classifiers, Schwab et al. (2018) found that, when children had not yet learned the classifier system, they used only one classifier in production but selected either classifier in a 2AFC task. Taken together, these studies show different results, suggesting that children's ability to recognize or rate minority forms is more erratic than their (highly consistent) productions. Ferdinand et al. (2019) also found that adults regularized more in their productions than was reflected when they were asked to estimate the frequency of variants in the input. To further investigate these questions, we added a grammaticality rating task, similar to the ones used in Hudson Kam & Newport, on the final day of the procedure in each of the remaining experiments. While a rating task provides only a rough way of looking at language knowledge and representation in young children in light of their often limited abilities to perform such a task (Gleitman et al., 1973), it can provide some additional insight into these questions.

Finally, we changed the distribution of determiner use from 67% dominant, 33% minority (as in Experiment 1) to 60% dominant, 40% minority. This change was made in anticipation of Experiment 3, in which we expose children to a 40% dominant determiner. This will enable us to compare learners' use of a 40% determiner when it is the minority form (in the present experiment) with their determiner use when it is the majority form (Experiment 3).

Methods

Participants

Fourteen younger children (mean = 5.98, range = 5.07– 6.71), sixteen older children (mean = 7.69 years, range = 6.97– 8.67), and twelve adult controls (mean = 20.53, range = 18.91– 22.93) participated in this experiment. All participants were recruited and compensated as in Experiment 1.

Description of the language

The artificial language was as in Experiments 1 and 2, except that it contained only five nouns (*daffin* (dog) was not used) and only transitive verbs (the 4 transitive verbs from Experiment 1 and an additional verb *daygin* (hit)) to facilitate balancing the determiners equally throughout the language. We manipulated the determiners as before, except the *dominant* determiner was used 60% of the time and the *minority* determiner was used the remaining 40% of the time. Aside from this ratio, determiner usage was completely probabilistic and unpredictable. To ensure that there were no unintended patterns underlying the determiner distribution, the 60%-40% split was true not only for the overall distribution of determiners, but also for each noun, for each syntactic position, and for the combination with each verb. As before, which determiner was the dominant determiner was counterbalanced across subjects. All other aspects of the language were entirely consistent and regular.

Stimuli

Exposure materials. The exposure materials were as in Experiment 1, except we presented 25 rather than 30 scenes (5 per verb), all scenes were transitive, and each of the 5 nouns occurred 5 times as a subject and 5 times as an object.

Testing materials. Three sets of 10 novel transitive scenes were created to elicit productions during testing. In each set, each verb was used twice, and each noun was used twice as a subject and twice as an object. Participants rotated through the three sets on the three days of testing, receiving a new set of test sentences each day. Which test set was used on which testing day was counterbalanced across subjects.

Grammaticality rating materials. Ten additional novel sentence-scene pairs were used for grammaticality ratings, divided into two groups of 5 sentences. In each group, each verb appeared once and each noun appeared once as a subject and once as an object. Children received only one of the two groups of sentences; adult participants received both groups. Within each group, each sentence was presented 4 times over the course of experiment. Each time, one of the two determiners in the sentences was either the dominant form, the minority form, a novel form, or entirely omitted. The other determiner in the sentence was always the dominant form. The position of the manipulated determiner (subject or object) was balanced throughout the course of the rating task.

Procedure

The procedure on days 1–4 was identical to that used in Experiment 1. On day 5, the grammaticality-rating task was added after the final production test. First, participants practiced with examples in English and scenes involving puppets that were not used in the experimental task. Some of the English sentences were grammatical, but others contained agreement or word order errors. Participants listened to an English sentence describing the scene and were asked if it sounded like a "good" sentence in English. They could indicate how good they thought the sentence was by clicking on one of four smiley faces, which corresponded to ratings of 1–4, below the video: frown, no expression, small smile, or grin. After the practice, participants completed the same task with the set of test sentences.

Results

Vocabulary testing

As in Experiment 1, both children and adults performed well on the vocabulary test. Adults and older children knew all five nouns (mean = 5.0, sd = 0.0 for both groups) while younger children knew an average of 4.6 (sd = 0.91).

Production data

To determine whether use of the dominant determiner changed with age, we first used mixed-effects logistic regression to predict use of the dominant form by Age Group (adult, older child, younger child) and Day. Model comparison indicated that the best fitting random effect structure allowed the effect of day to vary by participant (χ^2 (1) = 14.97, $p < .01$). There was no main effect of Day, and no significant interactions with Day (all p's > 0.39). Because there was no effect of Day on determiner use, Figure 2 shows participants' mean and individual use of the dominant determiner by language condition across all 3 testing days.

Overall, participants used the dominant form more than expected by chance (where chance = 50%, $\beta = 0.80$, $SE = 0.17$, $z = 4.61$ $p < .001$). There was a significant main effect of Age Group χ^2 (2) = 7.75, $p = .02$). The coefficient estimates indicate that older children do not differ from adults ($\beta = -0.05$, $SE = 0.42$, $z = -0.118$, $p = .91$), but younger children do use the dominant form marginally more than adults ($\beta = 0.79$, $SE = 0.44$, $z = 1.80$, $p = .07$).

To determine whether children or adults used the dominant determiner more often than it was present in the input, we next conducted a t-test against the level of the dominant determiner in the input (60%). Adults used the dominant determiner in 63.06% of all determiners produced ($SE = 2.64\%$) and older children used the dominant determiner in 60.09% of all determiners produced ($SE = 5.45\%$); neither is significantly different from the probability with which the dominant determiner was present in the input (adults: t(11) = 1.16, $p = .27$; older children: t(15) = 0.02, $p = .99$). In contrast, younger children used the dominant form significantly more than was present in the input (mean = 73.49%, $SE = 5.33\%$, t(13) = 2.53, $p < .03$).
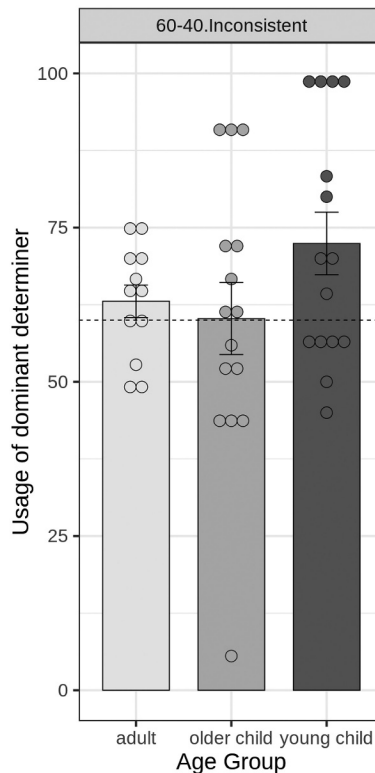

Figure 2. Mean production of dominant determiner in the 60%-40% inconsistent language (Experiment 2) in 3 age groups. Dots are individual participants; error bars are standard error. Dashed line indicates level of dominant determiner present in the input.

As in Experiment 1, we asked whether participants formed rule-like patterns, for example, whether they regularized the dominant or minority determiner or formed a rule based on lexical items or grammatical role. To assess this, we examined each participants' responses individually to determine whether a consistent response pattern appeared for that participant and coded each participant on each day as following a rule or not, as in Experiment 1. Participants were coded as following a rule if all or all but one of their productions on a given day were consistent with that rule. By the final day of testing, zero adults and one older child had formed rules (conditioned on grammatical role). In contrast, on this very strict criterion six of the younger children had formed rules (five maximized the dominant form and one formed a lexically conditioned rule).

To capture regularization beyond use of the dominant form, we next analyzed the entropy of participant productions. Recall that an entropy of 0 indicates that a participant's productions are completely regular (i.e. they used a single determiner in all productions), while an entropy of 1 indicates that the participant uses both determiners equally. As in Experiment 1, we calculated the entropy of each participant's productions on each day of testing and used mixed-effects regression to predict entropy by Age Group and Day. Analyzed in this way, we found a main effect of Day, with entropy increasing as day progressed ($\beta = 0.04$ $SE = 0.02$, χ^2 (1) = 5.95, $p = .02$). There was a significant main effect of Age Group, in which young children but not older children had significantly lower entropy than adults ($\beta_{young} = -0.17$, $SE_{young} = 0.08$, $\beta_{old} = -0.07$, $SE_{old} = 0.08$, χ^2 (2) = 5.94, $p = .05$) – that is, young children's productions were significantly more regular – with a mean entropy of 0.42 for young children, 0.61 for older children, and 0.62 for adults.

Grammaticality ratings

The results from the grammaticality ratings task are shown in Figure 3. Because subjects varied in the range of ratings they used, ratings were normalized by converting them into z-scores for each subject, and the mean of these z-scores is displayed. While all adults participated in the rating test, only twelve older and ten younger children were able to participate. The remaining children missed the rating test (e.g., a parent picked them up early) or did not participate due to experimenter error. The three age groups behaved similarly. Recall that two of the determiner manipulations involved forms that were legal in the language (dominant and minority forms) and two involved forms that never appeared in the input (omitting a determiner or using a novel determiner). All three age groups showed a clear difference in their ratings of those forms that are legal and those forms that are not. In all cases, illegal forms have a negative z-score, while legal forms have a positive one.

We tested this pattern statistically using ordinal regression, in which we analyzed rating (as an ordered factor from 1 to 4) with determiner Type (dominant, minority, novel, or omitted) and Age Group (adults, older children, younger children) as fixed effects and by-participant random intercepts and random slopes for Type. There was no main effect of Age Group ($\chi^2(2) = 0.02$, $p = .99$). However, there was a significant main effect of Type ($\chi^2(3) = 534.90$, $p < .001$), with both the dominant and minority determiners being significantly more likely to be highly rated than omitted determiners (dominant: $\beta = 4.30$, $SE = 0.28$, $z = 15.50$, $p < .001$; minority: $\beta = 3.70$, $SE = 0.26$, $z = 14.46$, $p < .001$). Participant ratings for novel determiners were not significantly different from omitted determiners ($\beta = 0.13$, $SE = 0.20$, $z = 0.63$, $p = .53$). Overall, then, the ratings suggest that the learners of all ages are aware of both the dominant and the minority forms. There was also a significant interaction between Age Group and Type ($\chi^2(6) = 52.19$, $p < .001$). As shown in Figure 3, the overall pattern of rating is the same for all age groups, with the two attested determiners (dominant and minority) rated higher than the two unattested determiners (novel and blank), but this difference is larger for adults than for young children.

Discussion

In accord with previous results, young children and adults behaved quite differently when they were taught a language with inconsistent variation. Adults reproduced this variation, in terms of how often
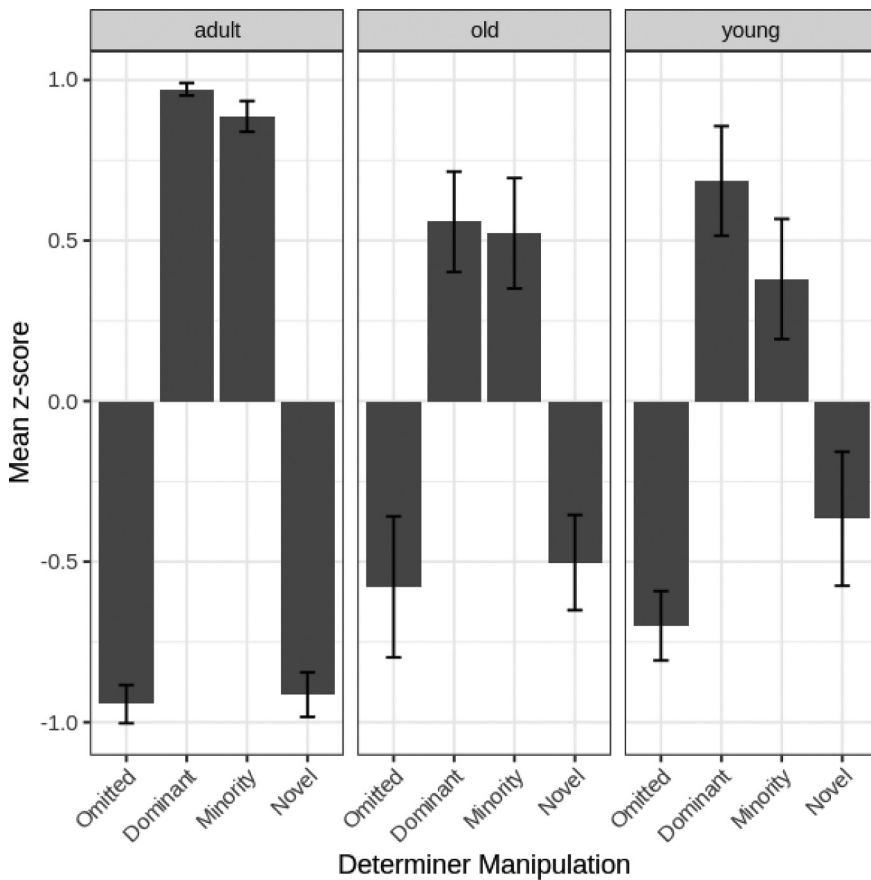
Figure 3. Mean z-score for each type of determiner item in the grammaticality rating task, for three age groups in the 60%-40% inconsistent language (Experiment 2).

each determiner was produced as well as in the lack of predictability or rule regarding when each determiner form was produced. In contrast, younger children used the dominant determiner more frequently than was in their input, and many (6 of 15) introduced strong rule-like patterns of determiner use. The older children behaved more like adults: they produced the dominant determiner as often as it appeared in their input; they did not introduce regularity as the younger children did. They did not regularize the dominant determiner, and only one had a rule-like production pattern (in this case a more complicated type of rule, using one determiner for subject nouns and another for object nouns). Overall these findings, while tentative due to small sample size, suggest that by age 7 or 8 children may show differences from young children in the way they learn inconsistent input, with many already learning miniature languages in a similar manner to adults.

This pattern of production was not mirrored in the grammaticality ratings. Despite rarely using the minority determiner in their own productions, young children appear to recognize it as an acceptable form of the determiner: they rated it almost as highly as the majority determiner that they used so frequently in production. These results suggest that children are not regularizing the dominant determiner in their productions simply because they have failed to store the minority form from their input.

One possibility is that regularization is entirely a production phenomenon. Another possibility, however, is that production – a more difficult and resource-demanding behavior than rating – more closely reflects the relative strengths and weaknesses of learners' knowledge of the determiner forms.

On this account, the dominant determiner may not be the only determiner form that is learned from inconsistent input, but it may be learned substantially better due to the relative inconsistency and low frequency of the minority form.

Although these results are suggestive, it is also important to keep in mind that children had difficulty understanding the grammaticality rating task (see Gleitman & Gleitman 1970; Gleitman et al., 1973; and other findings on children's difficulties with metalinguistic judgments) and, as we will see, show different patterns of results in ratings across studies here and in previous work.

Experiment 3

In each of the previous experiments and consistent with other work (e.g., Hudson Kam, 2015; Hudson Kam & Newport, 2005, 2009 Samara et al., 2017), children presented with inconsistent input have regularized. But how far does this phenomenon extend? That is, how low an input frequency will still be regularized? Is the dominant form – the one that is regularized – whatever form is relatively more frequent than others in a given context, or is there some threshold for regularization? One result from deaf children learning ASL from parents who are not fluent suggests that, at least in natural language learning, forms can be used surprisingly infrequently – with great inconsistency in the other forms alternatively used in the same obligatory contexts – and yet still be regularized by the young native signing child. For example, Ross's (2001) subject Sarah learned ASL entirely from her single hearing mother, who had learned ASL to a limited degree and produced ASL motion morphemes only about 15% of the time when required; the other 85% of her productions were highly variable and inconsistent. But Sarah at age 10 produced 76% consistent ASL motion morphemes. In spoken language, new varieties of English have been argued to emerge by precisely this process (Cheshire et al., 2011; Kerswill & Williams, 2000; Trudgill et al., 2000). For example, Trudgill et al. (2000) argue that New Zealand English is the result of children selecting the dominant form from inconsistent parent input across English, Scottish, Irish, Welsh, and Australian.

One way to begin to address this question in our experimental paradigm is to ask whether children will continue to regularize the dominant form, even if that form no longer appears in the majority of contexts. In Experiment 3, the determiner distribution was further manipulated in precisely this way. Here the dominant determiner appeared in only 40% of its contexts (that is, in 40% of the noun phrases), a proportion that is far lower than in any of our previous experiments. Indeed, this is the probability with which the *minority* determiner was used in Experiment 2. If young children are sensitive to the absolute frequency with which a determiner is used, they should produce the *dominant* determiner in this experiment with roughly the same (very low) probability as they produced the *minority* determiner in Experiment 2. On the other hand, if children are sensitive to the *relative* distribution of determiners (and not their absolute frequency), then young children should treat the dominant determiner in the current experiment as they have the dominant determiner in our other inconsistent languages, by greatly regularizing its use.

How do we lower the probability of the dominant determiner to under 50%? We did this by including an even lower frequency minority determiner (here 20%) as well as numerous "scatter" determiners, each occurring 4%, to fill out the rest of the distribution. To ensure that the addition of scatter determiners *per se* would not alter the results, we first ran a version with 60% dominant determiner, 32% minority determiner, and 8% scatter. The results of the 60% – 32% – scatter experiment were the same as the inconsistent language of Experiment 1: young children regularized their input and used the dominant determiner most of the time. We then proceeded here to test younger children, older children, and adults on a 40–20-scatter determiner language, examining both production and ratings as in Experiment 2. Note that, under these more complex circumstances, we also predict that adults should make the language more regular – though not to the same extent that young children will. This may be reflected in somewhat increased use of the dominant form, or in the number of scatter determiners they produce.

Methods

Participants

Eleven younger children (mean = 5.68, range = 4.96– 6.57), nine older children (mean = 7.92 years, range = 7.30– 8.61), and eight adult controls (mean = 20.27, range = 18.45– 23.38) participated in this experiment. An additional three younger children started the experiment but stopped participating before production testing had begun. All participants were recruited and compensated as in Experiment 1.

Description of the language

Other than the number and distribution of the determiners, the language in Experiment 3 was identical to that used in the Experiment 2. Unlike the previous three experiments, the language in Experiment 3 included 12 determiners: *ka, po, ny, gaw, lee, woo, vay, ba, tow, moy, dou*, and *suh*. In order to make forms maximally distinctive, each determiner had a unique consonant and vowel. *Ka* and *po* were always the dominant and minority forms; the remaining determiners were always the scatter forms. All determiner forms were distributed in an inconsistent manner. The dominant form of the determiner was used in 40% of noun phrases and the minority form in 20% of the noun phrases. The scatter forms (the remaining 10 other determiner forms) were each used in 4% of the noun phrases (once in each syntactic position per day). Aside from these ratios, determiner usage was completely probabilistic and unpredictable. Which determiner was dominant and which was minority was counterbalanced across subjects.

Stimuli and procedure

The set of visual scenes used for exposure, production testing, and grammaticality ratings were the same as in Experiment 2. The sentences presented along with each of the video clips changed in accordance with determiner distribution of the current language. The grammaticality rating task was altered to include an additional determiner manipulation, that of a scatter determiner. In all other ways the procedure was unchanged.

Results

Vocabulary testing

As in the previous experiments, all participants performed well on the vocabulary test. All adults and older children named all 5 nouns correctly. Young children knew an average of 4.82 of 5 nouns by the final day of testing (sd = 0.60).

Production data

Figure 4 shows the production of the dominant determiner for each age group. Data were scored as in previous experiments. Adults used the dominant determiner in 44.52% of productions (SEM = 3.47%), the minority determiner in 20.65% of productions (SEM = 3.02%), and scatter determiners in 34.82% of productions (SEM = 5.11%). For adults, then, determiner productions matched the input probabilities; even the proportion of scatter determiners produced was similar to the proportion in their input ($t(7) = 1.30$, $p = .23$). In contrast, younger children's productions were again quite different than their input: they used the dominant determiner in 81.10% (SEM = 6.65%) of productions, the minority determiner in 5.50% (SEM = 2.65%) of productions, and scatter determiners in 13.40% (SEM = 4.47%) of productions. As in previous experiments, then, the younger children greatly regularized, using the dominant determiner much more frequently than it occurred in the input ($t(10) = 6.18$, $p < .001$). This is particularly notable here since the dominant determiner was only present in 40% of the input sentences. The older children were in between: they used the dominant determiner in 49.45% of productions (SEM = 2.52%). Although this is significantly different from the input ($t(8) = 3.76$, $p < .01$), it comes nowhere near the degree of regularization exhibited by the younger children. Older children use the minority determiner in 23.86% of all productions (SEM = 3.70%), which
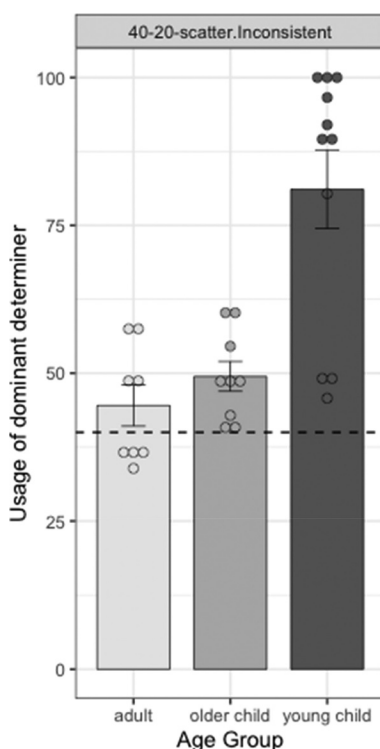
Figure 4. Mean production of dominant determiner in the 40–20-scatter language (Experiment 3) across 3 age groups. Dots are individual participants; error bars are standard error. Dashed line indicates level of the dominant determiner present in the input.

matches their input quite closely. The remaining 26.69% (SEM = 5.01%) of their productions were scatter determiners.

To ask whether use of the dominant determiner differed significantly by age, we ran a mixed-effect logistic regression predicting use of the dominant determiner by Age Group (adult vs. older child vs. young child) and Day and by-participant random intercepts.[6] Our results suggest that use of the dominant determiner was significantly greater than chance overall (where chance = 50%, β = 0.53, SE = 0.17, z = 3.08, p < .01). There was a main effect of Age Group, with young children but not older children using the dominant form significantly more than adults (β_{young} = 2.07, SE_{young} = 0.43, β_{old} = 0.21, SE_{old} = 0.42, χ^2 (2) = 21.77, p < .001). We ran a second model, releveled to test whether younger children were significantly different than older children (reference level: older children) and found that young children used the dominant form significantly more than older children (β = 1.87, SE = 0.42, z = 4.44, p < .001[7]). There was a main effect of Day, with log-odds of using the dominant form significantly decreasing as testing day increased (β = −0.27, SE = 0.08, χ^2 (1) = 8.66, p < .01). However, an interaction between Age Group and Day revealed that the effect of Day depended on Age Group, with young children but not older children significantly differing from adults (β_{young} = −0.47, SE_{young} = 0.20, β_{old} = 0.02, SE_{old} = 0.17, χ^2 (2) = 7.25, p < .03).

As in Experiments 1 and 2, we next asked whether participants' choice of determiner followed a rule-like pattern (e.g., by using the dominant, minority or scatter form on all their nouns, always conditioning on lexical item, or forming another rule-like pattern.). First, we examined the productions of each participant to determine whether a consistent response pattern appeared for that

---

[6]We attempted to conduct model comparison to find the best fitting random effects structure, but the model allowing day to vary by participant (isDominant ~ 1 + (1 + Day | Subject)) resulted in a singular fit
[7]We applied Bonferroni correction here to correct for multiple comparisons.

participant, and then coded whether each participant followed a rule (that is, followed the rule in all, or all but one, of their productions). By the final day of testing, no adults or older children had formed a rule, but 4 of 11 young children had fully regularized the dominant form.

As in Experiments 1 and 2, we also quantified the degree to which participants regularized using entropy, where 0 indicates productions are completely regular (i.e. the participant used a single determiner on all trials). First, we calculated the entropy of the input language to be 1.98 (40% dominant, 20% minority, and 4% each of 10 scatter determiners). Then, we calculated the entropy of each participant's productions on each day of testing and used mixed-effects regression to predict entropy by Age Group and Day with by-participant random intercepts. Contrary to the model testing use of the dominant form, there was no main effect of Day or interaction with Day (all p's > 0.24) in this model. There was, however, a significant main effect of Age Group, in which young children but not older children had significantly lower entropy than adults (β_{young} = −0.61, SE_{young} = 0.10, β_{old} = −0.05, SE_{old} = 0.11, χ^2 (2) = 28.96, p < .001) – that is, young children's productions were significantly more regular than the input (input entropy = 1.98) – with a mean entropy on the final day of testing of 0.43 for young children, 0.92 for older children, and 0.95 for adults. In a second model we tested the specific comparison between older children and younger children and found that younger children regularized significantly more than older children (β = −0.55, SE = 0.10, t(25.06) = −5.49, p < .001[8]). This suggests that while the youngest children may have used the dominant form less on the final day of testing (as evidenced by the effects of Day in that model), they continued to regularize in their productions (as evidenced by no effect of Day in our analysis of entropy here).

Finally, of particular interest in this experiment, with a 40–20-scatter Language, is whether children treated the dominant determiner – which was heard in 40% of exposure noun phrases and was the most consistently used form – differently than the minority determiner in Experiment 2 (with a 60–40 Language), which was also heard in 40% of exposure noun phrases but was not the most consistently used form. The data are shown in Figure 5. To test this, we analyzed use of the 40% determiner across experiments, predicting use of the determiner by Age Group (adult, older child, young child), Experiment (2 v 3), and Day. We used model comparison to determine the best fitting random effects structure and found that model fit was significantly improved by allowing day to vary by participant (χ^2 (2) = 17.73, p < .001). Our results indicate that use of the 40% determiner was not significantly different from chance overall (where chance = 50%, β = −0.12, SE = 0.13, z = −0.90, p = .37). There were no significant main effects of Day (χ^2 (1) = 2.04, p = .15) or Age Group (χ^2 (2) = 1.02, p = .60) and no significant interactions with Day (all p's>0.05). However, we did find a significant main effect of Experiment, with participants using the 40% form significantly more in Experiment 3, when it was the dominant determiner, than in Experiment 2, when it was the minority determiner (β = 1.35, SE = 0.25, χ^2 (1) = 15.48, p < .001). A significant interaction between Age Group and Experiment that young children but not older children used the 40% form significantly more than adults when it was the dominant compared to when it was the minority determiner (β_{young} = 2.92, SE_{young} = 0.63, β_{old} = 0.15, SE_{old} = 0.15, χ^2 (2) = 25.13, p < .001).

In sum, then, the results in 40–20-scatter are very similar to those in the earlier experiments, with young children regularizing to a surprising degree here even though the dominant determiner was both inconsistent and relatively low frequency. Further, young children regularized this 40% determiner here (where it was the majority determiner), but they did not regularize a 40% determiner in Experiment 2 (where it was the minority determiner). This suggests that young children's regularization is not dependent on the absolute frequency of input forms, but rather on the contrastive use of forms within each language.

---

[8]We releveled the model with older children as the reference and applied Bonferroni correction here to correct for multiple comparisons.

Grammaticality ratings

Grammaticality ratings for this language are shown in Figure 6. Unfortunately, while all adults participated in the rating test, only two older and eight younger children were able to participate. The remaining children missed the rating test (e.g., a parent picked them up early) or were not tested due to experimenter error. As such, the results from this rating test, particularly those with older children, should be interpreted with caution.

We tested whether ratings differed by Age Group and determiner Type using ordinal regression, in which we analyzed rating (as an ordered factor from 1 to 4) with determiner Type (dominant, minority, scatter, novel, or omitted) and Age Group (adults, older children, younger children) as fixed effects and by-participant random intercepts. There was a no main effect of Age Group (χ^2(2) = 1.26, p = .53), indicating that all three age groups showed similar patterns of rating overall. There was a main effect of Type (χ^2(4) = 356.20, p < .001), in which all attested determiner types (dominant, minority, scatter) were rated significantly higher than novel determiners (all p's < 0.001) while the omitted sentence type was rated significantly lower than the novel determiners (p < .01). There was also an interaction between Type and Age Group (χ^2(8) = 60.36, p < .001).

These results reflect what we see in Figure 6: Adult participants rate all three occurring forms of the determiner (dominant, minority, scatter) higher than the two forms that never appeared in the input
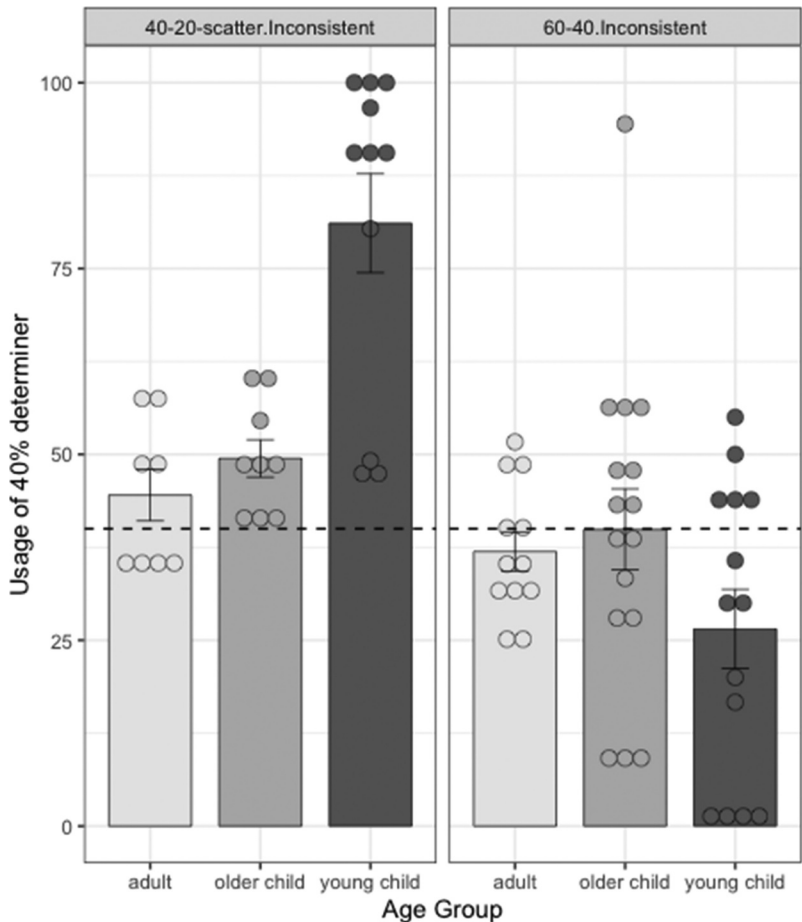

Figure 5. Mean production of the 40% determiner in the 40–20-scatter language (Experiment 3) and in the 60–40 language (Experiment 2) by age group. Dots are individual participants; error bars are standard error of the mean. The dashed line indicates the presence of the determiner in the input (40%).

(omitted, novel). The two forms that occurred more frequently were rated more highly than the less frequent scatter forms. In contrast, the ratings from the younger children are more equivocal. The ratings for the dominant and scatter forms are significantly higher than those for omitted determiners and novel determiners, but the rating of the minority determiner is much lower and not different from either of those. While adults and older children rate dominant and minority forms equally highly, young children rate the minority form substantially lower than the dominant one. This follows the pattern seen in their productions.

Discussion

In this experiment we found further evidence that young language learners regularize inconsistent linguistic input, even when the dominant determiner is relatively infrequent, while adult learners do not. As in our previous experiments, young children used the dominant determiner with much greater frequency and consistency than was present in their input. The 40–20-scatter distribution demonstrates the robustness of this phenomenon. Despite the much lower use of the dominant determiner in the input (40% here vs. 60% or 67% in previous experiments), younger children learning the 40–20-scatter condition regularized to almost the same degree as previously. When learning a language with a 67%-33% determiner split, young children used the dominant determiner in 80% of all productions; when learning a language with a 60%-40% determiner split, young children used the dominant determiner in 72.43% of all productions. In the current language, they used it in 81.10% of all productions. Despite having less evidence of its use (40% vs. 67% or 60%), young children used it with essentially the same degree of consistency. These results demonstrate that young children do not
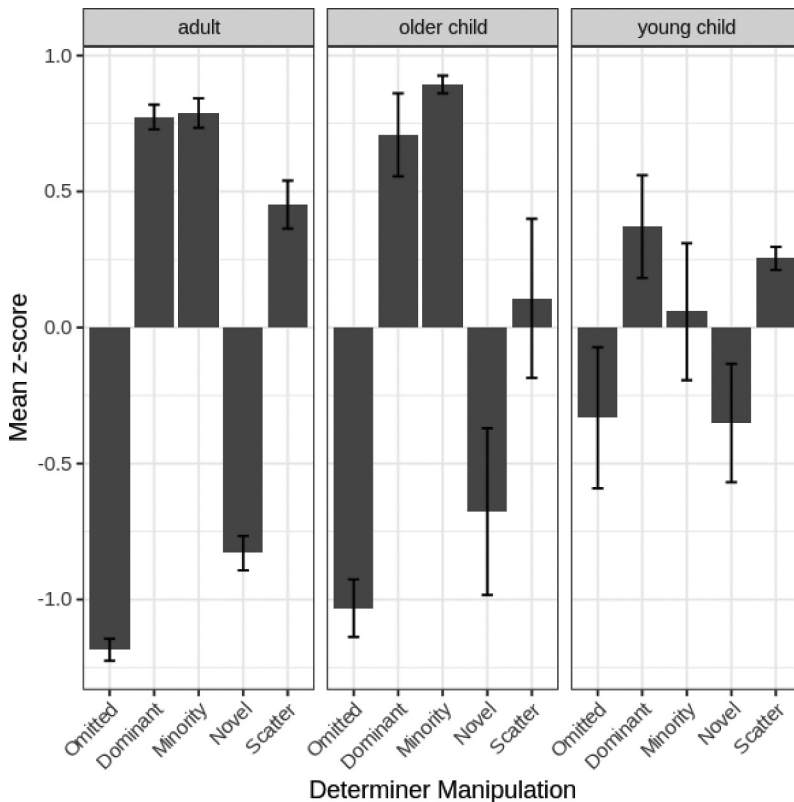

Figure 6. Mean z-score for each type of determiner item in the grammaticality rating task, for three age groups in the 40–20-scatter language (Experiment 3).

regularize the dominant form based on its use in the majority of input contexts or based on its absolute frequency of use in the input. The frequency of the dominant form in the exposure corpus of this experiment (40%) was equal to that of the minority form in Experiment 2, but in Experiment 2 young children seldom produced the minority form. Apparently young children regularize the form that is dominant relative to other forms used in the same contexts – and then use that form in a highly consistent way, regardless of its absolute input frequency.

In contrast, older children are more like adults: they use the dominant determiner significantly more often than it was in the input – which adults do not – but significantly and substantially less than young children do. Adults again match the frequency of the dominant determiner in their input.[9]

The grammaticality ratings also revealed an interesting pattern. Adults and older children rated the dominant and minority determiner more highly than the two forms they had not heard during exposure (novel determiner and omitted determiner), with scatter determiners rated between. However, younger children rated the dominant determiner significantly higher than the minority determiner, corresponding to their productions, whereas the older children and adults rated the dominant and minority determiners similarly.

General discussion

In previous natural and artificial language studies (Hudson Kam, 2015; Hudson Kam & Newport, 2005, 2009; Ross, 2001; Ross & Newport, 1996; Samara et al., 2017; Sankoff & Laberge, 1973; Senghas & Coppola, 2001; Singleton & Newport, 2004), children exposed to inconsistent linguistic input did not acquire these inconsistencies, but instead changed the language, forming regular rules instead of inconsistent usages. Such a phenomenon in the formation of pidgin and creole languages and in historical language change has been hypothesized to account for the formation of new grammatical structures (Fischer, 1978; Culbertson & Newport, 2015, 2017; Kiparsky, 1970; Newport, 1999; Roberts, 1999). In the present experiments we have brought this phenomenon into the lab in order to understand the circumstances under which it occurs and to gain insight into what the underlying mechanisms might be.

In the inconsistent language of Experiment 1, a simplified and improved training procedure used with 5-year-old children and adults showed that, while adults reproduced determiners in the same inconsistent 67%-33% alternation of their input, children produced the dominant determiner almost exclusively and used it consistently, in a rule-like way. In contrast, in the lexically consistent language, when the input contained determiners at the same probabilities but distributed in a deterministic way – when 67% of the nouns were always used with one determiner and 33% of the nouns were always used with another determiner (as in a gender system) – both children and adults acquired and reproduced the lexical rules. These results show that children produce forms in a consistent and regular way; when the input usage is not consistent, they regularize it in their own productions. The results of the lexically consistent language also showed that young children's regularizations are not merely due to their failure to acquire low frequency minority forms: children can acquire low frequency forms, as long as they are used in a consistent fashion.

Experiment 3 showed that, even when the dominant form was used infrequently – in a 40%-20%-scatter distribution – young children regularized the dominant form. Apparently, children do not regularize because the minority form is heard too infrequently to be learned or because they need more exposure to it than adults. Rather, regularization seems to be driven by specific characteristics of input variation. The contrast between Experiment 2 (where the *minority* determiner occurred in 40% of the

---

[9]Adult learners do not regularize substantially here, although Hudson Kam & Newport (2009) found that adults did regularize more when presented with a 60-scatter distribution. There are two important differences between the scatter distribution of Experiment 3 and that in Hudson Kam & Newport: Here the majority determiner was only present 40% of the time (rather than 60%), and here there was also a 20% minority determiner.

noun phrases but was not regularized) and Experiment 3 (where the *majority* determiner occurred in 40% of the noun phrases and was regularized) suggests that regularization occurs for the relatively dominant form and not on the basis of its frequency per se.

Altogether these results corroborate and refine the claim that children change inconsistent languages, making them more regular. What are the mechanisms that underlie this tendency of young children, and when does it happen? We turn now to the suggestions found in our results regarding these issues.

Developmental changes

Experiments 2 and 3 also provide some evidence about when children begin to lose their tendency to regularize inconsistent input. Surprisingly, the clear difference between child- and adult-like productions begins to diminish as early as 7- to 8-years-old. There are several cognitive processes undergoing maturational change over this age range that could be relevant.

What causes regularization?

One possible hypothesis about why young children regularize inconsistent input is that, in the view of some investigators, children have special and specific knowledge about the properties of natural language that underlies their ability to learn languages so well (Chomsky, 1965; Hornstein, 2009). On this view, children innately expect languages to contain deterministic rules and may impose this expectation on their input; as they mature, this knowledge and the associated ability to learn languages declines. While this is not our own view, it is compatible with the present results.

In contrast, several investigators have suggested that developmental changes in cognitive abilities, not specific to language, may underlie changes in the ability to acquire language – and in particular, that the cognitive limitations of very young children may be essential ingredients in their remarkable language learning abilities (Elman, 1993; Goldowsky & Newport, 1993; Newport, 1990). This view too is compatible with the present results. An important question here is what particular cognitive abilities might be responsible for regularization.

Several studies have asked whether working memory capacity (or some other aspect of memory entailed in storing linguistic input) plays a role in regularization. On such a hypothesis, children may store the dominant form in memory more effectively due to its more consistent usage; but as they mature, they are better able to store multiple forms involved in linguistic variation. There are a number of experimental supports for this hypothesis. As the complexity of variation in a language increases, or when given a concurrent capacity-limiting task to perform, even adults begin to regularize in a way similar to young children (Cochran et al., 1999; Hudson Kam & Newport, 2009). This effect goes away when the task is made easier (Hudson Kam & Chang, 2009). However, purely increasing the difficulty of the task does not induce adults to regularize as children do (Perfors, 2012; Perfors & Burns, 2010), nor has it been shown that young children regularize less when the difficulty of the task is decreased (though see Samara et al. (2017), in which children regularized less as their exposure to a much simpler language increased) . Perfors has suggested that cognitive limitations must be combined with a prior bias toward regularization in order to obtain the regularization effects we have found in young children.

Another possibility is inhibitory control, which is also still developing at this age and has been shown to affect other aspects of linguistic processing in children (Trueswell et al., 1999). Developmental changes in inhibition of the dominant form could be a source of the behavioral differences seen here: not in whether children can learn the non-dominant form of the determiner, but in how well they can control their impulse to produce the dominant form.

Yet another possibility is that children may consolidate or "bin" the statistics of linguistic information over smaller portions of the input; in particular, young children may store and learn the input in a lexically-based manner, while older children and adults may combine input information over a class of lexical items. Such a difference could explain the ease with which young children learned the lexically-consistent language in Experiment 1, which adults learned more slowly. Such a lexically-

based account is consistent with Tomasello (2000) and others, who have suggested that children's knowledge of syntax is initially item-based (though see Fisher, 2002, for an opposing view).

In inconsistent languages, determiner variation was explicitly designed to hold across various levels of the language; the same variation was present for all nouns, verbs, grammatical roles, etc. However, if children bin the data separately for individual lexical items, there will be less information in each bin; and this loss of information may lead produce greater uncertainty and might play a role in the tendency to regularize. Older learners may have a greater ability to calculate the distributions in different ways, either because they can consider different distribution factors simultaneously or because they can track more factors that condition the distribution. This is consistent with studies of adult language processing showing that adult speakers are sensitive to complex contingencies in their input (e.g., Amato & MacDonald, 2010; Bicknell, et al 2008).

Finally, another possibility is that children and adults use different learning strategies. For example, perhaps adults are testing explicit hypotheses and children are learning more implicitly. While we did not test the distinction between explicit and implicit learning in the present work, we often observe differences between adults and children in verbalizing their hypothesized rules. It is important to note, however, that while adults are often quite talkative during sentence exposure, attempting to guess the rule behind the determiner distribution, they do not necessarily produce or judge sentences in accord with their stated hypotheses, so it is not clear that their articulated strategies are the ones they are in fact using during learning. Moreover, the older children are more like younger children in this regard, with neither group talking aloud about hypotheses regarding determiner distribution. Though of course not conclusive, these anecdotes give the impression that differences in learning strategies may not be the main reason for the change in production over age.

Production versus ratings

The results from the current studies also highlight an important distinction between results from production tasks, which show effects of regularization, and perception tasks, which do not always show the same effects. One should proceed with caution in concluding too much from the grammaticality rating data, since the results across studies here and elsewhere (see also Hudson Kam & Newport, 2005, 2009) do not show a consistent pattern. However, it does seem clear from ratings that children are not simply failing to retain any knowledge about the low frequency determiner forms, since they are (in at least some of our studies) able to judge correctly that these are more acceptable than forms never heard in their input. Children's sensitivity to aspects of the input that are absent in their own productions is attested in other areas of language acquisition (Gerken et al., 1990). However, we do not believe that regularization is a "mere" production effect. Even if regularization occurs only in production, it would nonetheless affect language change. For example, in natural language regularization, such as the emerging Nicaraguan Sign Language or children learning from non-native input, productions from one generation serve as input to the next generation of learners.

In addition, as we have suggested above, production is a more difficult and resource-demanding behavior than ratings and therefore may more closely reflect the relative strength and weakness of learners' knowledge of determiner forms. On this account, the dominant determiner may not be the only determiner form that is learned from inconsistent input, but it may be learned substantially better than other forms, due to the greater inconsistency and lower frequency of minority forms. An important test of this hypothesis would be to examine children's performance with dominant and minority forms in other demanding tasks.

What do these findings say about natural language acquisition?

The current studies show that children regularize inconsistent input, a type of variation primarily present only in rare cases of language acquisition. An important question is whether this tendency is also present in more typical cases of language development. For example, what about learning the

variation in forms that is conditioned by lexical exceptions (as in the regular and irregular past tense forms in English) or by grammatical context (as in gender marking)? In these cases, a vast literature suggests that children do regularize the dominant form in very early acquisition but then acquire correct irregulars as they mature (e.g., Berko 1958; Brown, 1973; Kuczaj, 1977; Marcus et al., 1992 on -*ed* overregularization). In natural language circumstances with more inconsistent input – such as in creole communities or during periods of historical change – are children's regularizations only temporary? It is important to note the striking difference between the inconsistencies of well-developed languages and those in the linguistic environments of children like Singleton & Newport's Simon (2004), who are learning their native languages from non-native parents. The appearance of -*ed* in English, for example, is inconsistent across verbs; but each verb in a past tense context is consistently marked with -*ed* or with its own lexically exceptional past tense marker. This somewhat confusing pattern likely leads to regularization until children have learned the language well enough to notice and acquire the distinctions between individual verbs. As we saw in our lexically consistent language, under such conditions children do not regularize; rather, they are able to learn which lexical items go with which forms. In contrast, Simon's input from his non-native signing parents are more like our inconsistent input experiments, where forms are used in a truly inconsistent way, with no conditioning context to be noticed and learned (Singleton & Newport, 2004). This is the type of circumstance in which children's tendency to regularize can permanently alter the structure of the language.

An additional question is how these results relate to learning sociolinguistic variation that is present in natural languages. For example, children do learn to reproduce phonetic variation present in their caregivers' speech (Labov, 1989; Smith et al., 2007), and some research has suggested that this variation is acquired very early (Smith et al., 2013; but see Miller, 2013, for conflicting results). However, recent work by Samara et al. (2017) finds that young children tend to regularize grammatical variation that is inconsistent or conditioned on speaker identity. Ongoing research in our lab finds that young children first regularize and only later acquire the conditions as well as the forms of sociolinguistic variation (Sneller & Newport, 2020); these results on phonological variation are much like what we find in the experiments here for grammatical variation.

Much more work is required in order to fully address these issues. In the meantime, the studies presented here have provided insight into one reason children may be more adept language learners than adults: they have a striking tendency to regularize their linguistic input, altering the statistics from their original form. This tendency is present only in younger children and seems to diminish quickly once children reach school-age. Paradoxically, the inability to learn certain types of inconsistent statistics may make young children better able to learn language. Natural languages of the world include many types of variation: unconditioned regular rules, rules that are perfectly or partially conditioned by context, and – perhaps predominantly in emerging languages or at times of heavy language contact and change – forms that are used in highly inconsistent ways. According to the present results, young children begin learning by acquiring and producing the most regular forms – the dominant form in a given context – but when forms are used deterministically, young children are capable of acquiring the conditioning. More work is needed to determine how and when they go on to acquire the more complex conditions under which multiple variants may occur, such as probabilistic conditioning in sociolinguistic variation. This bias toward regularizing may be a key process in simplifying and sequencing parts of natural language acquisition and may help to explain why children are better language learners than adults. Perhaps most important, in circumstances of highly inconsistent and emerging languages, children may indeed be the forces that create regular linguistic systems.

Acknowledgments

the review and proofing process.

Disclosure statement

Funding

ORCID

Kathryn D. Schuler http://orcid.org/0000-0003-2962-731X
Sarah Furlong http://orcid.org/0000-0003-0683-7197
Elissa L. Newport http://orcid.org/0000-0002-2783-5062

References

Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory*, 21(3), 435–483.

Amato, M. S., & MacDonald, M. C. (2010). Sentence processing in an artificial language: Learning and using combinatorial constraints. *Cognition*, 116(1), 143–148.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, 9(4), 321–324. https://doi.org/10.1111/1467-9280.00063

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. arXiv preprint arXiv:1506.04967.

Berko-Gleason, J. (1958). The child's learning of English morphology. *WORD*, 14(2–3), 150–177. https://doi.org/10.1080/00437956.1958.11659661

Bickerton, D. (1984). The language bioprogram hypothesis, plus commentaries and author's response. *Behavioral and Brain Sciences*, 7(2), 173–221. https://doi.org/10.1017/S0140525X00044149

Bicknell, K., Demberg, V., & Levy, R. (2008). Local coherences in the wild: An eye-tracking corpus study. Language, 54, 363–388.

Boudreault, P., & Mayberry, R. I. (2006). Grammatical processing in American sign language: Age of first-language acquisition effects in relation to syntactic structure. *Language and Cognitive Processes*, 21(5), 608–635. https://doi.org/10.1080/01690960500139363

Brown, R. (1973). *A first language: The early stages*. Harvard University Press.

Chambers, K. E., Onishi, K. H., & Fisher, C. (2003). Infants learn phonotactic regularities from brief auditory experience. *Cognition*, 87(2), B69–B77. https://doi.org/10.1016/s0010-0277(02)00233-0

Cheshire, J., Kerswill, P., Fox, S., & Torgersen, E. (2011). Contact, the feature pool and the speech community: The emergence of multicultural London English. *Journal of Sociolinguistics*, 15(2), 151–196. https://doi.org/10.1111/j.1467-9841.2011.00478.x

Chomsky, N. (1965) Aspects of the Theory of Syntax, Cambridge, Massachusetts: MIT Press.

Cochran, B. P., McDonald, J. L., & Parault, S. J. (1999). Too smart for their own good: The disadvantage of a superior processing capacity for adult language learners. *Journal of Memory and Language*, 41(1), 30–58. https://doi.org/10.1006/jmla.1999.2633

Culbertson, J., & Newport, E. L. (2015). Harmonic biases in child learners: In support of language universals. *Cognition*, 139, 71–82. https://doi.org/10.1016/j.cognition.2015.02.007

Culbertson, J., & Newport, E. L. (2017). Innovation of word order harmony across development. *Open Mind: Discoveries in Cognitive Science*, 1(2), 91–100. https://doi.org/10.1162/OPMI_a_00010

Culbertson, Jennifer & Gagliardi, Annie & Smith, Kenny. (2017). Competition between phonological and semantic cues in noun class learning. Journal of Memory and Language. 92, 343-358. doi:10.1016/j.jml.2016.08.001.

Culbertson, J., Jarvinen, H., Haggarty, F., & Smith, K. (2019). Children's sensitivity to phonological and semantic cues during noun class learning: Evidence for a phonological bias. Language 95(2), 268-293. doi:10.1353/lan.2019.0031.

Chappell, H., & Verstraete, J. C. (2019). Optional and alternating case marking: Typology and diachrony. *Language and Linguistics Compass*, 13(3), e12311.

D'Arcy, A., & Tagliamonte, S. A. (2015). Not always variable: Probing the vernacular grammar.*Language Variation and Change*,*27*(3), 255–285.

DeGraff, M. (1999). *Language creation and language change: Creolization, diachrony, and development (introduction and conclusion)*. MIT Press.

Elman, J. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, *48*(1), 71–99. https://doi.org/10.1016/0010-0277(93)90058-4

Fischer S.(1978), Sign language and creoles. In P. Siple (Ed.), *Understanding language through sign language research*. Academic Press.

Fisher, C. (2002). The role of abstract syntactic knowledge in language acquisition: *A reply to Tomasello.Cognition*, 82, 259–278.

Ferdinand, V., Kirby, S., & Smith, K. (2019). The cognitive roots of regularization in language. Cognition, 184, 53–68.

Gerken, L., Landau, B., & Remez, R. (1990). Function morphemes in young children's speech perception and production. *Developmental Psychology*, *26*(2), 204–216. https://doi.org/10.1037/0012-1649.26.2.204

Gerken, L. A. (2004). Nine-month-olds extract structural principles required for natural language. *Cognition*, *93*(3), B89–B96. https://doi.org/10.1016/j.cognition.2003.11.005

Gerken, L. A. (2006). Decisions, decisions: Infant language learning when multiple generalizations are possible. *Cognition*, *98*(3), B67–74. https://doi.org/10.1016/j.cognition.2005.03.003

Gleitman, L., Gleitman, H., & Shipley, E. (1973). The emergence of the child as grammarian. *Cognition*, *1*(2), 137–164. https://doi.org/10.1016/0010-0277(72)90016-9

Goldowsky, B. N., & Newport, E. L. (1993). Modeling the effects of processing limitations on the acquisition of morphology: The less is more hypothesis. In J. Mead (ed.), *The proceedings of the 11th West Coast conference on formal linguistics*. Stanford, CA: CSLI.

Gomez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*(5), 431–436. https://doi.org/10.1111/1467-9280.00476

Gleitman, L. R., & Gleitman, H. (1970). Phrase and paraphrase: Some innovative uses of language.New York: W. W. Norton.

Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*,*70*(2), 109–135.

Hall, R. A. (1966). *Pidgin and creole languages*. Cornell University Press.

Hornstein, N. (2009). A theory of syntax: Minimal operations and universal grammar. Cambridge University Press.

Hudson Kam, C., & Newport, E. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology*, *59*(1), 30–66. https://doi.org/10.1016/j.cogpsych.2009.01.001

Hudson Kam, C. L. (2015). The impact of conditioning variables on the acquisition of variation in adult and child learners. *Language*, *91*(4), 906–937. https://doi.org/10.1353/lan.2015.0051

Hudson Kam, C. L., & Chang, A. (2009). Investigating the cause of language regularization in adults: Memory constraints or learning effects? *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *35*(3), 815–821.

Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, *1*(2), 151–195. https://doi.org/10.1080/15475441.2005.9684215

Kroch, A., & Small, C. (1978). Grammatical ideology and its effect on speech. Linguistic variation: Models and methods, 45755.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*(1), 60–99. https://doi.org/10.1016/0010-0285(89)90003-0

Kegl, J. A., Senghas, A., & Coppola, M. (1999). Creation through contact: Sign language emergence and sign language change in Nicaragua. In M. DeGraff (Ed.), *Language creation and language change: Creolization, diachrony, and development* (pp. 179–237). MIT Press.

Kerswill, P., & Williams, A. (2000). Creating a new town koine: Children and language change in Milton Keynes. *Language in Society*, *29*(1), 65–115. https://doi.org/10.1017/S0047404500001020

Kiparsky, P. (1970). Historical linguistics. In W. O. Dingwall (Ed.), *A survey of linguistic science*. (pp.576–642). University of Maryland Press.

Kocab, A., Senghas, A., & Snedeker, J. (2016). The emergence of temporal language in Nicaraguan sign language. *Cognition*, *156*, 147–163. https://doi.org/10.1016/j.cognition.2016.08.005

Koulaguina, E., & Shi, R. (2013). abstract rule learning in 11- and 14-month-old infants. *Journal of Psycholinguistic Research*, *42*(1), 71–80. https://doi.org/10.1007/s10936-012-9208-4

Kuczaj, S. (1977). The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior*, *16*(5), 589–600. https://doi.org/10.1016/S0022-5371(77)80021-2

Labov, W. (1989). The child as linguistic historian. *Language Variation and Change*, *1*(1), 85–97. https://doi.org/10.1017/S0954394500000120

Labov, W. (2007). Transmission and diffusion. *Language*, *83*(2), 344–387. https://doi.org/10.1353/lan.2007.0082

Marcus, G., Pinker, S., Ullman, M., Hollander, M., Rosen, T., & Xu, F. (1992). Overregularization in language acquisition. *Monographs of the SRCD*, Serial no. 228.

Mayberry, R. I., & Fischer, S. D. (1989). Looking through phonological shape to lexical meaning: The bottleneck of nonnative sign language processing. *Memory & Cognition*, *17*(6), 740–754. https://doi.org/10.3758/BF03202635

Maye, J., Werker, J. F., & Gerken, L. A. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), B101–11. https://doi.org/10.1016/S0010-0277(01)00157-3

McWhorter, J. H. (2005). *Defining creole*. Oxford University Press.

Miller, K. (2013). Acquisition of variable rules: /s/-lenition in the speech of Chilean Spanish-speaking children and their caregivers. *Language Variation and Change*, *25*(3), 311–340. https://doi.org/10.1017/S095439451300015X

Mintz, T. H., Newport, E. L., & Bever, T. G. (2002). The distributional structure of grammatical categories in speech to young children. *Cognitive Science*, *26*(4), 393–425. https://doi.org/10.1207/s15516709cog2604_1

Morgan, E., & Levy, R. (2016). Abstract knowledge versus direct experience in processing of binomial expressions. *Cognition*, *157*, 384–402. https://doi.org/10.1016/j.cognition.2016.09.011

Mufwene, S. S. (2007). Population movements and contacts in language evolution. *Journal of Language Contact*, *1*(1), 63–92. https://doi.org/10.1163/000000007792548332

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science*, *14*(1), 11–28. https://doi.org/10.1207/s15516709cog1401_2

Newport, E. L. (1999). Reduced input in the acquisition of signed languages: Contributions to the study of creolization. In M. Degraff (Ed.), *Creolization, Diachrony, and Language Acquisition*. MIT Press.

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance: I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*(2), 127–162. https://doi.org/10.1016/S0010-0285(03)00128-2

Perfors, A., & Burns, N. (2010). Adult language learners under cognitive load do not over-regularize like children. In S. Ohlsson & R. Catrambone (eds), *Proceedings of the 32nd annual conference of the cognitive science society* (pp. 2524–2529). Portland, Oregon.

Perfors, A. (2012). When do memory limitations lead to regularization? An experimental and computational investigation. *Journal of Memory and Language*, *67*(4), 486–506. https://doi.org/10.1016/j.jml.2012.07.009

Perfors, A. (2016). Adult regularization of inconsistent input depends on pragmatic factors. *Language Learning and Development*, *12*(2), 138–155. https://doi.org/10.1080/15475441.2015.1052449

Reali, F., & Griffiths, T. L. (2009). The evolution of frequency distributions: Relating regularization to inductive biases through iterated learning. *Cognition*, *111*(3), 317–328. https://doi.org/10.1016/j.cognition.2009.02.012

Roberts, J. (1999). Going younger to do difference: The role of children in language change. University of Pennsylvania Working Papers in Linguistics, 6(2), 10. Available at: https://repository.upenn.edu/pwpl/vol6/iss2/10 .

Ross, D. S., & Newport, E. L. (1996). The development of language from non-native linguistic input. In A. Stringfellow & D. Cahana-Amitay (Eds.) *Proceedings of the 20th annual Boston University conference on language development*. Cascadilla Press: Somerville, MA.

Ross, D. S. (2001). *Disentangling the nature-nurture interaction: The effects of exposure to non-native input on the language acquisition process*. Unpublished doctoral dissertation, University of Rochester.

Reeder, P. A., Newport, E. L., & Aslin, R. N. (2013). From shared contexts to syntactic categories: The role of distributional information in learning linguistic form-classes. *Cognitive Psychology*, *66*, 30–54.

Schuler, K. D., Yang, C., & Newport, E. L. (2016). Testing the Tolerance Principle: Children form productive rules when it is computationally more efficient to do so. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (Eds.) Proceedings of the 38th Annual Conference of the Cognitive Science Society (pp. 2321–2326). Austin, TX: Cognitive Science Society.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Samara, A., Smith, K., Brown, H., & Wonnacott, E. (2017). Acquiring variation in an artificial language: Children and adults are sensitive to socially conditioned linguistic variation. *Cognitive Psychology*, *94*, 85–114. https://doi.org/10.1016/j.cogpsych.2017.02.004

Sankoff, G. (1979). The genesis of a language. In K. C. Hill (Ed.), *The genesis of language*. Karoma Publishers, p. 32-47.

Sankoff, G., & Laberge, S. (1973). On the acquisition of native speakers by a language. *Kivung*, 6, 32–47. https://doi-org.proxy.library.upenn.edu/10.9783/9781512809589-014

Schwab, J., Lew-Williams, C., & Goldberg, A. (2018). When regularization gets it wrong: Children over-simplify language input only in production. *Journal of Child Language*, *45*(5), 1054–1072. https://doi.org/10.1017/S0305000918000041

Senghas, A., & Coppola, M. (2001). Children creating language: How nicaraguan sign language acquired a spatial grammar. *Psychological Science*, *12*(4), 323–328. https://doi.org/10.1111/1467-9280.00359

Singleton, J. L., & Newport, E. L. (2004). When learners surpass their models: The acquisition of American sign language from inconsistent input. *Cognitive Psychology*, *49*(4), 370–407. https://doi.org/10.1016/j.cogpsych.2004.05.001

Slobin, D. L. (1973). Cognitive prerequisites for the development of grammar. In D. L. Slobin & C. Ferguson (Eds.), *Studies of child language development*. Holt, Rinehart, & Winston, 175-208.

Smith, J., Durham, M., & Fortune, L. (2007). 'Mam, my trousers is fa'in doon! ': Community, caregiver, and child in the acquisition of variation in a Scottish dialect. *Language Variation and Change*, *19*(1), 63–99. https://doi.org/10.1017/S0954394507070044

Smith, J., Durham, M., & Richards, H. (2013). The social and linguistic in the acquisition of sociolinguistic norms: Caregivers, children, and variation. *Linguistics*, *51*(2), 285–324. https://doi.org/10.1515/ling-2013-0012

Smith, K., Perfors, A., Fehér, O., Samara, A., Swoboda, K., & Wonnacott, E. (2017). Language learning, language use and the evolution of linguistic variation. *Philosophical Transactions of the Royal Society B*, *372*(1711), 20160051. https://doi.org/10.1098/rstb.2016.0051

Smith, K., & Wonnacott, E. (2010). Eliminating unpredictable variation through iterated learning. *Cognition*, *116*(3), 444–449. https://doi.org/10.1016/j.cognition.2010.06.004

Sneller, B., & Newport, E. L. (2020). Acquisition of grammatically and socially conditioned phonological variation. *Poster presented at the Boston University Conference on Language Development*, Boston, MA, November 2020.

Tomasello, M. (2000). The item-based nature of children's early syntactic development. Trends in cognitive sciences, 4 (4), 156–163

Trudgill, P., Gordon, E., Lewis, G., & Maclagan, M. (2000). Determinism in new-dialect formation and the genesis of New Zealand English. *Journal of Linguistics*, *36*(2), 299–318. https://doi.org/10.1017/S0022226700008161

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, *73*(2), 89–134. https://doi.org/10.1016/S0010-0277(99)00032–3

Thompson, S. P., & Newport, E. L. (2007). Statistical learning of syntax: The role of transitional probability. *Language Learning and Development*,*3*, 1–42.

Thothathiri, M., & Rattinger, M. G. (2016). Acquiring and producing sentences: Whether learners use verb-specific or verb-general information depends on cue validity. *Frontiers in Psychology*,*7*, 404.

Wonnacott, E. (2011). Balancing generalization and lexical conservatism: An artificial language study with child learners. *Journal of Memory and Language*,*65*(1), 1–14.

Wonnacott, E., Newport, E. L., & Tanenhaus, M. K. (2008). Acquiring and processing verb argument structure: Distributional learning in a miniature language. *Cognitive Psychology*,*2008*(56), 165–209.

Weber-Fox, C., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, *8*(3), 231–256. https://doi.org/10.1162/jocn.1996.8.3.231