Published in final edited form as:

J Exp Psychol Learn Mem Cogn. 2009 May ; 35(3): 815–821. doi:10.1037/a0015097.

Investigating the cause of language regularization in adults: Memory constraints or learning effects?

Carla L. Hudson Kam and Ann Chang

Department of Psychology, University of California, Berkeley.

Abstract

When language learners are exposed to inconsistent probabilistic grammatical patterns they sometimes impose consistency on the language instead of learning the variation veridically. We hypothesize that this regularization results from problems with word retrieval, not learning per se. We test one prediction of this – that easing the demands of lexical retrieval leads to less regularization. We exposed adult learners to a language containing inconsistent probabilistic patterns and tested them using either a standard production task or one of two tasks that reduced the demands of lexical retrieval. As predicted, participants tested using the modified tasks more closely matched the probability of the inconsistent items than those tested using the standard task.

Keywords

Language acquisition; probability learning; retrieval; language production

Typically children learn a language that looks very much like the input they received. This is not always the case, however; under certain circumstances the learner ends up speaking a language differently than those who provided the input. Although the most well-known cases involve the creation of grammatical patterns in an emerging language (e.g. Senghas, 2000; Senghas & Coppola, 2001; Senghas, Coppola, Newport, & Supalla, 1997), sometimes the changes involve the imposition of consistency and regularity on previously existing inconsistent patterns (Sankoff & Laberge, 1973; Singleton & Newport, 2004), something known as regularization.

One explanation for this is that children, and only children, regularize languages because they innately know that human languages do not contain this kind of unpredictable variation (Becker & Veenstra, 2003; DeGraff, 1999; Lumsden, 1999). This explanation is rooted in a theoretical position that proposes that language is based primarily on domain specific knowledge and mechanisms. While this explanation for regularization is consistent with studies showing that children have a strong tendency to impose regularity on languages in contrast to adults who can learn probabilistic patterns, it cannot account for the finding that adults too will impose consistency under certain circumstances (Hudson Kam & Newport, 2005, in press).

However, there is a growing body of work suggesting that at least some aspects of acquisition can be explained without positing innate knowledge specific to language (see Newport & Aslin, 2000), and we here explore the possibility that this might also be the case for regularization. That is, that regularization might emerge from more general aspects of cognition. Specifically, we propose that regularization results from memory constraints, or rather, aspects of language

production that are sensitive to memory constraints – namely word retrieval. When retrieval is difficult, the most easily accessible form is over-retrieved, resulting in regularization. Our proposal then attempts to explain regularization as resulting from general cognitive constraints rather than anything specific to language or language learning.

Background

Newport and her colleagues have conducted detailed studies examining the acquisition of American Sign Language (ASL) by a deaf child whose only input was from his parents, both late-learners of ASL (Newport, 1999; Singleton & Newport, 2004; Ross & Newport, 1996). As is typical of late learners (Adamson, 1988; Johnson, Shenkman, Newport & Medin, 1996; Newport, 1984, 1990), the parents' signing contained inconsistent and probabilistic, rather than consistent and predictable, grammatical patterns. Simon tended not to reproduce the inconsistency present in his input, however. Instead, his signing was highly consistent – he regularized the language. Simon is not unique in this ability, however: regularization has been shown in other children learning both signed (Ross, 2001) and spoken languages (Kotsinas, 1988).

Regularization has also been shown in experiments. Hudson Kam and Newport (2005), for instance, found that children tend to impose consistency on artificial languages containing inconsistently occurring forms, while adults are able to learn the probabilities associated with the inconsistent forms. In particular, when given input in which a determiner (article) occurred probabilistically with nouns, i.e., occurred only 60% of the time, children produced the determiners in a highly consistent fashion, whereas adult learners matched the probability of occurrence in their own productions, i.e. produced determiners about 60% of the time. However, when adults were exposed to a language in which a main determiner alternated unpredictably with a number of other lower frequency forms called noise determiners they used the main determiners more often than they had heard them – they too began to regularize (Hudson Kam & Newport, in press).

Retrieval and Regularization

We are hypothesizing that the underlying source of regularization is limitations on working memory or executive function. There are a number of different processes or abilities subsumed under these terms (see, e.g., Baddeley, 1986), but the aspect we investigate here is retrieval. When retrieval is difficult, the most easily accessible form is likely to be retrieved repeatedly. In the service of language production, the same form is therefore (inappropriately) retrieved repeatedly, resulting in regularization. Retrieval difficulty of this form can result from cognitive limitations internal to the learner or from the to-be-retrieved material itself. Thus children would be expected to regularize 'easy' variation more than adults, due to their overall lower working memory capacities (Cowan, 1997; Gathercole, 1998). And when the targets of retrieval are multiple competing forms where one is more frequent than the others, even adults might be expected to regularize – the more frequent form is called to mind more easily than the others. On this argument regularization reflects aspects of production processes, rather than reflecting differences in what is learned.¹

A similar proposal has been made to explain individual differences in non-linguistic probability-learning. Kareev, Lieberman, and Lev (1997) found that people with smaller

¹Retrieval can be more difficult when learning is less successful, with weaker representations or weaker links to those representations. In this way individual differences in learning ability might lead to variation in retrieval efficiency and ease. The ability to suppress, another aspect of working memory/executive function, may also be a factor in regularization when the task involves suppressing a stronger form in order to select a weaker form. However, ultimately it is the outcome of the retrieval process we are testing here, whether or not the root cause retrieval is representational strength, retrieval efficiency, or suppression.

working memory capacity were more likely to overestimate probabilities than people with larger capacity. Dougherty and Hunter (2003) suggested that this is because participants with lower capacity are less able to call to mind non-target alternatives at the time of prediction or estimation, especially when the alternatives are low in frequency compared to the target.

Here we test one of the predictions of this hypothesis, namely that making production easier and less reliant on retrieval should result in less regularization and therefore better probability-matching. We exposed adult learners to a language containing variability that they have been previously shown to regularize, but changed the way we administered the production test to reduce the demands of word retrieval. We then compared the productions to those of participants exposed to the same language but who received a more demanding production task, and participants exposed to simpler (learnable) inconsistency. The question was whether and to what degree participants given the easier production task would regularize the inconsistent variation present in the language.

Method

Participants

Forty-eight native-English-speakers, average age 19.9 years ($SD = 2.4$), participated. They were recruited from a subject pool via an email describing the study, or after having responded to recruitment posters. They were paid for their participation.

The Language

The language was designed in conjunction with a set of objects and materials and the relationships that they could readily enter into. The basic vocabulary comprised: 36 nouns referring to objects and substances (e.g. tree, sand) divided into two classes, 7 intransitive verbs referring to actions, relationships, and properties (e.g. move, blue), 5 transitive verbs referring to actions and locative relationships (e.g. hit, be under), 1 negative word, and 2 main determiners (articles), 1 for each noun class. Nouns were assigned to a class on a completely arbitrary basis, with 20 nouns in class one and the remaining 16 in class two.² The only consequence of class is determiner selection. Some participants also heard eight lower frequency noise determiner (ND) forms, described in more detail under Input Manipulation. (This is the same basic language used in earlier studies. A complete word list can be found in Hudson Kam & Newport, 2005).

Word order is (NEG)V-S-(O). (Brackets indicate that the constituent is not obligatory in all sentences.) As is typical for VSO languages, the determiner follows the noun within the NP (Greenberg, 1963). This ordering ensured that participants would have to learn aspects of the language's grammar as well as vocabulary. An example sentence is presented in 1.

(1) flimm mauzner kaw ferluka poe
hit boat_{NC1} DET₁ girl_{NC2} DET₂
'The boat hits the girl.'

The language contains over 13,200 semantically possible sentences. Two hundred thirty (115 intransitive, 115 transitive) were in the exposure set.

²This is atypical for actual languages, but is not unlike the way adult learners treat noun classes (Cain, Weber-Olsen, & Smith. 1987).

Presentation

Participants were exposed to the language by videotape for eight sessions, each lasting 25–29 minutes and containing approximately 115 sentences. Participants were seated in front of a video monitor, on which they watched a scene or event. They then heard a sentence describing the scene. Sentences were spoken at a normal rate with English prosody and phonology and sounded very natural. Participants were asked to repeat each sentence after hearing it. They were told that this pronunciation practice would be helpful as they would have to produce their own sentences at the end of the experiment. (This was not actually the case for participants in one condition, however, for consistency they also received these instructions.) There was no explicit instruction in grammar or vocabulary, and participants never saw anything written during exposure.

The complete exposure set was presented four times over the eight exposure sessions. Testing occurred in an additional session. Participants completed the entire experiment over 9–12 days.

Experimental Manipulation

Two different manipulations were included. First, participants were exposed to one of two different versions of the language, one contained simple presence/absence inconsistency that adults are known to learn, the other, more complex inconsistency which adults regularized in previous experiments. Second, we manipulated the production task in order to vary the difficulty of retrieval: some participants received the standard production task used in previous experiments, others were tested using one of two modified procedures designed to ease the demands of production.

Input manipulation—For all participants, the main determiners occurred with 60% of the in-class noun phrases. That is, 60% of the time participants heard a noun from class one it occurred with the appropriate main determiner, likewise for nouns in class two. The manipulation occurred in the other 40% of the noun phrases. In the 0-ND inconsistency language (zero noise determiners, or presence/absence inconsistency) the remaining noun phrases occurred without determiners. In the 8-ND inconsistency language, these nouns were instead accompanied by one of eight lower frequency NDs, each of which occurred with 5% of the noun phrases in each class.³

These percentages were true of noun phrases in general, as well of the various syntactic positions, exposure sessions, and sentence types. For example, 60% of intransitive subjects, transitive subjects, and transitive objects occurred with main determiners. The NDs were similarly evenly distributed. However, occurrence percentages were not precisely the same for each noun; individual nouns occurred with main determiners 41%–74% of the time, for example. Importantly, each presentation of a particular exposure set sentence had the potential to be different from the other three. Thus, the only patterns of determiner use were the noun classes and the different overall frequencies of the main and noise forms. All other aspects of the language were completely consistent.

Testing manipulation—Manipulations also occurred in the sentence production task. In this task, participants saw a novel scene and had to produce the corresponding (novel) sentence. There were 24 test sentences, 12 transitive, 12 intransitive, resulting in 36 possible determiners. Responses were video-recorded and later examined for determiner use.

In the *Standard* test, participants were told that they would see a scene on the monitor and hear the first word (i.e., the verb) of the corresponding sentence. Their task was to produce a sentence

³Like the main determiners, the NDs were monosyllabic. Four had a CV structure and four had a CVC structure.

describing the scene, beginning with the word they had been given. Participants were asked to indicate where a word they could not recall should go in the sentence (e.g., by saying X), allowing us to include data from incomplete responses. In this version of the test, participants had to recall the nouns and determiners themselves, assemble the constituents of the sentence into the appropriate order, and then finally produce the sentence (see Bock, 1995).

In the *Flashcard* version, the researcher began by laying out 3×5 index cards on which all the verbs, nouns, and determiners in the language were written. Participants were then told that they would see a scene on the video monitor and hear the first word in the sentence and that their task was to select the words necessary to describe the scene, and place them in order on the table front of them. Participants were asked to indicate where a word they could not recall should go in the sentence, and were told that if they felt the same word should go in more than one place, to just tell the researcher. This version of the test is more of a recognition test, as all possible words were provided for participants. Participants were not provided with a sentence template to use with the cards, thus they still had to actively recall at least some aspects of the novel language such as word order as well as how many words were required. However, overall, the flashcard version is much less cognitively demanding, especially with respect to word retrieval.

The flashcards reduce the demands of active recall, however, they also reduce the amount of maintenance required. We controlled for this effect in the *Nouns-provided* test. Here, after seeing the scene to be described, participants heard the verb and noun(s) appropriate to the sentence. They were told to produce a sentence (orally) describing the scene using the words they had heard, adding whatever other words they felt necessary, if any. Two lists of randomized orders of the verb and noun(s) were created, with half of the participants hearing each. In providing the content words to be used in the sentence, we reduced the amount of effort the speaker has to exert to access and retrieve words. However, the determiners were not directly cued for participants; they still had to be actively retrieved. Thus, they had to keep the words they were given in mind (maintenance), recall the determiners (retrieval), and actively call to mind details of the language's structure to use in constituent assembly (retrieval and manipulation). However, unlike the Standard test, they did not have to recall the phonological forms of the nouns. According to theories of speech production, function words are selected after content words (see, e.g., Bock, 1995). The resource savings from giving participants the content words should, therefore, enable them to devote more resources to retrieving the determiners.

Resulting conditions—Two groups of participants received the Standard test, one was exposed to 0-ND inconsistency (0-ND Standard), the other to the lower frequency NDs (8-ND Standard). Likewise two groups of participants performed a modified production test, both heard the noise determiners, one received the Flashcard test (8-ND Flashcard), the other the Nouns-given test (8-ND Nouns-given).

If regularization results solely from the input, all three conditions exposed to the 8-ND language should show evidence of regularization, regardless of method of testing. If instead regularization results from aspects of the production process, those given the less demanding production tests (8-ND Flashcard, 8-ND Nouns-given) should show behaviour closer to probability-matching, much like participants in the 0-ND Standard condition.

Other Tests

Determiner judgement—This test assessed participants' knowledge of determiners through judgements rather than production. Participants listened to sentences presented via a computer and judged them on a four-point scale. Presentation was audio only – there was no corresponding visual scene. Participants were instructed to give the sentence a high rating when

it sounded just like a sentence from the language that they had been learning and a low rating when it sounded completely unlike a sentence from the language. If they thought a sentence was mostly, but not completely, like or unlike sentences from the language, they were to use the middle of the scale. Participants had 3 seconds in which to respond, and responses were recorded by the experimenter.

There were 36 test items, consisting of three variations of 12 base sentences: one contained the appropriate main determiner, one contained a ND, and one had no determiner. The items were randomly ordered with the constraint that two versions of the same base sentence could not follow each other. All test sentences were novel.

Sentence structure—This task assessed participants' knowledge of other aspects of the language, in particular, basic sentence construction and verb-subcategorization. Participants listened to 16 pairs of sentences and were asked to select the member of the pair that sounded most like a sentence from the language. The two sentences in each pair were versions of the same base sentence, one grammatical, one ungrammatical. Half the pairs tested whether participants knew that transitive verbs required two nouns and intransitives only one, the others tested whether participants knew that a verb was required in every sentence. Participants responded by circling 1 or 2 on an answer sheet. There was a 1-second pause between the two sentences in a pair, and a 5-second pause between pairs. The location of the grammatical sentence was randomized, as was the ordering of items, with the constraint that no more than two pairs could occur in a row that tested the same rule and were of the same transitivity. Again, the test sentences were novel; they had not occurred in the exposure set or the other tests.

Results and Discussion

Sentence Production

Figure 1 shows the percentage of correct nouns accompanied by the appropriate main determiner, the main measure of interest. Given our hypothesis we would expect participants exposed to the complex inconsistency but tested using the modified production tasks to match their input much more closely than those exposed to the complex inconsistency but given the standard production task, possibly to the point of being indistinguishable from those exposed to the simpler inconsistency.

An ANOVA revealed a significant effect of condition ($F_{(3,44)} = 6.4, p = .001$). As in previous work, the 8-ND Standard and 0-ND Standard conditions were significantly different from each other ($t'_{(44)} = 3.65, p = .004$). With respect to our specific hypotheses, further comparisons showed that as predicted, participants in the 8-ND Standard condition produced significantly more main determiners than those in the 8-ND Flashcard condition ($t'_{(44)} = 3.5, p = .006$) and the 8-ND Nouns-provided condition ($t'_{(44)} = 3.57, p = .005$). In contrast, the difference between participants in 0-ND Standard condition and those in the 8-ND Flashcard condition is not significant ($t'_{(44)} = 0.538, p = 1.000$), similarly for the 8-ND Nouns-provided condition ($t'_{(44)} = 0.083, p = 1.000$). Importantly, the 8-ND Nouns-provided and 8-ND Flashcard conditions are not significantly different from each other ($t'_{(44)} = 0.074, p = 1.000$).

We next examined how closely participants matched the main determiner input percentage (60%), and found that only participants in the 8-ND Standard condition produced main determiners significantly more (or less) often than they heard them (0-ND Standard: $t_{(11)} = -1.45, p = .175$; 8-ND Standard: $t_{(11)} = 4.65, p = .001$; 8-ND Flashcard: $t_{(11)} = -2.03, p = .067$; 8-ND Nouns-provided: $t_{(11)} = -1.36, p = .2$). Thus, as predicted by our retrieval hypothesis, participants exposed to complex inconsistency were able to more closely match their input when production was made easier.

We also examined what participants produced the rest of the time. Figure 2 shows the percentage of productions that were things other than correct main determiners: including no determiner (zero), incorrect main forms, NDs, and ‘other’. Other includes anything other than main or noise determiners. Most were blends such as ‘koe’ or ‘paw’, but some participants occasionally used ‘mang’, the word for block, in place of a determiner. Given the low individual input frequency of the NDs (5%), we show the percentage of NDs produced in syllable-structure categories, CV and CVC, each of which comprised four different NDs and represented 20% of the input.

Several patterns are apparent in these data. The first concerns nouns produced without determiners; these were common for participants in the 0-ND Standard condition but quite rare for all other participants. This is reasonable, as only that group was exposed to any nouns without determiners. The low frequency of incorrect main determiners also stands out. Although participants did occasionally produce ‘kaw’ where they should have said ‘poe’ or vice versa, this was very infrequent, suggesting that participants were able to learn the noun classes (or at least which determiners belonged with which nouns).⁴

Looking just at the conditions exposed to NDs, although the testing manipulation enabled better input-matching for the main determiners, participants did not necessarily match the probabilities of the NDs. In all three conditions, although production of CV forms did not differ significantly from the input (8-ND Standard: $t_{(11)} = -1.8$, $p = .099$; 8-ND Flashcard: $t_{(11)} = 1.68$, $p = .12$; 8-ND Nouns-provided: $t_{(11)} = 3.27$, $p = .078$), production of CVC forms was significantly below the input proportion (8-ND Nouns-provided: $t_{(11)} = -12.17$, $p < .001$; 8-ND Flashcard: $t_{(11)} = 3.05$, $p = .011$; 8-ND Standard: $t_{(11)} = -4.76$, $p = .001$). This could be due to the relative phonological heaviness of the CVC forms: in natural languages function words tend to be phonologically light, thus, the CVC forms might be dispreferred as determiners.

There were also some interesting differences between the three 8-ND-exposed conditions in their production of incorrect main and ‘other’ forms. In particular, the 8-ND Nouns-provided participants produced more of both than participants in the other two 8-ND conditions (incorrect main – 8-ND Standard: $t_{(22)} = 2.33$, $p = .03$; 8-ND Flashcard: $t_{(22)} = 2.16$, $p = .042$; ‘other’ – 8-ND Standard: $t_{(11,24)} = 3.38^5$, $p = .006$; 8-ND Flashcard: $t_{(11,20)} = 3.39^5$, $p = .006$). These results are quite sensible in light of the testing tasks. Participants in the 8-ND Standard condition only produced nouns they could recall. Those in the 8-ND Flashcard condition could use recognition to ‘produce’ nouns that they knew less well, but they could also use recognition for the determiners. In contrast, those in the 8-ND Nouns-provided condition were given the nouns, and so had to produce all nouns in every sentence whether they knew them well or not and without the extra support for determiners given in the 8-ND Flashcard condition.⁶

Determiner judgement

Figure 3 shows the ratings participants gave to sentences of various types. Condition is not significant ($F_{(3,44)} = 1.173$, $p = .331$), sentence type is ($F_{(2,88)} = 125.172$, $p < .001$), however,

⁴This contrasts with findings showing that learners have difficulty acquiring classes without multiple correlated cues (e.g. Braine et al., 1990; Brooks et al., 1993). It may be the case that our relatively more naturalistic exposure (without teaching) encouraged participants to learn in a different way. Moreover, our participants saw each noun in more varied contexts than other studies, which might encourage abstraction. Alternatively, it is possible that our participants did not actually learn classes. We had no other reflex of class and so no paradigm to test. (The test sentences were novel, however, so the particular combinations of words were new.) Notably, the determiners were usually produced as separate (prosodic) words, suggesting that, at the very least, participants did not simply view them as part of the noun itself.

⁵Adjusted df used due to significant heterogeneity of variance.

⁶There were also some differences in their production of NDs. Participants in the 8-ND Flashcard condition produced significantly more CV NDs than the other two 8-ND groups (8-ND Standard: $t_{(22)} = -2.42$, $p = .024$; 8-ND Nouns-provided: $t_{(16,2)}^5 = 2.73$, $p = .015$), and more CVC forms than participants in the 8-ND Nouns-provided condition ($t_{(22)} = -2.42$, $p = .024$).

so is the interaction ($F_{(6,88)} = 49.194, p < .001$), reflecting the fact that the ratings given to the various sentence types are not the same in the four conditions.

Planned comparisons revealed that all conditions rated the sentence types they had heard 60% of the time significantly higher than those they heard 40% of the time (8-ND Flashcard: $F_{(1,11)} = 13.99, p = .003$; 8-ND Nouns-provided: $F_{(1,11)} = 11.34, p = .006$; 8-ND Standard: $F_{(1,11)} = 12.2, p = .005$; 0-ND Standard: $F_{(1,11)} = 12.19, p = .005$), which were in turn rated significantly higher than sentence types they had not heard (8-ND Flashcard: $F_{(1,11)} = 25.27, p < .001$; 8-ND Nouns-provided: $F_{(1,11)} = 18.747, p = .001$; 8-ND Standard: $F_{(1,11)} = 89.23, p < .001$; 0-ND Standard: $F_{(1,11)} = 304.15, p < .001$). However, which type of sentence was the type heard less frequently and which was the type not heard differed across conditions (hence the interaction). For the three 8-ND conditions sentences with NDs were the less-frequent type and those with no determiners were the type not encountered. This was reversed for the 0-ND condition. Ratings in all conditions, then, reflect the input frequencies.

Sentence Structure

Importantly, the production differences cannot be explained by differences in language knowledge more generally. Performance on the sentence structure test was high across conditions (0-ND Standard = 14.58, 8-ND Standard = 14.75, 8-ND Flashcards = 14.25, 8-ND Nouns-provided = 15.5; max = 16), and condition was not significant ($F_{(3,44)} = 1.01, p = .398$). The overall mean of 14.77 (SD = 1.82) was significantly above chance ($t_{(47)} = 25.71, p < .001$).

General Discussion

The present experiment was designed to test the hypothesis that regularization of inconsistencies in language results from general cognitive processes, namely difficulties with self-directed word retrieval. Our results generally support this proposal. We found that participants given the modified testing did not regularize the main determiners like those given the standard production task. Instead, they closely matched the input probabilities of the main determiners, much like participants exposed to simpler inconsistency. Importantly, the judgements given to sentences of various types reflected the input frequencies for participants in all conditions, again suggesting sensitivity to the probabilities not always apparent in production.

Our hypothesis was based in part on work in the probability-estimation literature, in particular Dougherty and Hunter's (2003) proposal. Although theirs is as an explanation of individual variation, it also predicts differences dependent on the stimuli, specifically, that when there are few alternatives people should be better able to call to mind all of the possibilities, when there are many low frequency alternatives each is more difficult to recall and so the more frequent alternative is recalled instead. This is exactly what we find: people exposed to 0-ND match the probabilities more closely than those exposed to 8-NDs when retrieval or recall is required. When retrieval pressures are lessened, however, they too can demonstrate knowledge of the underlying probabilities, at least of the more frequent forms, showing a distinction between knowledge and behaviour (see Gaissmaier, Schooler, & Rieskamp, 2006). Our general hypothesis also predicts that we should be able to induce greater regularization by increasing the cognitive demands. Another non-linguistic probability-learning study is suggestive on this point. Wolford, Newman, Miller, and Wig (2004) found maximizing in a 2-choice task when subjects performed a simultaneous n-back task. Related, Bybee and Slobin (1982) found that adults show over-regularization of morphological forms even for words they already know when speaking under less-than-ideal conditions, such as rushed production (see also McDonald, 2006). Whether this is true for regularization as well remains to be seen.

Although the results support our hypothesis, we must acknowledge that children might regularize for different reasons than adults. Children have different memory abilities than adults *all* the time, not just at retrieval. Thus, it is quite possible that their initial encoding of the material may differ quite substantially, possibly to the point where lower frequency forms are difficult enough to retrieve one would want to say that children's knowledge is qualitatively, rather than quantitatively, different. Indeed, Newport's (1990) Less-is-More hypothesis posits that children's cognitive limitations actually make them better at finding grammatical patterns in languages, and there is some experimental evidence for this (Elman, 1993; Kersten & Earles, 2001; Pitts Cochran, McDonald, & Parault, 1999). Interestingly, however, retrieval problems have also been suggested as an explanation for children's over-regularizations in typical acquisition; children produce forms such as *sitted* because they fail to retrieve the correct form *sat* (Marcus et al., 1992). It is worth noting that, in contrast to the regularization we are examining which appears to be long-lasting (Ross & Newport, 1996; Singleton & Newport, 2004), over-regularizations do not persist. This might be due to their depending on markedly different processes, but we think it more likely that over-regularizations disappear because the relevant ambient language is completely consistent and regular (the past tense of *sit* is always *sat*), something not true of the inconsistencies we are studying. When there is a true generalization it eventually will be the strongest representation, when there is not, a regularization can persist.

In sum, we found that regularization in adult learners is reduced when word-retrieval demands are reduced, consistent with the idea that regularization results from constraints imposed by domain-general cognitive processes, rather than domain-specific knowledge.

Acknowledgments

This research was supported by NIH grant HD 048572 to C. Hudson Kam.

We wish to thank Amy Finn and Whitney Goodrich for their assistance, and Chris Kam for comments on the manuscript.

References

- Adamson, HD. Variation theory and second language acquisition. Washington, DC: Georgetown University Press; 1988.
- Baddeley, AD. Working memory. Oxford: Oxford University Press; 1986.
- Becker A, Veenstra T. The survival of inflectional morphology in French-related creoles. *Studies in Second Language Acquisition* 2003;25:283–306.
- Bock, K. Sentence production: From mind to mouth. In: Miller, JL.; Eimas, PD., editors. *Speech, language, and communication*. New York: Academic Press; 1995. p. 181-216.
- Braine MDS, Brody RE, Brooks PJ, Sudhalter V, Ross JA, Catalano L, Fisch SM. Exploring language acquisition in children with a miniature artificial language: Effects of item and pattern frequency, arbitrary subclasses, and correction. *Journal of Memory and Language* 1990;29:591–610.
- Brooks PJ, Braine MDS, Catalano L, Brody RE, Sudhalter V. Acquisition of gender-like noun subclasses in an artificial language: The contribution of phonological markers to learning. *Journal of Memory and Language* 1993;32:76–95.
- Bybee, J.; Slobin, D. Why small children cannot change language on their own: Evidence from the English past tense. In: Alqvist, A., editor. *Papers from the 5th International Conference on Historical Linguistics*; Amsterdam: John Benjamins; 1982. p. 29-37.
- Cain J, Weber-Olsen M, Smith R. Acquisition strategies in a first and second language: Are they the same? *Journal of Child Language* 1987;14:333–352. [PubMed: 3611246]
- Cowan, N., editor. *The development of memory in childhood*. Hove, East Sussex: Psychology Press; 1997.

- DeGraff, M. Creolization, language change, and language acquisition: An epilogue. In: DeGraff, M., editor. *Language creation and language change: Creolization, diachrony, and development*. Cambridge, MA: MIT Press; 1999. p. 473-543.
- Dougherty MRP, Hunter J. Probability judgement and subadditivity: The role of working memory capacity and constraining retrieval. *Memory and Cognition* 2003;31:968-982.
- Elman JL. Learning and development in neural networks: The importance of starting small. *Cognition* 1993;48:71-99. [PubMed: 8403835]
- Gaissmaier W, Schooler LJ, Rieskamp J. Simple predictions fueled by capacity limitations: When are they successful? *Journal of Experimental Psychology: Learning, Memory, and Cognition* 2006;32:966-982.
- Gathercole SE. The development of memory. *Journal of Child Psychology and Psychiatry* 1998;39:3-27. [PubMed: 9534084]
- Greenberg, JH. Some universals of grammar with particular reference to the order of meaningful elements. In: Greenberg, JH., editor. *Universals of language*. Cambridge, MA: MIT Press; 1963. p. 73-113.
- Hudson Kam CL, Newport EL. Regularizing unpredictable variation: The roles of adult and child learners in language variation and change. *Language Learning and Development* 2005;1:151-195.
- Hudson Kam CL, Newport EL. Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology*. (in press)
- Johnson J, Shenkman K, Newport EL, Medin D. Indeterminacy in the grammar of adult language learners. *Journal of Memory and Language* 1996;35:335-352.
- Kareev Y, Lieberman I, Lev M. Through a narrow window: Sample size and the perception of correlation. *Journal of Experimental Psychology: General* 1997;126:278-287.
- Kersten AW, Earles JL. Less really is more for adults learning a miniature artificial language. *Journal of Memory and Language* 2001;44:250-273.
- Kotsinas U-B. Immigrant Children's Swedish: A New Variety? *Journal of Multilingual and Multicultural Development* 1988;9:129-140.
- Lumsden, JS. Language acquisition and creolization. In: Degraff, M., editor. *Creolization, diachrony, and language acquisition*. Cambridge, MA: MIT Press; 1999. p. 129-157.
- Marcus G, Pinker S, Ullman M, Hollander M, Rosen T, Xu F. Overregularization in language acquisition. *Monographs of the Society for Research in Child Development*. 1992Serial no. 228
- McDonald JL. Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language* 2006;55:381-401.
- Newport EL. Constraints on learning: Studies in the acquisition of American Sign Language. *Papers and Reports on Child Language Development* 1984;23:1-22.
- Newport EL. Maturation constraints on language learning. *Cognitive Science* 1990;14:11-28.
- Newport, EL. Reduced input in the acquisition of signed languages: Contributions to the study of creolization. In: Degraff, Michel, editor. *Creolization, diachrony, and language acquisition*. Cambridge, MA: MIT Press; 1999. p. 161-178.
- Newport, EL.; Aslin, RN. Innately constrained learning: Blending old and new approaches to language acquisition. In: Howell, SC.; Fish, SA.; Keith-Lucas, T., editors. *Proceedings of the 24th Annual Boston University Conference on Language Development*; Somerville, MA: Cascadia Press; 2000. p. 1-21.
- Pitts Cochran B, McDonald JL, Parault SJ. Too smart for their own good: The disadvantage of a superior processing capacity for adult language learners. *Journal of Memory and Language* 1999;41:30-58.
- Ross DS. Disentangling the nature-nurture interaction in the language acquisition process: Evidence from deaf children of hearing parents exposed to non-native input (Doctoral Dissertation, University of Rochester, 2001). *Dissertation Abstracts International* 62-07B 2001:3402.
- Ross, DS.; Newport, EL. The development of language from non-native linguistic input. In: Stringfellow, A.; Cahana-Amitay, D.; Hughs, E.; Zukowski, A., editors. *Proceedings of the 20th annual Boston University Conference on Language Development*; Boston: Cascadia Press; 1996. p. 623-645.
- Sankoff G, Laberge S. On the acquisition of native speakers by a language. *Kivung* 1973;6:32-47.

- Senghas, A. The development of early spatial morphology in Nicaraguan Sign Language. In: Howell, SC.; Fish, SA.; Keith-Lucas, T., editors. Proceedings of the 24th Annual Boston University Conference on Language Development; Somerville, MA: Cascadilla; 2000. p. 696-707.
- Senghas A, Coppola M. The creation of Nicaraguan Sign Language by children: Language genesis as language acquisition. *Psychological Science* 2001;12:323–328. [PubMed: 11476100]
- Senghas, A.; Coppola, M.; Newport, EL.; Supalla, T. Argument structure in Nicaraguan Sign Language: The emergence of grammatical devices. In: Hughes, E.; Hughs, M.; Greenhill, A., editors. Proceedings of the 21st Annual Boston University Conference on Language Development; Boston: Cascadilla; 1997. p. 550-561.
- Singleton JL, Newport EL. When learners surpass their models: The acquisition of American Sign Language from impoverished input. *Cognitive Psychology* 2004;49:370–407. [PubMed: 15342259]
- Wolford G, Newman SE, Miller MB, Wig GS. Searching for patterns in random sequences. *Canadian Journal of Experimental Psychology* 2004;58:221–228. [PubMed: 15648726]

Figure 1.
Mean production of correct main determiner forms by condition.

Figure 2.
Mean production of other forms by condition.

Figure 3.
Mean ratings by condition.