BRIEF REPORTS

Investigating the Cause of Language Regularization in Adults: Memory Constraints or Learning Effects?

Carla L. Hudson Kam and Ann Chang
University of California, Berkeley

When language learners are exposed to inconsistent probabilistic grammatical patterns, they sometimes impose consistency on the language instead of learning the variation veridically. The authors hypothesized that this regularization results from problems with word retrieval rather than from learning per se. One prediction of this, that easing the demands of lexical retrieval leads to less regularization, was tested. Adult learners were exposed to a language containing inconsistent probabilistic patterns and were tested with either a standard production task or one of two tasks that reduced the demands of lexical retrieval. As predicted, participants tested with the modified tasks more closely matched the probability of the inconsistent items than did those tested with 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. Whereas 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 here we 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

Regularization in Language Acquisition

Newport and her colleagues have conducted detailed studies examining the acquisition of American Sign Language (ASL) by a deaf child the authors referred to by the pseudonym "Simon," whose only input was from his parents, both late learners of ASL (Newport, 1999; Ross & Newport, 1996; Singleton & Newport, 2004). 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 inconsis-

Correspondence concerning this article should be addressed to Carla L. Hudson Kam, Department of Psychology, 3210 Tolman Hall, #1650, University of California, Berkeley, Berkeley, CA 94720. E-mail: clhudson@berkeley.edu

tently occurring forms, whereas 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, that is, 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, that is, 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, that is, 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, then 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 differences in what is learned.[1]

A similar proposal has been made to explain individual differences in nonlinguistic probability learning. Kareev, Lieberman, and Lev (1997) found that people with smaller 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 nontarget alternatives at the time of prediction or estimation, especially when the alternatives are low in frequency compared with the target.

In the current study, we tested 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 with those of (a) participants exposed to the same language but who received a more demanding production task and (b) 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 in the experiment. They were recruited from a participant pool via an e-mail 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 (NC). Nouns were assigned to a class on a completely arbitrary basis, with 20 nouns in Class 1 and the remaining 16 in Class 2.[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), where NEG represents negative, V represents verb, S represents subject, and O represents object. (Brackets indicate that the constituent is not obligatory in all sentences.) As is typical for VSO languages, the determiner (DET) follows the noun within the noun phrase (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 here:

flimm	mauzner	kaw	ferluka	poe
hit	boat$_{NC1}$	DET$_1$	girl$_{NC2}$	DET$_2$

"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.

Presentation

Participants were exposed to the language by videotape for eight sessions, each lasting 25–29 min 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

___

[1] Retrieval 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 is representational strength, retrieval efficiency, or suppression.

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

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 these participants 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 containing simple presence/absence inconsistency that adults are known to learn and the other containing more complex inconsistency, which adults regularized in previous experiments. Second, we manipulated the production task to vary the difficulty of retrieval: Some participants received the standard production task used in previous experiments, and others 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 that participants heard a noun from Class 1, it occurred with the appropriate main determiner, likewise for nouns in Class 2. 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 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 in. (approximately 7.5 × 17.5 cm) index cards on which all of 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 in 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 the number of words 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. In this test, 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 were 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 had to exert to access and retrieve words. However, the determiners were not directly cued for participants; they still had to be actively retrieved. Thus, the participants 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 case of 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-provided test (8-ND nouns-provided).

If regularization results solely from the input, then participants in all three conditions exposed to the 8-ND language should show

---

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

evidence of regularization, regardless of method of testing. If, instead, regularization results from aspects of the production process, participants given the less demanding production tests (8-ND flashcard, 8-ND nouns-provided) should show behavior closer to probability matching, much like participants in the 0-ND standard condition.

Other Tests

Determiner judgment. This test assessed participants' knowledge of determiners through judgments rather than production. Participants listened to sentences presented via a computer and judged them on a 4-point scale. Presentation was by audio recording 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 s 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 an 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. The sentence structure 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 intransitive verbs 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-s pause between the two sentences in a pair and a 5-s 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 expected participants exposed to the complex inconsistency but tested with 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.
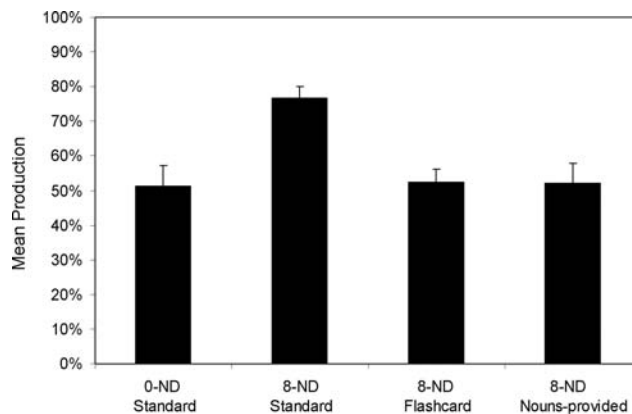

Figure 1. Mean production of correct main determiner forms by condition.

An analysis of variance (ANOVA) revealed a significant effect of condition, $F(3, 44) = 6.40$, $p = .001$. 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 did those in the 8-ND flashcard condition, $t'(44) = 3.50$, $p = .006$, and those in 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 was not significant, $t'(44) = 0.54$, $p = 1.000$; similarly, for the 8-ND nouns-provided condition, $t'(44) = 0.08$, $p = 1.000$. Importantly, the 8-ND nouns-provided and 8-ND flashcard conditions were not significantly different from each other, $t'(44) = 0.07$, $p = 1.00$.

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, $t(11) = -1.36$, $p = .200$. 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 consisted of things other than correct main determiners, including no determiner (zero), incorrect main forms, NDs, and "other," where "other" included 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, consonant–vowel (CV), and consonant–vowel–consonant (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
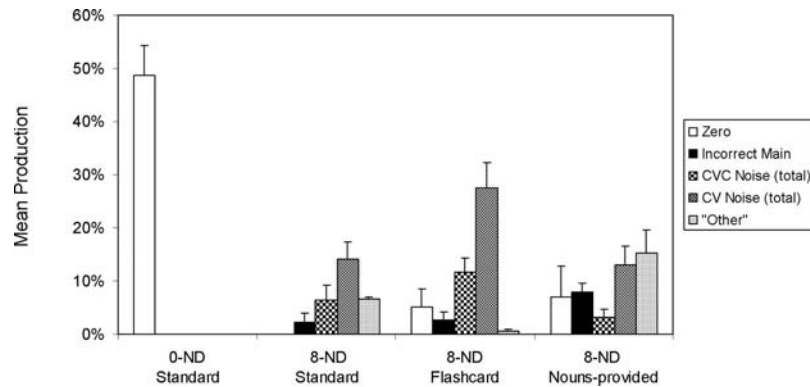
Figure 2. Mean production of other forms by condition. CVC = consonant–vowel–consonant; CV = consonant–vowel.

participants in the 0-ND standard condition but quite rare for all other participants. This is reasonable, as only the 0-ND standard 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, we found that 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, $t(11) = -1.80$, $p = .099$; 8-ND flashcard, $t(11) = 1.68$, $p = .120$; 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 have been dispreferred as determiners.

There were also some interesting differences between the three 8-ND–exposed conditions in participants' production of incorrect main and "other" forms. In particular, the 8-ND nouns-provided participants produced more of both types of responses than did participants in the other two 8-ND conditions: incorrect main, 8-ND standard, $t(22) = 2.33$, $p = .030$; 8-ND flashcard, $t(22) = 2.16$, $p = .042$; "other," 8-ND standard, $t(11.24) = 3.38$, $p = .006$; 8-ND flashcard, $t(11.20) = 3.39$, $p = .006$.[5] 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 therefore 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 Judgment

Figure 3 shows the ratings participants gave to sentences of various types. Condition was not significant, $F(3, 44) = 1.17$, $p = .331$, sentence type was significant, $F(2, 88) = 125.17$, $p < .001$, however, as was the interaction between them, $F(6, 88) = 49.194$, $p < .001$, reflecting the fact that the ratings given to the various sentence types are not the same in the four conditions. Planned comparisons revealed that participants in all of the conditions rated the sentence types they had heard 60% of the time significantly higher than those they had 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, $F(1, 11) = 12.20$, $p = .005$; 0-ND standard, $F(1, 11) = 12.19$, $p = .005$, which, in turn, were rated significantly higher than sentence types they had not heard: 8-ND flashcard, $F(1, 11) = 25.27$, $p < .001$; 8-ND nouns-provided, $F(1, 11) = 18.75$, $p = .001$; 8-ND standard, $F(1, 11) = 89.23$, $p < .001$; 0-ND standard, $F(1, 11) = 304.15$, $p < .001$. However, the type of sentence heard less frequently and the type not heard

---

[4] This contrasts with findings showing that learners have difficulty acquiring classes without multiple correlated cues (e.g., Braine et al., 1990; Brooks, Braine, Catalano, Brody, & Sudhalter, 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.

[5] Adjusted degrees of freedom were used due to significant heterogeneity of variance.

[6] There were also some differences in their production of NDs. Participants in the 8-ND flashcard condition produced significantly more CV NDs than participants in the other two 8-ND groups: 8-ND standard, $t(22) = -2.42$, $p = .024$; 8-ND nouns-provided, $t(16.2) = 2.73$, $p = .015$ (see Footnote 5 regarding degrees of freedom), and more CVC forms than participants in the 8-ND nouns-provided condition, $t(22) = -2.42$, $p = .024$.
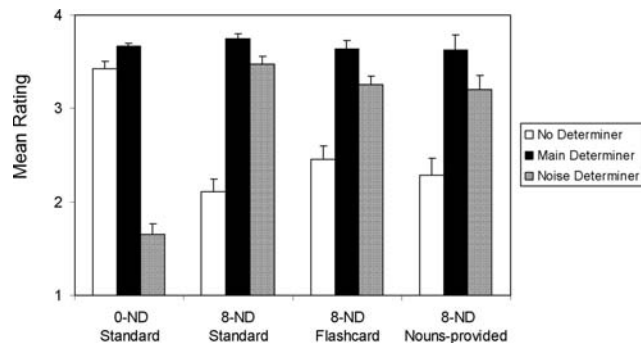
Figure 3. Mean ratings by condition.

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, reflected 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; maximum = 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 the participants 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 judgments 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 behavior (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 nonlinguistic probability learning study is suggestive on this point. Wolford, Newman, Miller, and Wig (2004) found maximizing in a 2-choice task when participants performed a simultaneous *n*-back task. Relatedly, 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.

References

Adamson, H. D. (1988). *Variation theory and second language acquisition.* Washington, DC: Georgetown University Press.

Baddeley, A. D. (1986). *Working memory.* Oxford: Oxford University Press.

Becker, A., & Veenstra, T. (2003). The survival of inflectional morphology in French-related creoles. *Studies in Second Language Acquisition, 25,* 283–306.

Bock, K. (1995). Sentence production: From mind to mouth. In J. L. Miller & P. D. Eimas (Eds.), *Speech, language, and communication* (pp. 181–216). New York: Academic Press.

Braine, M. D. S., Brody, R. E., Brooks, P. J., Sudhalter, V., Ross, J. A., Catalano, L., & Fisch, S. M. (1990). 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, 29,* 591–610.

Brooks, P. J., Braine, M. D. S., Catalano, L., Brody, R. E., & Sudhalter, V. (1993). Acquisition of gender-like noun subclasses in an artificial language: The contribution of phonological markers to learning. *Journal of Memory and Language, 32,* 76–95.

Bybee, J., & Slobin, D. (1982). Why small children cannot change language on their own: Evidence from the English past tense. In A. Alqvist (Ed.), *Papers from the 5th International Conference on Historical Linguistics* (pp. 29–37). Amsterdam, the Netherlands: John Benjamins.

Cain, J., Weber-Olsen, M., & Smith, R. (1987). Acquisition strategies in a first and second language: Are they the same? *Journal of Child Language, 14,* 333–352.

Cowan, N. (Ed.). (1997). *The development of memory in childhood.* Hove, East Sussex: Psychology Press.

DeGraff, M. (1999). Creolization, language change, and language acquisition: An epilogue. In M. DeGraff (Ed.), *Language creation and language change: Creolization, diachrony, and development* (pp. 473–543). Cambridge, MA: MIT Press.

Dougherty, M. R. P., & Hunter, J. (2003). Probability judgment and subadditivity: The role of working memory capacity and constraining retrieval. *Memory and Cognition, 31,* 968–982.

Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition, 48,* 71–99.

Gaissmaier, W., Schooler, L. J., & Rieskamp, J. (2006). Simple predictions fueled by capacity limitations: When are they successful? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 966–982.

Gathercole, S. E. (1998). The development of memory. *Journal of Child Psychology and Psychiatry, 39,* 3–27.

Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of language* (pp. 73–113). Cambridge, MA: MIT Press.

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

Hudson Kam, C. L., & Newport, E. L. (in press). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology.* doi: 10.1016/j.cogpsych.2009.01.001

Johnson, J., Shenkman, K., Newport, E. L., & Medin, D. (1996). Indeterminacy in the grammar of adult language learners. *Journal of Memory and Language, 35,* 335–352.

Kareev, Y., Lieberman, I., & Lev, M. (1997). Through a narrow window: Sample size and the perception of correlation. *Journal of Experimental Psychology: General, 126,* 278–287.

Kersten, A. W., & Earles, J. L. (2001). Less really is more for adults learning a miniature artificial language. *Journal of Memory and Language, 44,* 250–273.

Kotsinas, U.-B. (1988). Immigrant children's Swedish: A new variety? *Journal of Multilingual and Multicultural Development, 9,* 129–140.

Lumsden, J. S. (1999). Language acquisition and creolization. In M. Degraff (Ed.), *Creolization, diachrony, and language acquisition* (pp. 129–157). Cambridge, MA: MIT Press.

Marcus, G., Pinker, S., Ullman, M., Hollander, M., Rosen, T., & Xu, F. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development, 57*(4, Serial No. 228).

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language, 55,* 381–401.

Newport, E. L. (1984). Constraints on learning: Studies in the acquisition of American Sign Language. *Papers and Reports on Child Language Development, 23,* 1–22.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science, 14,* 11–28.

Newport, E. L. (1999). Reduced input in the acquisition of signed languages: Contributions to the study of creolization. In Michel Degraff (Ed.), *Creolization, diachrony, and language acquisition* (pp. 161–78). Cambridge, MA: MIT Press.

Newport, E. L., & Aslin, R. N. (2000). Innately constrained learning: Blending old and new approaches to language acquisition. In S. C. Howell, S. A. Fish, & T. Keith-Lucas (Eds.), *Proceedings of the 24th annual Boston University Conference on Language Development* (pp. 1–21). Somerville, MA: Cascadilla Press.

Pitts Cochran, B., 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,* 30–58.

Ross, D. S. (2001). 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, 62–07*B, 3402.

Ross, D. S., & Newport. E. L. (1996). The development of language from non-native linguistic input. In A. Stringfellow, D. Cahana-Amitay, E. Hughs, & A. Zukowski (Eds.), *Proceedings of the 20th annual Boston University Conference on Language Development* (pp. 623–645). Boston: Cascadilla Press.

Sankoff, G., & Laberge, S. (1973). On the acquisition of native speakers by a language. *Kivung, 6,* 32–47.

Senghas, A. (2000). The development of early spatial morphology in Nicaraguan Sign Language. In S. C. Howell, S. A. Fish, & T. Keith-Lucas (Eds.), *Proceedings of the 24th annual Boston University Conference on Language Development* (pp. 696–707). Somerville, MA: Cascadilla.

Senghas, A., & Coppola, M. (2001). The creation of Nicaraguan Sign Language by children: Language genesis as language acquisition. *Psychological Science, 12,* 323–328.

Senghas, A., Coppola, M., Newport, E. L., & Supalla, T. (1997). Argument structure in Nicaraguan Sign Language: The emergence of grammatical devices. In E. Hughes, M. Hughs, & A. Greenhill (Eds.), *Proceedings of the 21st annual Boston University Conference on Language Development* (pp. 550–561). Somerville, MA: Cascadilla.

Singleton, J. L., & Newport, E. L. (2004). When learners surpass their models: The acquisition of American Sign Language from impoverished input. *Cognitive Psychology, 49,* 370–407.

Wolford, G., Newman, S. E., Miller, M. B., & Wig, G. S. (2004). Searching for patterns in random sequences. *Canadian Journal of Experimental Psychology, 58,* 221–228.