

Novelty and Regularization: The Effect of Novel Instances on Rule Formation

Elizabeth Wonnacott and Elissa L. Newport
University of Rochester

1 Background

The current studies are part of a program of research exploring the learning of inconsistent linguistic structures. Inconsistent input is unusual in normal language acquisition, but may occur in certain atypical language environments, for example when the learner is exposed to a pidgin language or to the language system of a late learner. Singleton and Newport (2004) have studied the language development of a single child who received ASL input from his late learner parents. This child, whom they named Simon, was profoundly deaf and received no other ASL input. Singleton and Newport studied Simon's and his parents' production of a particular class of morphemes in ASL (morphemes of movement). The parents were found to use these morphemes *inconsistently*, producing them around 70% of the time in obligatory contexts. In contrast, Simon used these forms *very* consistently, around 90% of the time in obligatory contexts. This level of consistency was equivalent to that seen in the productions of control children who had been exposed to native ASL. Thus Simon did not reproduce the probabilistic structure of his input, but instead reorganized that input so that his own linguistic system was rule-governed. We will refer to this process as 'regularization.'¹

More recently the learning of inconsistent input has been studied in the lab using an artificial language learning paradigm. This allows us to explore the precise circumstances in which regularization will occur. There are two motivations behind this work. First, we hope to illuminate the processes at work when children such as Simon are exposed to and regularize inconsistent linguistic input. Second, we hope to reveal learning biases which may apply in *normal* language acquisition, biases which might not be visible when learners are exposed to complete linguistic input.

* This research was supported in part by NIH grant DC00167. We are grateful to the Newport/Aslin lab group for helpful suggestions on both our research and this presentation.

1. Singleton and Newport call this 'frequency boosting'.

1.1 Previous findings: Hudson and Newport (2001, 2005)

Hudson and Newport (2001, 2005) conducted a series of experiments investigating the learning of inconsistencies. They created a set of miniature artificial languages in which *determiner* forms occurred with nouns with varying levels of inconsistency. For example, they had a 60% condition, in which 60% of the occurrences of each noun were accompanied by a determiner and 40% were accompanied by no determiner. They found that adult subjects who were exposed to such a language did *not* regularize, but instead matched the statistics of their input. For example, in the 60% condition, they themselves used determiners 60% of the time in their own productions.

However, Hudson and Newport also discovered conditions in which subjects *did* regularize. For adult subjects, regularization depended on the form of the inconsistency in the input. When two variant forms were used in free alternation, e.g. the determiner present versus determiner absent condition described above, the most frequent form was *not* regularized. However, when many different determiners were used, but one form was much more frequently and consistently used than the others, learners *did* regularize that most consistent form. They also found that, across different conditions, children were more likely to use rule-based productions than adults.

2 The present studies

In this work we explore further circumstances in which *adult* subjects will regularize inconsistencies. Hudson and Newport showed that adult learners are highly sensitive to the statistics of their input and tend to reproduce these veridically in their own productions. Under what circumstances do learners pay less attention to the details of their input and more attention to general patterns?

In these studies we ask the following questions: What happens if adult learners are asked to produce utterances for which there are no specific instances in the input? Will they now form rules to guide their productions? We investigated these issues by teaching subjects an inconsistent structure using one set of vocabulary, and then asking them to use that structure with novel vocabulary. Across two experiments we examined the effects of this contrast on two different structures, word order (Experiment 1) and determiners (Experiment 2), each used inconsistently and tested in similar ways.

3 Experiment 1: Regularizing word order inconsistencies

3.1 Method

Subjects. The subjects were 24 native English speakers, all students at the University of Rochester. Subjects were randomly assigned to one of three conditions, with 8 subjects in each condition. All subjects were run and tested individually and were paid for their participation.

Description of the Language. The miniature artificial language had 4 transitive verbs (*glim* [hit], *norg* [kiss], *flern* [hug], *frag* [headbutt]) and 5 nouns referring to 5 puppet animals (*flugat* [bee], *blergen* [lion], *tombat* [giraffe], *slagum* [ladybug], *nagid* [elephant]). All sentences were 3 words long, but word order was inconsistent. For example, ‘The bee hit the lion’ could be expressed in either of the following ways:

glim flugat blergen (VSO: HIT BEE LION)

glim blergen flugat (VOS: HIT LION BEE)

The input language was structured so that 66% of sentences had VOS word order and 33% had VSO word order. (We chose VOS to be the more regular word order, since this was judged to be the least like English of the two orders).

Procedure. *Stage 1: Vocabulary Training.* Subjects were taught the complete set of nouns by having them view scenes in which the 5 puppet animals performed intransitive actions, each accompanied by the appropriate sentence presented aurally. Because intransitive actions were expressed in short two-word sentences and used verbs that were not tested subsequently, the main function of this portion of the training was to teach subjects the vocabulary. These were all the nouns that subjects needed to know when they produced their own sentences at the end of the experiment. Subjects were trained on these sentences until they could name all the animals on a vocabulary test in which they saw the animal and produced the appropriate noun. All of our subjects succeeded in doing this.

Stage 2: Sentence Exposure. 24 transitive sentences were presented aurally, along with short video clips showing puppets performing transitive actions (e.g. they heard *glim flugat blergen* and saw a video of the bee hitting the lion). The input was structured so that each scene occurred three times: twice with a VOS sentence and once with a VSO sentence. Thus the VOS structure occurred:

- For 66% of all scenes.
- For 66% of occurrences of *each* scene.
- For 66% of occurrences of each verb.

Crucial to the design is that only 2 out of the 5 possible nouns were used in these sentences (*flugat* and *blergen*). This means that 3 nouns were reserved for testing. In addition, only 3 of 4 possible verbs were used in these sentences (*glim* [hit], *norg* [kiss], *flern* [hug]), reserving one verb for testing.

Stage 3: Production Test. Subjects were presented with a set of scenes along with the first word of the sentence, i.e. the verb. Their task was to produce the complete transitive sentence. Three different tests were given to three different groups of subjects in three different conditions:

(1) *Old Words Test:* 6 scenes were presented, all of which had previously occurred during sentence exposure.

(2) *New Words Test:* 6 scenes were presented, all of which were entirely new, involving the 3 nouns used in vocabulary training but not during sentence exposure, and the new verb. (Subjects were warned that sentences would contain

a new verb, but since it was given to them in the test, they could readily use it in their own productions).

(3) *Mixed Test*: 12 scenes, with the test items from *Old Words Test* and *New Words Test* combined into one long test set, with the two types of test items interspersed. (This condition is thus a replication of the first two conditions using a within-subjects design).

In all test conditions subjects produced sentences by repeating the verb and providing the two nouns. For example, they if they saw a video depicting *BEE HIT LION* and heard 'glim' they could produce the sentence 'glim flugat blergen' or 'glim blergen flugat'. The order in which they placed the subject and object nouns gave a VSO or VOS structure to their production.

3.2 Results

We began by scoring how often subjects used the form that was dominant in their input. Figure 1 shows the percentage of sentences produced with VOS word order in the *Old Words Test* and *New Words Test* conditions.

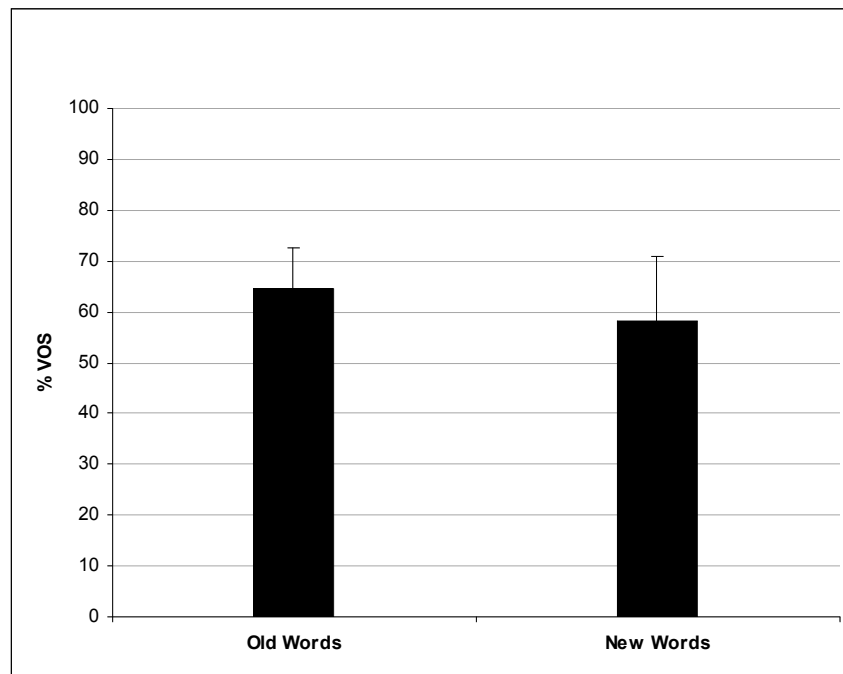


Figure 1: % of sentence productions with VOS order

In the *Old Word Test* condition, as in Hudson and Newport, subjects seem to be probability matching: they used the VOS word order 65% of the time,

approximately as often as they heard it used in the input. In the *New Words Test* condition, subjects used VOS 58% of the time, though there is more individual variation. On this measure there was no significant difference between the conditions, ($t=0.33$, $p=0.74$, $df=14$) and no indication of regularization in either condition.

However, what if some subjects were regularizing the non-dominant VSO word order? (This would actually be in line with what Hudson and Newport have found in child learners, who often regularized non-dominant patterns in the input). If different subjects used different word orders, our method of scoring wouldn't capture this regularization. Rather, we need to look at the performance of individual subjects. We therefore rescored the data, asking how many subjects used one word order consistently and regularly, regardless of whether this was the dominant word order in the input². Figure 2 shows the percentage of subjects in the *Old Words Test* and *New Words Test* conditions who used one word order, either VOS or VSO, across *all* of their productions.

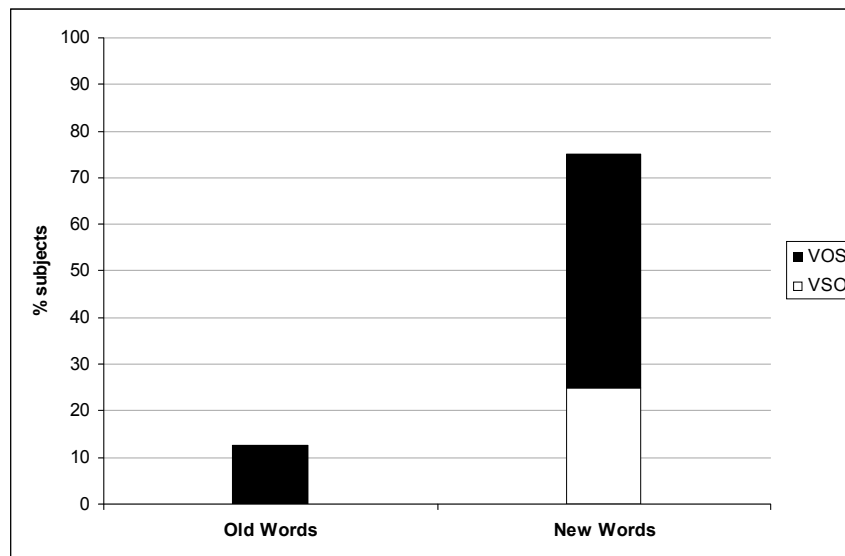


Figure 2: % of subjects using one word order in 100% of their productions

It can be seen that the majority of subjects in the *New Words* condition *were* in fact regularizing, though a third of these subjects were using the non-dominant VSO order. Overall, only 13% of subjects regularized word order in

2. Note that we also checked for regularization on a verb by verb basis, but subjects did not regularize in this manner.

the Old Words condition, while 75% of subjects regularized in the New Words Condition. This difference is significant ($\chi^2=6.34$, $df=1$, $p=0.01$). Thus there is a significant effect of using novel vocabulary. In this case the majority of subjects appear to have formed rules to guide their productions.

We now consider the results from the *Mixed Test* condition. This is a within-subjects version of the previous two conditions. This condition allows us to see if producing both types of test items affects subjects' performance in either condition. In Figure 3 the black bars show the number of subjects in the *Mixed Test* condition who used one word order across all their productions of the different types of test items: Old Word and New Word. For comparison, the white bars show the comparable results from the previous conditions.

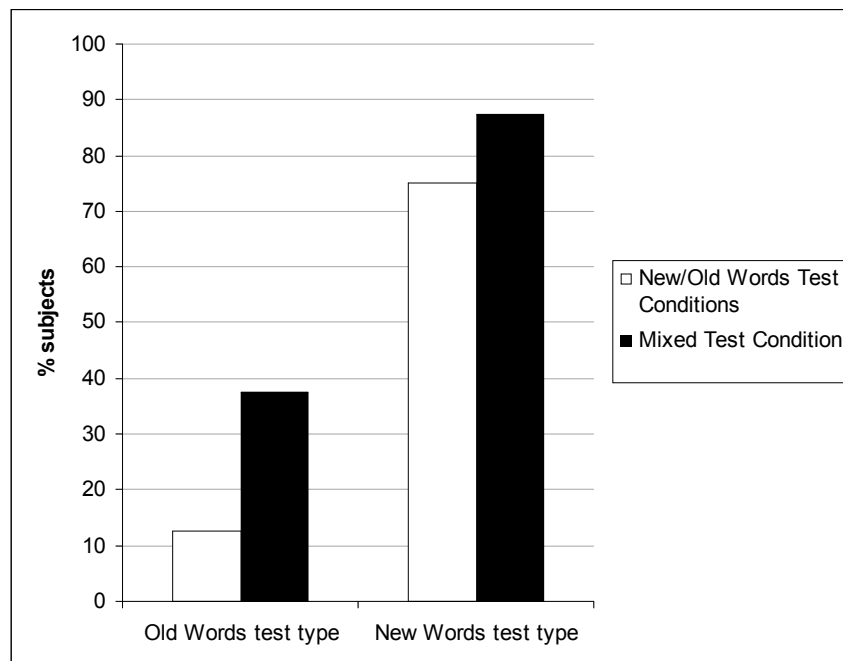


Figure 3: Mixed Test condition: % of subjects who regularized each test type, compared to results from previous conditions

For new word test items it is clear that getting the mixed test has not weakened regularization for novel words: 88% of subjects in the *Mixed Test* condition regularized, compared to 75% in the *New Words* condition ($\chi^2=0.41$, $df=1$, $p=0.52$). Interestingly, there is an increase in the number of subjects regularizing old word test items in the mixed condition: 38% of subjects compared to 13% in the *Old Words Test* condition. The increase is not significant ($\chi^2=1.33$, $df=1$, $p=0.25$) but is suggestive, particularly in light of

equivalent results in Experiment 2 (see below). We hope to run more subjects in the future to see if this effect is real. If so, a tentative explanation is that if we can we induce subjects to form a rule by having them use the structure with novel vocabulary, at least some subjects will also apply that rule when producing old word sentences.

4 Discussion of Word Order experiments

The regularization seen with new word test items contrasts with Hudson and Newport's finding that adult subjects did *not* regularize one of two competing variant forms. Of course that study did not test with novel vocabulary. In addition, the inconsistent structure in their language was determiner usage, not word order. Thus in Experiment 2 we investigated whether adult subjects would also regularize variable determiner usage when using novel vocabulary.

5 Experiment 2: Regularizing determiner inconsistencies

5.1 Method

Subjects. Subjects were 39 native English speakers, all students at the University of Rochester. Subjects were randomly assigned to one of three conditions, with 13 subjects in each condition. All subjects were run and tested individually and were paid for their participation.

Description of the Language. The language was exactly like that used in the Word Order experiments, except that the word order was now fixed as VSO (chosen to be like that used by Hudson and Newport), and nouns could now be followed by a determiner form 'ka.' Examples of possible sentence for the scene BEE HIT LION are:

glim flugat blergen ka
glim flugat ka blergen
glim flugat blergen
glim flugat ka blergen ka

In the exposure set, determiners occurred with nouns 66% of the time.

Procedure. The procedure was exactly as in Experiment 1, with two exceptions. First, the language/sentence exposure set differed as already described. Second, in this experiment the vocabulary was trained by providing actual vocabulary training, rather than doing this training indirectly via intransitive sentence training.

Stage 1: Vocabulary Training. Subjects were taught the complete set of nouns by viewing static pictures of the 5 puppet animals along with their names. Subjects practiced the vocabulary until they could name all the animals on a vocabulary test in which they saw each picture and produced the appropriate noun. All of our subjects succeeded in doing this.

Stage 2: Sentence Exposure. The exposure set was composed of the same scenes used in the Word Order experiment, i.e. those involving the 2 nouns *flugat* and *blergen* and the 3 verbs *glim*, *norg*, and *flern*, but in this experiment the scenes were accompanied by VSO sentences with optional determiners. The input set was structured so that:

- determiners appeared with nouns 66% of the time
- each noun occurred with a determiner 66% of the time
- determiners occurred equally often with subjects and objects.

Again there were three test conditions: *Old Words Test*, *New Words Test* and *Mixed Test* conditions. The tests given in each condition were identical to those in the Word Order experiment. Again subjects produced their own sentences, but in this experiment we looked to see whether they produced determiners with each of the two nouns.

5.2 Results

We began by simply scoring how often subjects produced determiners with nouns. Figure 4 shows the percentage of nouns produced with a determiner by subjects in the *New Words Test* and *Old Words Test* conditions.

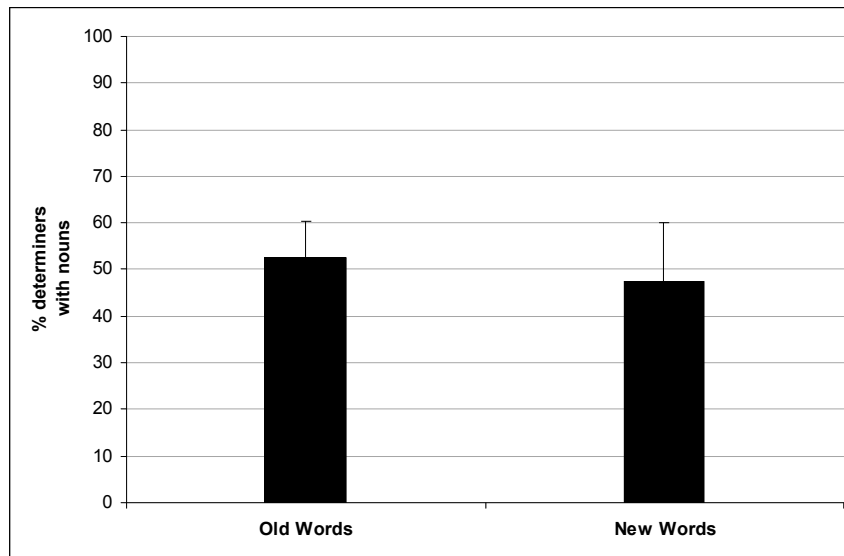


Figure 4 : % of nouns produced with determiners

52% of nouns were accompanied by determiners in the *Old Words Test* and 47% in the *New Words Test*. This difference is not significant ($t=0.44$, $df=24$, $p=0.67$). This method of scoring thus yields no evidence of regularization. As in the Word Order experiment, subjects appear to be probability matching.

(Subjects are slightly undershooting the probabilities of determiner occurrence in sentence input, but this is probably due to their exposure to bare nouns during vocabulary training. Hudson and Newport also saw this undershooting when they trained on vocabulary in the same way.)

However, just as in the Word Order experiment, it could be the case that subjects in the *New Word Condition* were regularizing their own individual patterns. In fact, given the input they received in this experiment, there are several different ways in which subjects could regularize the use of determiners. We observed the following types of regularization:

- Regularization of determiners in subject position, i.e. always produce determiners with the subject noun or never produce the determiner with the subject noun (usage on object noun can be variable)
- Regularization of determiners in object position, i.e. always produce determiners with the subject noun or never produce the determiner with the subject noun (usage on subject noun can be variable)
- Regularization of determiners in both sentence positions, i.e. always produce determiners with nouns or never produce determiner with nouns

Figure 5 shows the number of subjects who regularized in one of these ways in each condition.

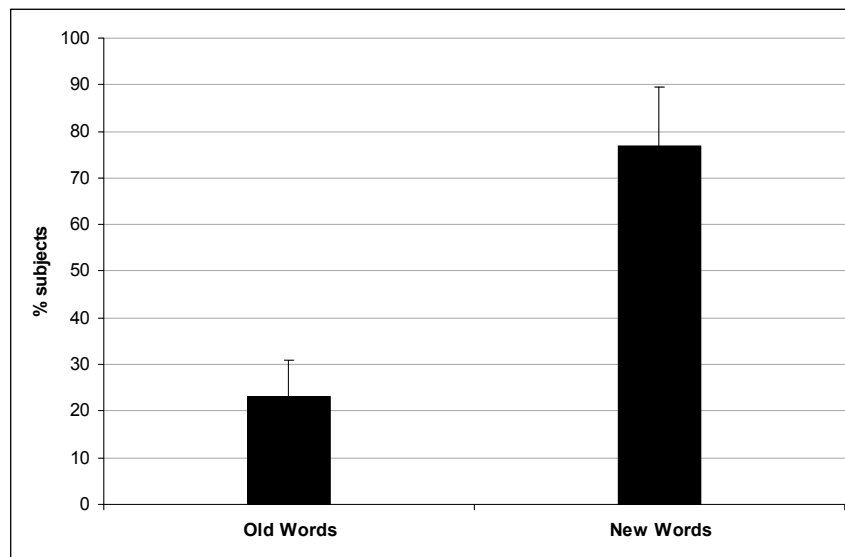


Figure 5: % of subjects using regular pattern

As in the word order experiments, then, many more subjects regularized in the *New Word Test* condition than in the *Old Words Test* condition (77% compared to 23%, a significant difference ($\chi^2=7.5$, $df=1$, $p=0.006$)).

Figure 6 also includes results from the *Mixed Test* condition. As before, we have separated subjects' performance on the different types of test item.

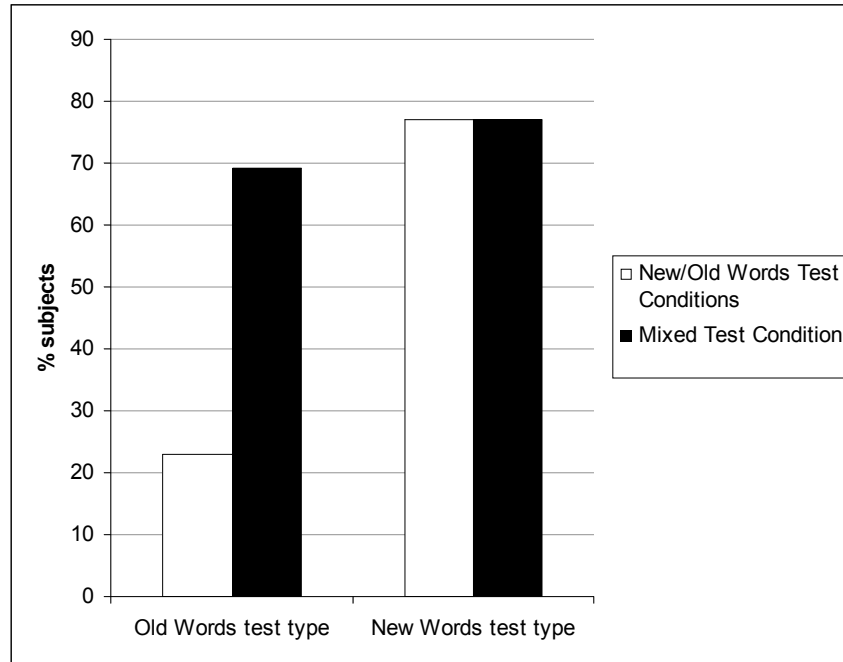


Figure 6: Mixed Test condition: % of subjects who regularized, compared to results from previous conditions

Again regularization of new word test items has not been affected by subjects seeing both test types: 77% of subjects regularized across new word test items in both the *New Words Test* and *Mixed Test* conditions. Interestingly, for old word test items we see the same pattern that we saw more weakly in the Word Order experiment. Regularization for old word sentences was more likely for subjects in the *Mixed Test* than in the *Old Words Test* condition: 69% of subjects regularized versus 23%, ($\chi=5.6$, $df=1$, $p=0.02$). In this experiment the difference is significant, suggesting that there is an effect of seeing new word test items even for performance on the old word test items.

6 Discussion

In these studies we asked adult subjects to abstract from their input and use an inconsistent structure with novel vocabulary. In this situation, when they did not have access to particular instances in their input, adult subjects were likely to regularize and form rules to guide their productions. This contrasts with the

situation in which these subjects were asked to use only old vocabulary, when they tended to match the probabilistic structure of their input.

One interpretation of these results is that learners are constrained by *two* biases. On the one hand, they are very sensitive to their input and want to match its statistics as closely as possible in their own productions. This bias explains the probability matching seen in the Old Words Conditions, and in the original Hudson and Newport experiments. However, we suggest that there is also a second bias: learners expect constructions in a language to be rule-governed. When we ask subjects to form sentences with novel words, we are asking them to produce utterances that are very different from those they heard in the input. This moves their focus away from the input, and they may show the second bias, forming rules to guide their productions. Once a rule has been created, it may also be used for old word test items, if subjects are asked to produce those as well.

7 Future Work

Our next step is to conduct these studies with child subjects. We know from the Simon study, and also from Hudson and Newport, that children regularize in a broader range of circumstances than adults. Thus it is possible that, in contrast to adults, children will regularize *both* new and old word sentences in these experiments. Two accounts would be compatible with these results. The bias toward regularizing their input, rather than preserving its details might be inherently stronger for children than for adults. In addition, children might be less capable of retaining the details of their input, and thus would be more likely to fall back on their bias to regularize (cf. Newport, 1990, for the Less is More hypothesis). However our results turn out, we expect that testing learners on both new and old vocabulary will help us to understand further the process of regularizing inconsistent input.

References

- Hudson Kam, Carla L., & Newport, Elissa L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, in press.
- Newport, Elissa L. (1990). Maturation constraints on language learning. *Cognitive Science*, *14*, 11-28.
- Hudson, Carla L. & Newport, Elissa L. (2001) Regularization during creolization: The learning of inconsistent linguistic input. Boston University Conference on Language Development, November 2001.
- Singleton, Jenny L., & Newport, Elissa L. (2004). When learners surpass their models: The acquisition of American Sign Language from inconsistent input. *Cognitive Psychology*, *49*, 370-407.