Title
Testing the Tolerance Principle: Children form productive rules when it is morecomputationally efficient to do so

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 38(0)

Authors
Schuler, Kathryn D.
Yang, Charles
Newport, Elissa L.

Publication Date
2016

Peer reviewed

Testing the Tolerance Principle: Children form productive rules when it is more computationally efficient to do so

Kathryn D. Schuler (kds71@georgetown.edu)
Department of Neurology, 4000 Reservoir Road NW
Washington, DC 20057 USA

Charles Yang (charles.yang@ling.upenn.edu)
Department of Linguistics, Computer Science &
Psychology, 3401 Walnut Street
Philadelphia, PA 19104 USA

Elissa L. Newport (eln10@georgetown.edu)
Department of Neurology, 4000 Reservoir Road NW
Washington, DC 20057 USA

Abstract

During language acquisition, children must learn when to generalize a pattern – applying it broadly and to new words ('add *–ed*' in English) – and when to restrict generalization, storing the pattern only with specific lexical items. One effort to quantify the conditions for generalization, the Tolerance Principle, has been shown to accurately predict children's generalizations in dozens of corpus-based studies. This principle hypothesizes that a general rule will be formed when it is computationally more efficient than storing lexical forms individually. It is formalized as: a rule *R* will generalize if the number of exceptions does not exceed the number of words in the category *N* divided by the natural log of *N* ($N/\ln N$). Here we test the principle in an artificial language of 9 nonsense nouns. As predicted, children exposed to 5 regular forms and 4 exceptions generalized, applying the regular form to 100% of novel test words. Children exposed to 3 regular forms and 6 exceptions did not extend the rule, even though the token frequency of the regular form was still high in this condition. The Tolerance Principle thus appears to capture a basic principle of generalization in rule formation.

Keywords: artificial language; language acquisition; productivity; morphology; computational modeling.

Introduction

When children learn a language, they do not just memorize words or sentences; they acquire the patterns by which words and sentences are formed. We call these *rules* – for example, "add *–ed*" for the past tense of a verb or "add *–s*" to make a noun plural. However, in some cases there are not broad patterns for how words change their form; sometimes there are many idiosyncratic ways in which individual words form the plural or the past tense, such as *mouse* / *mice* or *go* / *went* – and these must indeed be memorized.

Several types of evidence show that children acquire rules when they are available. For example, young children make overgeneralization errors, extending rules to words that are actually lexical exceptions (e.g. *daddy goed to the store; I saw two mouses yesterday)* (Marcus et al. 1992, Pinker 1992, 1995, Yang 2002, Maslen et al. 2004). Furthermore, when children are asked to produce novel lexical forms in controlled experiments, they can spontaneously apply these rules to nonsense words they have never heard before (Berko 1958). In the famous 'wug' test, children were exposed to novel items from various linguistic categories and were asked to provide their inflected forms. (For example, *This is a wug. Now there are two of them. There are two* ___.) Children demonstrated that they had acquired productive rules by applying regular inflections (e.g. *wugs*) in these novel cases (Berko 1958).

However, not all linguistic rules are productive. Some inflected forms are idiosyncratic to a single lexical item, as noted above. In addition, some rules apply to only a restricted subset of lexical items, like *sing/sang/sung* and *ring/rang/rung*, but are not productive in that they do not apply broadly to new words. Children presented with novel forms like *gling* (similar to irregulars like *ring* and *sing*) do not produce *glang* as the past-test form (Berko, 1958). Though there are a handful of examples of irregular rules being generalized during language acquisition (e.g. *wipe-wope*, *think-thunk*) (Pinker 1999), this type of over-irregularization is unattested in the CHILDES corpus (Yang 2015). In other data, over-regularization errors involving *-ed* are relatively common (about 8%; Maslen et al. 2004), whereas analogical errors following irregular patterns are exceedingly rare (no more than 0.2%; Xu & Pinker 1995). Importantly, this absence of over-irregularization is not exclusive to English acquisition; it has been documented in children's naturalistic production data in many other languages (see e.g. Clahsen and Penke 1992 on German, Allen 1996 on Inuktitut, Clahsen et al. 2004 on Spanish, Caprin and Guasti 2009 on Italian, Demuth 2003 on Bantu languages, Deen 2005 on Swahili, among others).

The tendency to generalize some rules but to restrict others motivates the question: what governs when children will form productive rules during language acquisition? Researchers have been investigating rule learning in language acquisition for many years, but most work has focused on the difference between learning rules and learning the exceptions (not on the difference between learning regular and irregular rules). Though it is generally agreed that learners must memorize the idiosyncratic exceptions (e.g., *go-went*), how regular and irregular rules are handled is still debated. Some have proposed that the irregular rules must be memorized in the same way that the exceptions are (Pinker and Ullman 2002). Others have proposed that everything must be memorized, including both the regular and irregular rules (McClelland and

Patterson 2002). While this argument is not resolved, the available behavioral evidence, most clearly from children's production, points to a strong distinction between rules that are productive and those that are not. There must therefore be a mechanism during learning that governs when productive rules will be formed.

Recently Yang has proposed the Tolerance Principle (2005, 2016) – a model of productivity based on the acquisition literature that takes this categorical distinction into account. It quantifies the precise number of exceptions that a productive rule can tolerate before it becomes computationally less efficient than storing all of the lexical items individually. The Tolerance Principle accurately predicts children's generalizations in corpus data from dozens of rules/patterns in a number of languages, demonstrating that it is a viable model of productivity in language acquisition. However, in order to have adequate data for predictions, corpus analyses combine data from multiple children at different ages, not all of whom may show the same behaviors; and one can only test those patterns that happen to occur in real languages. Here we ask whether the Tolerance Principle can predict when children will generalize a productive rule in an artificial language learning experiment, where we can manipulate the precise number of lexical items that obey a rule or are exceptions. Our results indicate that children do indeed form a productive rule when the Tolerance Principle predicts that they will, applying the regular form to 100% of novel test words. When the Tolerance Principle predicts that no productive rule should be formed, children do not extend the rule, even though the token frequency of the most regular form was still high in this condition. In contrast, adult participants learning the same artificial language appear to extend the rule at the same level as the rule's token frequency in the language, approximating a well-studied phenomenon known as *probability matching* (Hudson Kam & Newport, 2005, 2009). We present this as evidence that the Tolerance Principle appears to capture a basic principle of generalization in rule formation in children, and suggest that adults adopt a different strategy during rule learning.

The Tolerance Principle

The Tolerance Principle (Yang, 2005, 2016) is a learning model that quantifies the precise conditions for generalization during language acquisition. It hypothesizes that a general rule will be formed when doing so is computationally more efficient than storing lexical forms individually. The model computes this computational efficiency by calculating the time complexity of applying a rule compared with accessing individual lexical forms. To illustrate, imagine that a learner is faced with a potential rule – for example, the English 'add *–ed* to make a verb in the past tense.' The English learner has encountered many items that obey this rule (regular forms) as well as many items that do not (irregular forms or exceptions). To be maximally efficient in formulating the past tense of verbs, the learner can do one of two things:

(1) *Store all lexical forms individually*: store every item individually in a list ranked by frequency, searching the list every time there is an occasion to express the past tense of a verb.

(2) *Form a productive rule*: store only the exceptions in a frequency-ranked list. To express the past tense, the learner searches the list of exceptions first. If the target verb is not among these exceptions, the learner applies the rule 'add *–ed*.'

The Tolerance Principle computes the time complexity required for each of these operations and assumes that the learner will adopt the optimal (i.e., faster) strategy. Productive rules, then, are formed only when it is more computationally efficient for the learner.

Formally, if R is a rule that may apply to N lexical items and there are e exceptions to this rule, the time required to access the rule can be expressed as $T(N, e)$. If R is productive, as in (2) above, then the rule is not applied until the learner has first evaluated and rejected every exception (e) on the list. In other words, applying a productive rule consumes e units of time. The time required for exceptions, on the other hand, is determined by the lexical item's frequency (i.e., its rank in the list of exceptions). To compute the time complexity $T(N, e)$, Yang (2016:48) calculates "the weighted average of time units over the probabilities of these two sets of items." If R is unproductive, as in (1) above, then all N items are treated as exceptions and are listed in order of frequency. The time complexity under these circumstances can be expressed $T(N, N)$, as the number of exceptions e is equivalent to the number of items in the list N. It is conjectured that the learner compares the time complexity required to form a productive rule, $T(N, e)$, with the time complexity required when all N items are stored individually as lexical exceptions, $T(N, N)$. By solving this equation for e, the Tolerance Principle computes the precise number of exceptions that a productive rule can tolerate before its formation becomes computationally inefficient.[1] This solution is as follows:

(3) **Tolerance Principle**: Let R be a rule that is applicable to N items, of which e are exceptions. R is productive if and only if $e \leq \theta_N = N/\ln(N)$.

In other words, it is only more efficient to form a productive rule when the number of exceptions is less than the number of items divided by the natural log of the number of items. To illustrate, imagine a category of 9 items. Given a rule R that may apply to these 9 items, the Tolerance Principle predicts that 4.096 (or $\theta_9 = 9/\ln 9$) exceptions will be tolerated before forming a productive rule becomes less efficient than storing individual items. This means that learners will form a productive rule if there are 4 or fewer exceptions to the rule R, but not if there are 5 or more. Importantly, this implies that the distinction between forming a productive rule and storing individual lexical items is a categorical one.

[1] See Yang (2016) for the full derivation of the Tolerance Principle and the supporting psycholinguistic evidence from lexical and morphological processing.

There is a theoretical tipping point at which forming a productive rule becomes less computationally efficient than the alternative strategy. The Tolerance Principle allows us to compute this tipping point.

How well does this model hold up to empirical investigation? Yang has tested the model on corpus data in a number of rule acquisition scenarios and found that the Tolerance Principle predicts productive rule formation surprisingly accurately. For example, consider the English productive rule 'add *–ed* to make a verb past tense.' Yang analyzed the 5 million words of child-directed English from CHILDES (MacWhinney 2000) and found 1022 unique past-tense verbs. By the Tolerance Principle, the English 'add *–ed*' rule should tolerate 147 verbs that are exceptions in a class of 1022 lexical items. Yang's analysis found only 127, well below the tipping point (θ_{1022}=147) for computational efficiency. The irregular rules/patterns in English, however, do not fare so well. Even the irregular class that has the highest homogeneity, the *ing*-*ang* class (such as *sing*-*sang* and *ring*-*rang*), has too many exceptions. The CHILDES English input corpus has 8 verbs that end in *ing*, but only three change the past tense to *ang* (*ring*, *sing*, *spring*) and five do not (*bring*, *fling*, *sting*, *swing*, *wing*), exceeding the threshold of θ_8=3. Thus the *ing-ang* pattern is predicted to be unproductive, in accord with children's productions described in the Introduction.

Though Yang has provided substantial evidence from corpus analyses to validate his account, further empirical investigation is necessary to demonstrate that children indeed follow the Tolerance Principle in acquiring productive rules during language acquisition. Here we will apply two well-known acquisition paradigms – artificial language learning and the "wug" test – to submit the Tolerance Principle to further experimental scrutiny. We use an artificial language paradigm to precisely control the input to child learners, providing them with highly controlled frequencies and numbers of words that follow a rule and words that are exceptions. This allows us to create conditions where the Tolerance Principle predicts productive rule formation (as in the *–ed* example above), and where it predicts that a rule will be unproductive (as in the *ing*-*ang* example above). We ask whether the Tolerance Principle correctly predicts when a pattern in an artificial language is widespread enough for a child to form a productive rule, using a "wug" test to assess whether children have formed a productive rule (one that applies to novel lexical items) or have restricted generalization.

Behavioral Data

Participants

Fifteen children (mean = 7.48 years, range = 5.08 – 8.92 years) and twenty adult controls participated in this experiment. An additional 3 children began the experiment but did not complete it, and an additional 4 children participated but were excluded from analysis for failure to understand the task (quantified as a failure to produce the correct noun on at least 50% of the test trials). Children were recruited from Washington DC metro area schools and were run either at their school or in our lab. Adult participants were recruited and run online using Amazon Mechanical Turk[2]. All participants were native English speakers with normal hearing and normal to corrected-to-normal vision. Child participants received stickers and a set of small toys for their participation. Adult participants received compensation at a rate of $10/hour.

Stimuli

Description of the language We designed two artificial languages: one in which the Tolerance Principle predicts that learners should form a productive rule and one where learners should not form a productive rule. To do so, we first created a rule *R* for a category of 9 nonsense nouns. The rule was: "To make a noun plural, add *ka*." Next we used the Tolerance Principle to calculate the number of regular forms vs. exceptions a productive rule can tolerate in a category of 9 nouns. Using the predicted value of 4.096 exceptions, we created two conditions: one where a productive rule should be formed (5 regulars, 4 exceptions), and one where a productive rule should not be formed (3 regulars, 6 exceptions).

To create our exposure corpus, we assigned each noun a plural marker that either followed the rule (add *ka*) or was an exception (add *po, tay, lee bae, muy,* or *woo*), depending on the condition. Then we used these nouns and markers to create an exposure corpus of 72 sentences (24 singular and 48 plural). All sentences began with the same nonsense verb *gentif*, meaning "there is/are". Singular sentences were unmarked ("*gentif* + NOUN") and paired with one image of the corresponding object. Plural sentences were formed "*gentif* + NOUN + MARKER" and paired with 2, 4, or 6 images of the corresponding object. There were thus 18 possible sentences in the language: 1 singular and 1 plural sentence for each noun.

We generated the exposure corpus by allowing noun frequency to vary along a Zipfian distribution, with nouns taking the regular form (*ka*) as the most frequent in both conditions. Thus the second most frequent noun was presented half as often as the most frequent noun, the third most frequent noun was half as often as the second, and so on. This is important because the distribution of word frequency in natural language is approximately Zipfian, and the computation underlying the Tolerance Principle assumes that word frequency follows this pattern. Making the regular form the most frequent ensured that its token frequency was high in both conditions.

Procedure

Exposure Each participant was presented with the 72 exposure sentences in random order. On each trial, participants saw a picture of 1, 2, 4, or 6 instances of a noun

[2] Adults run on Mechanical Turk give the same results as adults run in the lab on the same paradigm as children.

and hear the corresponding singular (for 1) or plural (for 2, 4, or 6) sentence. They were asked to repeat the sentence aloud (or type it into a response box on Mechanical Turk) before moving on to the next trial. Every 18 trials they were given a short break. Children were offered a sticker during breaks to encourage them to continue in the experiment.

Production Test After exposure, we used a "wug" test to assess whether children had formed a productive rule (Berko, 1958). During this test, participants were given singular sentence-image pairs containing novel nonsense nouns they had not heard during exposure and were asked to provide the plural form. Each participant completed 12 production test trials, 2 for each of 6 novel nouns. To prevent participants from using a plural form based only on the precise number of instances shown in that trial, the test items contained 3 or 5 instances of the novel noun (whereas there were 2, 4, or 6 in the exposure set). Finally, all participants were given a rating test to ensure they had learned the nouns and markers they were exposed to.

Results & Analysis

For each production test trial, participants were asked to produce the plural form of a novel noun they heard only in a singular form. These novel productions allowed us to assess whether participants formed a productive rule. Recall that the Tolerance Principle predicts that there will be a categorical distinction between productive and unproductive (lexically specific) rules. In our artificial language, a productive rule should be formed if more than 4.096 nouns obey the rule (as in our 5 regular/4 exception condition), but not if fewer than 4.096 nouns do (as in our 3 regular/6 exception condition). When a productive rule is formed, it should be applied to 100% of novel nouns, as is the case for English past tense 'add *–ed*.'

To determine whether participants formed a productive rule, we performed a one-tailed t-test against the hypothesized value of 100%. Participants who have formed a productive rule should, according to the Tolerance Principle, mark these novel plural sentences with *ka* 100% of the time. On the other hand, participants who have not formed a productive rule should use the *ka* inflection significantly less than 100% (and perhaps no more frequently than other inflection forms are used).

Focusing first on the child production data, Figure 1 shows the percentage that each inflection type was produced during the production test for participants in the 5 regular/4 exception condition (5R/4E) and the 3 regular/6 exception condition (3R/6E). These data show that children in the 5R/4E condition mark novel nouns with the *ka* inflection on 91.7% of plural trials. This value is not statistically different from 100% ($t=1.00$, $p=0.18$). In contrast, in the 3R/6E condition, children mark novel nouns with the *ka* inflection on only 16.9% of plural trials. This value is substantially and significantly different from 100% ($t=6.81$, $p<0.0001$). Children thus appear to have formed a productive rule when the Tolerance Principle predicts that they will (in the 5 regular / 4 exception condition), but not when it predicts that they will not (in the 3 regular / 6 exception condition).

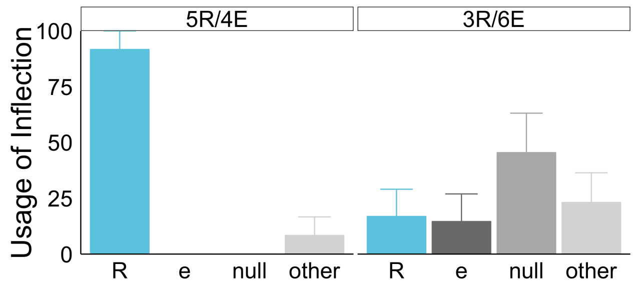

Figure 1: Percentage of each type of inflection added to novel nouns by children when their exposure contained 5 regulars/4 exceptions compared with 3 regulars/6 exceptions. *R* = the regular form; *e* = any exception form; null = no plural marker (unmarked); other = any marker not present in the exposure (e.g. English [+s]).

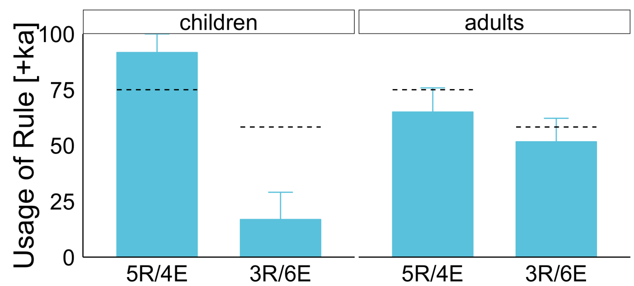

Figure 2: Percentage of regular inflection *ka* applied to novel nouns by children and adults when their exposure contained 5 regulars/4 exceptions compared with 3 regulars/6 exceptions. Dashed line indicates the token frequency of the *ka* inflection in the input.

Table 1: Number of children who applied the rule to 100% of plural test nouns or to fewer than 100% of plural test nouns in each condition.

	5R/4E	3R/6E
# children using rule 100%	6	1
# children using rule <100%	1	7

This strong result is further underlined by looking at the data from individual children. Table 1 shows the number of children who used the *ka* inflection on 100% of test trials in the 5 regular/4 exception condition compared with the 3 regular/6 exception condition. Six out of 7 children produced the *ka* inflection on 100% of production trials in the 5 regular/4 exception condition, while only one out of 8 children did so in the 3 regular/6 exception condition.

Turning next to the adult production data, we find a somewhat different pattern of results. Adults in the 5R/4E condition mark novel nouns with the *ka* inflection on 65.0% of plural trials; unlike children, this value is significantly different from our 100% productivity criterion (t=3.23,

$p<0.01$). Like children, adults in the 3R/6E mark novel nouns with *ka* significantly less than 100% of plural trials ($t=4.59$, $p<0.001$) - only 51.7%. That is, for adults, this contrast is much weaker, and when we compare the use of the *ka* inflection between adults to the children, as in Figure 2, we see striking differences between the two. The Tolerance Principle effect is much more pronounced in children, who exhibit a much more categorical response in their use of the *ka* inflection. Indeed, children, but not adults, show a significant difference between the use of *ka* in the 5R/4E condition and that in the 3R/6E condition (children: $t=4.91$, $p<0.001$, adults: $t=0.89$, $p=0.39$).

One possible explanation is that adults are not obeying the Tolerance Principle and are instead producing *ka* with the same frequency they heard this inflection in their input. This behavior is known as *probability matching* (Hudson Kam & Newport, 2005, 2009). Recall that the nouns in our artificial language follow a Zipfian distribution, with rule-following nouns being the most frequent in the distribution. Thus the token frequency of the *ka* inflection is fairly high in both conditions: 75% of the plural exposure sentences in the 5R/4E condition, and 58.3% of plural exposure sentences in the 3R/6E condition. To determine whether this could explain the difference between the two groups, we analyzed both child and adult use of *ka* against the token frequency of *ka* in the exposure for the two conditions. We found that only adults match the token frequency in both the 5R/4E ($t=0.92$, $p=0.19$) and 3R/6E conditions ($t=0.63$, $p=0.27$). In contrast, the child data is not consistent with a probability matching interpretation. Children in the 5R/4E condition produce the *ka* inflection significantly more than the input frequency ($t=2.00$, $p<0.05$) and in the 3R/6E condition produce the *ka* inflection significantly less than the input frequency ($t=3.40$, $p<0.01$).

Conclusions & Discussion

Here we have asked whether the Tolerance Principle accurately predicts when children will generalize a productive rule in an artificial language learning experiment and when they will restrict generalization. The Tolerance Principle is based on the hypothesis that productivity emerges when it is the most efficient strategy for learners to access lexical forms. The model allows us to calculate the number of lexical exceptions there can be to a given rule before it becomes more efficient for the learner to simply memorize each lexical form individually. For our 9-noun artificial category, the Tolerance Principle predicts that a productive rule will be formed when there are fewer than 4.096 lexical items that are exceptions. We found that, just as predicted, learners formed a productive rule when there were 4 lexical items that were exceptions to the rule, but not when there were 6. These results suggest that the Tolerance Principle has accurately captured something significant about the conditions for generalization during learning.

Importantly, the criterion we used to assess whether learners formed a productive rule was categorical. Our analysis asked whether learners extended the rule to 100% of the test trials – the most rigorous possible test of productivity. We found that, while both children and adults were more likely to extend a productive rule when there were 4 exceptions than when there were 6, only children displayed a categorical distinction between forming a productive rule and not forming one. As predicted by the Tolerance Principle, almost every child exposed to 5 regulars/4 exceptions extended the rule to 100% of test trials, while almost no children exposed to 3 regulars/6 exceptions did (see Table 1). These results for children are in accord with those cited above for regulars like *–ed* versus irregulars like *sing/sang*. While children in the 3R/6E condition did occasionally use *ka* on novel items, their results are only what one would expect under conditions of uncertainty. In the absence of a productive rule, children appear to be unsure how to mark novel forms, most often using no plural marker (null) and otherwise selecting at random from among the various markers they heard during exposure (the *ka* inflection as well as the exceptions).

The striking difference we observed between children and adults, shown in Figure 2, led us to ask whether an alternative model would better predict adult behavior in this task. The Tolerance Principle is based on the number of lexical items (types) that observe or violate a pattern, which is consistent with most approaches to productivity from a wide range of perspectives (e.g., Plunkett & Marchman 1991). By contrast, adults appear to follow more closely the token frequencies. Recall that the token frequency of the *ka* marker was high in both the 5R/4E and the 3R/6E conditions, due to the Zipfian distribution of noun frequency in our exposure corpus – 75% and 58.3%, respectively. We found that adults produced the *ka* marker with the same frequency they heard this marker in their input. This *probability matching* behavior has been observed in adult learners in many other experiments in our lab (e.g. Hudson Kam & Newport, 2005, 2009). There are several ways to interpret the differences between children and adults in the context of computational efficiency, particularly given the greater cognitive resources of adult learners (Newport, 1990). One possibility is that probability matching may be the more efficient computational strategy for adult learners. This would imply that the Tolerance Principle is exclusive to children, capturing a basic principle of generalization in rule formation for very young learners. A related possibility is that only children learn or produce forms categorically. On this interpretation, the difference between 5R/4E and 3R/6E influences both child and adults learners, but only children show this difference in such an extreme contrast in their output. A final possibility is that perhaps adults only engage probability-matching behavior when they are learning from a very small number of items; probability matching may be easy and efficient for adults when there are only a few items to keep track of. However, when there are many items to track, it may no longer be efficient to track and closely match input probabilities. This latter interpretation is supported by the results in studies on a different topic, the effect of inconsistent input. In these studies children and adults look very different in experiments when the learning task involves only a small number of items, with adults probability matching and children maximizing the majority form in their productions. However, when the number and complexity of the items and their variations become very large, adults begin to behave more like children (Hudson-Kam & Newport 2009). In ongoing work we are exploring this by giving adult participants our Tolerance Principle tasks with a very large set of nouns. We hypothesize that, at this higher level of complexity, adults may begin to exhibit the type of categorical behavior predicted by the Tolerance Principle.

As mentioned above, children and adults behave differently both in the present experiments and in the literature on inconsistent input. Though our results suggest that adults adopt a similar strategy to perform these two different tasks, children appear to be handling them differently. In the literature on inconsistent input, children produce the form that appears most often in their exposure nearly 100% of the time (Hudson-Kam & Newport 2005, 2009). While the results of our 5R/4E condition are similar to these findings, the results from our 3R/6E condition are quite different. Although the *ka* marker is the most frequent marker in the 3R/6E exposure (58% of tokens), children produce this form with much lower probability than they were exposed to. This suggests that children are not simply forming productive rules based on what appears most frequently in their input, as they do when they are faced with patterns that are inconsistently marked or probabilistic. In the present paradigm, in contrast with our studies of inconsistent input, each lexical item is marked in a consistent way across trials. When children are exposed to lexically consistent patterns, they form productive rules based on the number of lexical items that observe these patterns, as predicted by the Tolerance Principle. To examine these issues further, we are conducting an experiment in which token frequency is matched across conditions and another in which the rule is assigned to very low frequency nouns.

As noted above, the computations behind the Tolerance Principle are based on the assumption that productive rules emerge when they are the most efficient strategy for learners. Successful models of many cognitive and neural processes are based on the same notion. This implies that the Tolerance Principle may be applicable to rule acquisition and generalization in domains other than language. Future work is required to determine whether this exciting prospect is empirically supported.

Acknowledgments

This research was supported by NIH grant HD-37082 to E. Newport and by Georgetown University. We thank Sarah Furlong, John Davis, and Breana Downey for their important assistance in recruiting and running the participants, and the members of the Language and Development Lab for helpful comments on the research.

References

Allen, S. (1996). *Aspects of argument structure acquisition in Inuktitut.* Johns Benjamins.

Berko, J. (1958). The child's learning of English morphology, *Word,* 14, 150-177.

Caprin, C. and Guasti, M.T. (2009). The acquisition of morpohosyntax in Italian: A cross-sectional study. *Applied psycholinguistics*, 20, 23-52.

Clahsen, H. and Penke, M. (1992). The acquisition of agreement morphology and its syntactic consequences.: New evidence on german child language from the simone corpus. In Meisel, J., editor, *The acquisition of verb placement*, pp 181-234. Kluwer, Dordrecht.

Clahsen, H., Aveledo, F. & Roca, I. (2002). The development of regular and irregular verbs inflection in Spanish child language. *Journal of Child Language,* 29, 591-622.

Deen, U. (2005). *The acquisition of Swahili*. John Benjamins Publishing.

Demuth, K. (2003). The acquisition of Bantu languages. *The Bantu languages*, pp 209–222. Curzon Press. Surrey, UK.

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

Hudson Kam, C.L., & Newport, E.L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology*, 59, 30-66.

Marcus, G., Pinker, S., Ullman, M.T., Hollander, M., Rosen, J., & Xu, F. (1992). *Overregularization in language acquisition.* Monographs of the Society for Research in Child Development. University of Chicago Press, Chicago.

Maslen, R., Theakston, A.L., Lieven, E.V., and Tomasello, M. (2004). A dense corpus study of past tense and plural overregularization in English. *Journal of Speech, Language and Hearing Research*, 47(6): 1319-1333.

MacWhinney, B. (2000). *The CHILDES Project.* Lawrence Erlbaum.

McClelland, J. & Patterson, K. (2002). Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences,* 6, 465-472.

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

Pinker, S. (1999). *Words and rules: The ingredients of language*. Basic Books, New York.

Pinker, S., and Ullman, M.T. (2002) The past and future of past tense. *Trends in Cognitive Science*, 6, 456-463.

Plunkett, K., and Marchman, V. (1991). U-shaped learning and frequency effects in a multi-layered perception. *Cognition,* 38 (1): 43-102.

Yang, C. (2005). On productivity. *Linguistic variation yearbook*, 5(1), 265-302.

Yang, C. (2015). For and against frequencies. *Journal of Child Language.* 42, 287-293.

Yang, C. (2016). *The price of productivity*. MIT press.