Productivity depends on communicative intention and accessibility, not thresholds

Alexia Hernandez (alexiamh@princeton.edu)
Department of Psychology, Princeton University, Princeton NJ 08544
Sammy Floyd (sfloyd@princeton.edu)
Department of Psychology, Princeton University, Princeton NJ 08544
Adele E. Goldberg (adele@princeton.edu)
Department of Psychology, Princeton University, Princeton NJ 08544

Abstract

When do children extend a construction ("rule") productively? A recent *Threshold* proposal claims that a construction is productive if and only if it has been witnessed applying to a sufficient proportion of cases and sufficiently few exceptions. An alternative proposal, *Communicate and Access* (C&A), argues that children extend a construction productively because they wish to express an intended message and are unable to access a "better" (appropriate and more conventional) way to do so. Accessibility, in turn, is negatively affected by interference from competing alternatives. In a preregistered experiment, 32 4-6-year-old children were provided with exposure to 2 mini-artificial languages for which the two proposals make opposite predictions. Results support the C&A proposal: children were more productive after witnessing 3 rule-following cases than after 5, due to differences in interference. We conclude that productivity is encouraged by a desire to communicate a message and is constrained by accessibility and interference.

Keywords: productivity, communication, accessibility, Tolerance Principle, Sufficiency Principle

Introduction

When children learn a new noun, *wug*, they are quite adept at producing its plural, *wugs* (Berko, 1958). On the other hand, the *-th* nominalizing suffix (*warmth, width)* is not generally added to new cases *(?coldth; ?oldth*) outside the domain of ordinal numbers (*gazilionth*). A recent *Threshold* proposal has attempted to predict when rules "go" productive and when they do not (Yang, 2016). In particular, in order for a rule to be productive, a Tolerance Principle offers a ceiling on the number of witnessed exceptions and a Sufficiency Principle suggests a floor on the number of cases witnessed following the rule. The required calculations are based on the following 3 numbers:

1) # of cases which potentially follow a rule: N
2) # of witnessed exceptions to a rule: e
3) # of witnessed rule-following cases: M

Specifically, the upper bound on exceptions and lower bound on rule-following cases have been proposed according to the thresholds in (1) and (2) (Yang 2016):

  (1) Tolerance Principle (TP): $e \leqslant N/\ln N$
  (2) Sufficiency Principle (SP): $M \geq N - N/\ln N$

For instance, in a domain of size 9, for a rule to be used productively, the minimum number of cases that must be witnessed following a rule is 5 (Sufficiency Principle) and the maximum number of exceptional cases is 4 (Tolerance Principle) (Table 1; Yang 2016; Schuler, Yang, & Newport 2016).

Table 1: The Threshold numbers predicted by Sufficiency and Tolerance Principles (Yang, 2016; SYN '16).

Domain Size (N)	Minimum # of rule-following cases (M): N- N/lnN	Maximum # of exceptions (e): N/ln N
9	**5**	**4**

A prior study (Schuler, Yang, & Newport, 2016, hereafter SYN '16), aimed to test the predictions in Table 1, but as explained below, the results are open to a different interpretation. The alternative proposal, which we refer to as Communicate and Access (C&A), takes as its starting point the idea that learners aim to convey their messages while obeying the conventions of the language as best they can (Goldberg, 2019). In order to *be able to* use a new language to express an intended message in an appropriate way, children need to be able to *access* the appropriate form. Accessibility is positively affected by the availability of a target form (Bybee, 2010) and is negatively affected by interference from contextually relevant competitors (Bates & MacWhinney 1987; Montag et al. 2017). We report new data involving two new experimental conditions that unconfound the predictions of the two proposals.

SYN '16 aimed to test the predictions in Table 1 by exposing 5-8-year-old children to a rule that could potentially apply to 9 cases in one of two conditions. In a **5R/1-1-1-1E** condition, the rule applied to 5 cases and 4 other cases were witnessed being exceptional, with each exception being exceptional in its own way. In this case, the domain size (N) was 9, the number of cases witnessed following the rule (M) was 5, and the number of exceptions (e) was 4. Because each exceptional case was unique, we represent the 4 exceptions here as 1-1-1-1. This 5R/1-1-1-1E condition satisfied both the Tolerance and Sufficiency principles and, as predicted by SYN '16, children treated the rule as fully productive. In a **3R/1-1-1-1-1-1E** condition, children saw a rule applied to 3 cases and 6 other cases were witnessed being exceptional. Here the Sufficiency Principle was violated (at least 5 rule-following cases should be required for productivity), and there were more exceptions than allowed by the Tolerance Principle. As predicted by SYN '16, children did *not* extend

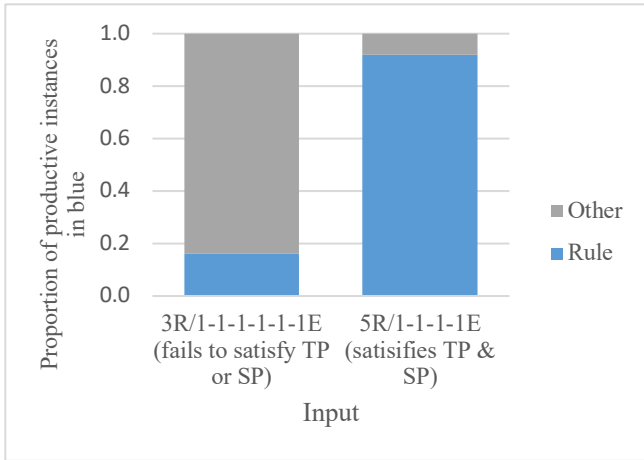the rule to new cases in this 3R/1-1-1-1-1-1E condition (see Figure 1).

Figure 1: Data reported by SYN (2016): Proportion of productive rule-following (blue) and Other (grey) responses. Children exposed to a rule with 3 rule-following and 6 unique exceptions (left) or 5 rule-following and 4 unique exceptions (right).

Critically, the pattern of results reported by SYN '16 is equally consistent with the Communicate & Access proposal. That is, children extend a construction productively when they wish to express an intended message and are unable to access a "better" (appropriate and more conventional) way to do it. From this perspective, productivity is the effect of producing a "good enough" option when no conventional form exists, or when none is sufficiently accessible at the moment of speaking. Accessibility, in turn, is negatively affected by interference from competing alternatives, which themselves may be more or less accessible (Harmon & Kapatsinski, 2017; Macdonald, 2013; Montag, Matsuki, Kim, & Macdonald, 2017).

If we compare the 5R/1-1-1-1E and 3R/1-1-1-1-1-1E "rules" which children were exposed to in SYN '16, the C&A proposal likewise predicts that the 5R/1-1-1-1E rule should be more productive than the 3R/1-1-1-1-1-1E rule, but for different reasons than the Threshold proposal suggests.

Instead of viewing language as requiring abstract rules that are subject to numerical thresholds which render the rules either categorically productive or unproductive, the C&A approach predicts that learners record imperfect (lossy) memory traces that relate linguistic words and phrases to their meanings in context. Therefore, in what follows we refer to emergent generalizations as *constructions* instead of rules in describing the C&A perspective. Other things being equal, a construction is more accessible in memory after being witnessed with a greater variety of distinct cases because variability increases accessibility within the range of witnessed variability, and decreases it outside the range of witnessed exemplars (Tenenbaum & Griffiths, 2001; Suttle & Goldberg, 2011). This follows from the fact that memory is associative and content-addressable. The fact that memory is associative entails that new memories are integrated with existing memories; the fact that memory is content-addressable means that existing clusters of memories are activated to the extent that they are relevantly similar for the purpose of task demands.

Conversely, accessibility is negatively impacted by interference from competing constructions, with interference increasing as the accessibility of the competing constructions increases: witnessing 6 exceptional alternative cases *interferes* with a construction more than witnessing only 4 exceptional cases. Since other things were held constant in SYN '16, the availability of the construction was higher, and interference was lower in the 5R/1-1-1-1E condition relative to the 3R/1-1-1-1-1-1E condition. Therefore, the C&A proposal concurs that the 5R/1-1-1-1E condition should be more productive.

To summarize, the results reported by SYN '16 cannot distinguish between the proposal based on thresholds as determined by Tolerance and Sufficiency Principles, on the one hand, and the Communicate and Access proposal, on the other (Table 2).

Table 2: Convergent predictions are made by Threshold and C&A proposals for the productivity of an unconditioned "rule" with domain size of 9 in conditions tested by SYN '16 on 5-8-year-olds.

M vs. e, M = #Rule-following cases e = exceptional cases		(Shared) Predictions and Results (SYN '16):
3R/ 1-1-1-1-1-1E	Threshold: Neither TP nor SP are satisfied	No systematic productivity
	C&A: Tentative constructional generalization competes with many alternatives: no clear winner emerges	No systematic productivity
5R/ 1-1-1-1E	Threshold: TP and SP are satisfied	Productivity
	C&A: Constructional generalization is more accessible than any alternative	Productivity

In order to compare the two proposals directly, we report a new experiment for which they make *opposing* predictions. Specifically, we exposed a group of 4-6-year-old children to 2 new mini-artificial languages. In a **3R/0E condition**, a novel "rule" was witnessed applying to 3 out of 9 cases with 0 exceptions. The Threshold proposal predicts that children in this condition will not use the rule productively because an insufficient number of rule-following cases are witnessed: recall that in a domain of 9, a minimum number of 5 cases is required for productivity. The C&A proposal predicts, on the other hand, since 0 exceptions were witnessed, there should be no competition. Therefore, the C&A proposal predicts that as long as children understand the function of the construction and are able to access it, a construction that is

witnessed applying to 3 cases and 0 exceptions *will* be used productively.

In a separate **5R/4E condition**, a second novel rule was witnessed applying to 5 out of 9 cases, while 4 other cases were exceptional. The only difference between this 5R/4E condition and the 5R/1-1-1E condition in SYN '16 is that here the 4 exceptional cases behaved alike. In both cases, there were 5 rule following cases and 4 non-rule following cases. Therefore, the Threshold proposal predicts the 5R/4E rule should be as categorically productive as the 5R/1-1-1E rule of SYN '16 was. The C&A proposal, on the other hand, predicts that the "exceptional" construction—which was applied to 4 entities—should interfere with learners' ability to access the higher type frequency construction—which was applied to 5 entities. Because there is only a 25% difference in availability (and interference) between the two patterns, and no conditioning factors that could systematically distinguish the two, interference should render the (slightly) more dominant construction—the one witnessed applying to 5 entities—less than fully productive. C&A further predicts that when the more dominant construction is not used, the competing, less dominant construction will be used instead.

To summarize, the Threshold proposal predicts that when exposed to the 3R/0E rule, children should treat it as completely unproductive, and when exposed to the 5R/4E rule, they should treat it as completely productive. The C&A proposal, on the other hand, predicts that the 3R/0E pattern should be productive because it has no competition. As long as children are able to understand the task, they should use the pattern productively. In the 5R/4E condition, C&A predicts that the dominant pattern should be subject to interference from the less dominant pattern and should therefore be less than fully productive. Again, if children fail to use the dominant construction, C&A predicts that they will use the competitor, less-dominant construction instead. The predictions of the two proposals are represented in Table 3.

Table 3: Predictions of the Threshold and Communicate and Access proposals.

NEW CONDITIONS:
M vs. e,
(M= #Rule-following cases
e = exceptional cases)
		PREDICTIONS:

3R/0E	**Threshold:** SP is not satisfied	Rule$_3$ should not be productive
	C&A: Tentative constructional generalization has no interference from alternatives	Construction$_{(3)}$ should be productive
5R/4E	**Threshold:** TP and SP are satisfied	Rule$_{(5)}$ should be productive
	C&A: Dominant construction is only 25% more accessible than interchangeable alternative construction	Construction$_{(5)}$ and Construction$_{(4)}$ should compete

Experiment

Preregistration at Open Science Framework (OSF). We preregistered a plan to test 16 children without counterbalancing the constructions across conditions, and to use *t*-tests against full and 0 productivity (following SYN '16). We subsequently preregistered a second design with another 16 children in order to counterbalance the constructions (plural vs. classifier) and in order to preregister a more appropriate mixed model (glmer) analysis. Data was collected for each experiment only after it was preregistered. Results are combined below, as is appropriate, but both groups of participants were also analyzed separately (the first group with and without the 5 additional children tested with slightly different instructions). The pattern of results reported below remain unchanged in these subgroups.

Methods

Participants

32 children between the ages of 4 and 6 ($M = 56$ months) are analyzed below. We changed the instructions after an initial 5 children were tested and these children are excluded from analysis. All but one child provided four critical responses, two in each condition. One child opted out after the first condition (which happened to be 3R/0E for this child). All children were tested at the Princeton University Baby Lab, two were bilingual English speaking and the rest were monolingual English speakers. All had normal hearing and vision and were born at full term (38+ week gestation). After each question, children received a sticker regardless of their response, and after the study, each child received a book and a prize, and the family received $10.

Procedure

The design was within-participants. In each of the two conditions, children were exposed to a mini-language that included 1 or 2 novel words, and 9 familiar English words naming each of 9 distinct kinds of animals or crayon colors. In the 3R/0E condition, a single novel form (*po*) was witnessed being used with 3 out of 9 items. In the 5R/4E condition, one form (*dax* or *fep)* was randomly assigned to 5 of the 9 items, and the other form was assigned to the remaining 4 items (see Figure 2).

The following were counterbalanced (in a nested fashion) across participants:

- order: whether children witnessed the 3R/0E or the 5R/4E condition first
- function: whether the rule/construction tested had a plural function or was used as a classifier
- item: whether the 9 items (or pairs of items) in the domain were crayons or animals
- dominant form: whether *dax* or *fep* was dominant form in the 5R/4E condition (*po* was consistently used in the 3/0 condition).

In each condition, the choice of which individual items (animals or crayons) was witnessed in the target

construction was randomly determined for each child, as was the order of presentation of items.

Pretest before each condition. Children were asked to count the 9 distinct entities in order to ensure that they recognized that the relevant domain size was 9. Children were then asked to name each distinct animal or crayon color. After each response, children received a sticker. All children succeeded in both tasks.

Exposure to a potential rule and exceptions. Children were then introduced to a puppet, Mr. Chicken, who, they were told, spoke a different language. Each child took part in both the 5R/4E and 3R/0E conditions as follows:

5R/4E condition: each child witnessed the rule applied once to each of 5 unique cases. 4 other cases were witnessed that were exceptional in that they did not follow the rule.

- When the novel forms were **classifiers,** Mr. Chicken picked up each of the objects and named the entity in "chicken language," saying the name of the entity immediately followed by a novel classifier, 5 of which followed the dominant pattern and 4 of which followed the exceptional pattern, ordered randomly. (*e.g., lion fep, monkey dax, zebra fep, giraffe dax…*). There were no conditioning factors that determined which novel classifier was used with each animal. Children were asked to repeat every novel form witnessed.
- When the novel forms were **plurals,** Mr. Chicken picked up one of each type of object, said its name and then picked up two of the same type, and used a novel suffix as a plural marker (e.g., *lion,* picking up one lion, *lion dax,* picking up two lions). Children repeated each singular and plural form. 5 entities were pluralized with one morpheme (*dax* or *fep*, counterbalanced) and the other 4 were pluralized with the other form. The items assigned to each novel plural marker were selected randomly, so there were no conditioning factors that determined which novel plural was used. Items were selected in random order (e.g., *lion, lion dax; monkey, monkey fep)*

3R/0E condition: each child witnessed the rule applied once to each of 3 unique cases. The other 6 entities were not witnessed either following the rule or being exceptional.

- When the novel form was a **classifier**, Mr. Chicken picked up 3 animals (or crayon colors) and named them with a novel classifier, *po* (e.g., *lion po, zebra po…*). Which animals were named was random for each child. Children were asked to repeat each label after hearing it.
- When the novel form was a **plural**, Mr. Chicken picked up and named one animal or crayon (e.g., *lion*) and then picked up two of the same animals or crayons which were labeled with the name and the plural morpheme, *po* (e.g., *lion po*). This was done for 3 types of animals or crayon colors, selected randomly. Children were asked to repeat each label after hearing it.

Production task

After initial exposure, children were asked to label another item the way Mr. Chicken would. Then, children were exposed again in the same way to the same condition and were asked to label a different item. This provided two responses for each condition. In the 5/4 condition, children were asked to label two never-before-seen items. In the plural condition, children labeled one of the remaining 6 items they hadn't heard labeled. Thus, children provided four critical responses, two in each condition.

Figure 2: Sample stimuli. 3R/0E condition with animals and 5R/4E condition with crayons. The two functions of the constructions, and animals vs. crayons were separately counterbalanced across
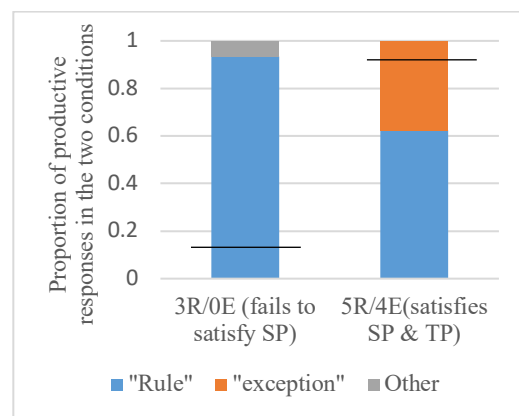
Results

Figure 3: Proportion of responses in 3R/0E and 5R/4E conditions. Dominant form ("rule," blue); less dominant form (orange; relevant in 5R/4E condition), or other (gray). Domain size = 9. Black lines indicate Threshold predictions for height of the rule-following cases (in blue).

The values indicated by blue are the proportion of cases in which the rule was used productively. The Threshold

proposal's predictions for the expected proportion of rule-following cases (in blue) are indicated by the black lines in Figure 2; they are predicted to be identical to those in Figure 1. However, in the 3R/0E condition, the proportion of responses in which children used the novel form productively was near ceiling (M = .935; Figure 2, left). Specifically, 29 out of the 32 children consistently used the novel form productively, 2 children used it on one out of two trials, and only 1 child failed to use it at all.

In the 5R/4E condition, 14 out of 31 children consistently used the dominant form productively for new novel objects, 5 children consistently used the slightly less dominant form productively, and another 12 children produced both of the novel forms (one on each trial) (M = .625; Figure 2, right).

We analyzed the data using the glmer package with Condition (3R/0E or 5R/4E) as the predictor and by-subject random intercepts and slopes, and random intercept for Function-Assignment (plural or classifier first): Correct ~ Condition + (1 + Condition |Subject) + (1|Function-Assignment), family=binomial, data). Recall that other random factors were counterbalanced. The Threshold proposal predicted that the 3R/0E condition should be categorically unproductive and the 5R/4E condition should be categorically productive. However, results show a significant difference between the two conditions in the *opposite* direction ($\beta = 8.824, z = -2.918, p = 0.0035$). This is consistent with the Communicate and Access proposal which predicted that children should be productive in the 3R/0E as long as they understood the task, since there was no interference from a competing form; and children should be markedly less productive in the 5R/4E condition since the 2 forms witnessed would compete with one another, as there were no conditioning factors available to distinguish them.

Prior work has found children over-rely on *either* of two options when the difference in type frequency is not overwhelming (Hudson Kam & Newport 2005; 2009; Schwab, Lew-Williams and Goldberg 2018), and this occurred in the 5R/4E condition. In fact, 19 out of 31 children used only one form or the other: 14 children only used the more dominant form, and another 5 only used the less-dominant form. For this reason, it is not particularly meaningful to compare children's performance in the 5R/4E condition to chance. A majority of children chose one of the two options and simply repeated that option for all cases. But it also not appropriate to describe children's behavior as treating the more dominant form as a rule, given that fewer than half of the children consistently used the dominant form (14/31). Moreover, the remaining 12 children used one of each form, which is a pattern of behavior regularly seen in adults, when two options are witnessed with nearly equal type frequency (Hudson Kam & Newport 2005). In a comparison of the age of the 19 children who used a single form and the 11 children who used both forms, we find on average, that the latter group was 4 months older (M = 54 vs. 59 months). Using a 1-tailed t-test, this result is marginally significant (t = 1.61, p = .059).

Discussion

Critically, children were more productive in the 3R/0E condition than in the 5R/4E condition, directly contradicting the Threshold proposal's predictions, while being consistent with the predictions of Communicate & Access. Moreover, the Threshold proposal makes clear predictions that were disconfirmed in each condition considered separately.

In the 3R/0E condition, the Threshold proposal predicted that children should have been completely *un*productive, as they witnessed fewer rule-following cases than the number demanded by the Sufficiency Principle, given the domain size of 9: i.e., they only witnessed 3 cases, when 5 is predicted to be the minimum number required. Nevertheless, children overwhelmingly used the novel construction productively. The Sufficiency Principle has generally been argued to require an unrealistically high number of rule-following cases be witnessed in order for productivity to be realized (Goldberg 2018, 2019), and children's behavior in the 3R/0E condition confirms this. It is highly unlikely that children misjudged the size of the domain of the construction, since there were exactly nine entities (or pairs of entities) in the display (Figure 2) and children accurately counted them at the beginning of each condition. In fact, if children had assumed that the domain only included the three items that had been witnessed in the novel construction, with the other cases falling outside of the construction's domain, then the construction should not have been applicable to the other cases, and yet children overwhelmingly *did* extend it to the randomly selected new entities at test.

In the 5R/4E condition, the Threshold proposal predicted full productivity of the dominant form (which was witnessed with 5 out of 9 cases), as both the Tolerance and Sufficiency principles were satisfied. While 45% did use the dominant form productively, another 16% used the "exceptional" form productively. The rest, 39% of children, used both forms, one with each of the new entities. Defenders of the Threshold proposal might argue that the last group of children interpreted the input as evidence for *two* distinct and exceptionless rules, one of which applied to 4 cases and the other of which applied to 5 cases. However, this would require distinct domains for the two rules, and yet no conditioning factors were provided. Recall that instances that appeared with the dominant form and instances that appeared with the less dominant form were selected at random and differed across children. And, although the difference in type frequency between the dominant and less dominant constructions was close (5:4), it falls squarely within the thresholds that were proposed for the more dominant construction to become productive as children had done in the 5R/1-1-1-1 case reported by SYN '16.

Is it possible to defend the Threshold proposal on the grounds that the children in the current experiment were more adult-like? That is, the Threshold proposal is specifically aimed at young children's behavior rather than adults', since adults are recognized to behave somewhat differently than children in artificial language paradigms (Boyd & Goldberg 2012; Hudson Kam & Newport 2005, 2009), perhaps relying

on strategies or metalinguistic awareness that is unavailable to children as they learn their first language. Notably, however, the children in the current work were almost 3 years younger than those tested by SYN '16 (56 *vs.* 90 months).

Results in both conditions are consistent with the Communicate and Access proposal. In the 3R/0E condition, only one option was provided and so there was no interference from any competitors. C&A predicts that as long as children are able to appreciate the convention and access the form, they should use the form productively for new cases, as they overwhelmingly did. The results of the 5R/4E condition are also consistent with the C&A proposal. Since there were no conditioning factors to distinguish the two constructions, and since the forms were nearly equal in dominance (type frequency), C&A predicted that children would have no good way to resolve the competition between them. In fact, 14 children consistently used the more dominant option, while 5 used the less dominant form. This over-reliance on a single form recalls prior work that investigated children's productions when faced with unconditioned variation (Kam & Newport, 2005; Singleton & Newport, 2004), or when faced with variation that is conditioned, but by factors that the children fail to recognize (Schwab, Lew-Williams, & Goldberg, 2018). In those studies, children tended to rely on a single option in production tasks, but recognized both forms as acceptable in judgment tasks. The discrepancy between production and judgment tasks suggests that the over-reliance on one form during production results from the challenge of accessing and choosing between multiple forms without any reason to prefer one over the other (Harmon & Kapatsinski 2017; Schwab, Lew-Williams, and Goldberg, 2018).

Recall that the Communicate and Access proposal takes as its starting point the idea that learners aim to convey their messages *while obeying the conventions of their language as best they can*. While it is simpler to over-rely on one option in the face of unconditioned variation, it is more conventional to use both options, since both options were witnessed. As expected, then adults should be more likely to match the relative frequencies in the input even when faced with unconditioned variation between two alternatives, because they are better able to access both forms and choose between them. And in fact adults do tend to be more successful than children at matching the input veridically in mini-artificial language experiments (Kam & Newport 2005; 2009; SYN '16). We see evidence that an over-reliance on a single form is *simplification* in the current work, in that 12 out of 31 children used both novel forms in the 5R/4E condition. Moreover, the children who used both forms in their own productions were marginally older than those who over-relied on a single option, by an average of four months. We take that as an indication that children attempted to successfully produce both options, with older children simply being more successful.

We therefore conclude then that interference—the nature of the exceptional cases—played a key role in whether a competing form was used productively. That is, the difference between the 5R/4E condition here and 5R/1-1-1-1

condition of SYN '16 is that the current class of exceptions all occurred with the same form, making the "exceptional" form itself accessible. And since the exceptional case was just as appropriate for expressing the intended message (i.e., there were no conditioning factors that made either more appropriate), and the "exceptional" cases were nearly as accessible as the "rule," the C&A proposal predicted that the exceptions would interfere with the productive use of the rule. And this is evident in the current results in that children were significantly less productive in the 5/4 condition than in the 3/0 condition.

Conclusion

The present work investigated the factors that underlie children's productive use of a novel rule or construction. We compared two proposals that make contrasting predictions. The first, a Threshold proposal, argues that rules are used productively as long as two thresholds are met: the proportion of potential cases that are witnessed obeying a rule must cross a threshold in order to satisfy a Sufficiency principle and the proportion of potential cases that are witnessed behaving exceptionally must remain below a threshold in order to satisfy a Tolerance principle (Yang, 2016). A Communicate and Access proposal instead appeals to the idea that a speaker's goal is to convey her intended message while obeying the conventions of her language as best she can. On this view, children extend constructions in new ways when they need to express a given message and they are unable to access a more conventional or better match. Accessibility of a construction increases as the variability of witnessed cases increases; and accessibility of the construction decreases as the accessibility of a competing construction increases.

In the current experiments, 4-6-year-old children were exposed to 2 mini-artificial languages. Each language provided exposure to a potentially productive rule, which was assigned a plural or classifier function. In one condition, a novel construction was witnessed applying to 3 out of 9 cases and 0 exceptions. The Threshold proposal predicted that children would *not* use this 3R/0E rule productively, as too few instances were witnessed to satisfy the Sufficiency principle. The Communicate and Access proposal predicted that children *would* use the construction productively because there was no better way to communicate their intended message; i.e., there was no interference from any competing alternative. As predicted by the C&A proposal, the construction was overwhelmingly used productively.

The other condition exposed children to 5 out of 9 cases following a rule, the 4 other cases being exceptions to that rule. The Threshold proposal predicted that in this 5R/4E condition, children should be fully productive, since a sufficient number of rule-following cases was witnessed, and a low enough number of exceptions was witnessed. Unlike in previous work (SYN '16), here the 4 exceptional cases all behaved alike. The Communicate and Access proposal predicted that there would be competition between the two constructions, and that this would interfere with the productivity of both. In fact, there was markedly less

productivity in the 5R/4E condition than in the 3R/0E condition, counter to what the Threshold proposal predicted and consistent with the C&A proposal. Children in the 5R/4E condition over-relied on the dominant construction (45%), *or* on the less dominant construction (16%), *or* they used both constructions (39%).

To summarize, our preregistered experiment contradicts the Threshold proposal while being consistent with Communicate and Access. We conclude that productivity is encouraged by the desire to communicate a message while obeying the conventions of the language. On this perspective, we do not extend a construction productively unless we are unable to access a "better" (more conventional and appropriate) way to express our intended message. Productivity of a construction is contrained by the accessibility of the construction, and accessibility is affected by both the variability of witnessed exemplars and interference from a competing construction (Goldberg, 2019). When there is no better way or when we are unable to access a better way at the moment of speaking, we have no choice but to extend appropriate constructions that *can* be accessed.

The C&A proposal takes a different perspective on prior findings that children tend to "regularize" their input, making it more systematic and therefore in some sense better. The C&A proposal suggests that "regularization" arises from a failure to successfully access a more conventional and appropriate alternative. C&A takes the position that both children and adults aim to conform to the conventions used by others who are considered to be knowledgeable. Adults are more successful at reflecting the input veridically given very limited exposure, but children aim to--and ultimately do--learn the conditioning factors of the constructions they are exposed to, and to a remarkable extent, successfully conform to the speech patterns used in their language communities. In fact, we saw adult-like behavior in a subset of (somewhat older) children in the current experiment who, in the 5R/4E condition, used both options.

The Threshold proposal faces other outstanding issues that are not addressed here. For example, exceptions are assumed to be searched serially and before rule-following cases, despite a lack of psycholinguistic evidence for this claim (Hernandez, 2019; Wittenberg & Jackendoff 2018; Kapatsinki 2018). The proposal assumes that exceptions are listed in order of frequency so that neither exceptions nor rule-following cases are allowed to cluster within our associative memory as proposed by the C&A and other accounts (Ambridge et al. 2018; Bybee 2010; Goldberg 2019; Kapatsinki 2018; McClelland& Patterson 2002). Without allowing instances to cluster in memory, it is entirely unclear how children are able to determine the domain of a rule, let alone calculate the size of the domain, as is required for the Threshold proposal to make any predictions at all.

To summarize, constructions (or "rules") do not "go productive" by crossing predetermined numerical thresholds. Rather, people extend constructions for new uses when doing so provides an accessible way to best express their intended messages.

References

Ambridge, B., Barak, L., Wonnacott, E., Bannard, C., & Sala, G. (2018). Effects of both preemption and entrenchment in the retreat from verb overgeneralization errors. *Collabra: Psychology*, *4*(1).

Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. In *Mechanisms of Language Acquisition* (pp. 157–193). L. Erlbaum Ass.

Berko, J. (1958). The child's learning of English morphology. *Word*, *14*(2–3), 150–177.

Boyd, J. K., & Goldberg, A. E. (2012). Young children fail to fully generalize a novel argument structure construction when exposed to the same input as older learners. *Journal of Child Language*, *39*(3), 457-481.

Bybee, J. (2010). *Language, usage and cognition*. Cambridge University Press.

Goldberg, A. E. (2018). The sufficiency principle hyperinflates the price of productivity. *Ling. Approaches to Bilingualism*, *8*(6), 727-732.

Goldberg, A.E. (2019) *Explain me this: creativity, competition, and the partial productivity of constructions*. Princeton University Press.

Freudenthal, D., Pine, J. M., & Gobet, F. A 2018. Computational Model of the Acquisition of German Case. *Proc. of Cog. Sci. Conference*. Madison, Wisc.

Harmon, Z., & Kapatsinski, V. (2017). Putting old tools to novel uses . *Cog. Psych. 98*, 22–44.

Hernandez, Alexia. (2019). Rule Generalization in Children: Testing a Threshold Proposal. Ling. Senior Thesis,Princeton.

Kam, CH & Newport, EL (2005). Regularizing unpredictable variation. *Lang. Learn.& Dev. 2*(2), 151–195.

Kam, C. L. H., & Newport, E. L. (2009). Getting it right by getting it wrong. *Cognitive Psychology*, *59*(1), 30-66.

Kapatsinski, V. (2018). On the intolerance of the Tolerance Principle. *Ling. Approaches to Bilingualism*, *8*(6), 738-742.

Kapatsinski, V. (2018). *Changing minds changing tools*. MIT Press.

MacDonald, M. C. (2013). How language production shapes language form & comprehension. *Frontiers in Psych. 4*: 1–16.

McClelland, J. L., & Patterson, K. (2002). Rules or connections in past-tense inflections. *TiCS 6*(11), 465-472.

Montag, JL, Matsuki, K., Kim, JY, & MacDonald, MC (2017). Language Specific and Language General Motivations of Production Choices *Collabra: 3*: 1–22.

Schuler, K. D., Yang, C., & Newport, E. L. (2016). Testing the Tolerance Principle . *Proc. of Cog Sci.*

Schwab, JF, Lew-Williams C, & Goldberg, AE (2018). When regularization gets it wrong. *JCL* 1–19.

Singleton, JL & Newport, EL(2004). When learners surpass their models. *Cog. Psych.49*(4), 370–407.

Suttle, L., & Goldberg, A. E. (2011). The partial productivity of constructions as induction. *Linguistics*, *49*(6), 1237-1269.

Tenenbaum, JB & Griffiths, TL (2001). Generalization, similarity, and Bayesian inference, *BBS:* 629–640.

Wittenberg, E. & Jackendoff, R. (2018). Formalist modeling and psychological reality. *Ling. Approaches to Bilingualism*, *8*(6), 787-791.

Yang, C. (2016). *Price of Linguistic Productivity.* MIT Press