Investigating a Productive Rule for English Singular Noun Possessives

Henry Garant
University of Pennsylvania
May 14th, 2019
Ling 270

Abstract

Jackson & Cottrell (1997) demonstrated the characteristic child language "U-shaped" learning patterns for the English past tense; however, they did not question the underlying reasons for the generation of "U-shaped" learning. This paper attempts to offer the process of forming generalized language rules as an explanation of the pattern of "U-shaped" learning. Using a calculus, the Tolerance Principle given by Yang (2005), this paper demonstrates the possible relationship between rule learning and the generation of the "U-shaped" pattern. To exhibit this relationship, this paper focuses on children learning the English rule of singular possessives: adding ('s) to indicate possession. In particular, child utterance data was taken from the CHILDES database to investigate how children learn this rule despite common exceptions. Ultimately, this paper provides evidence that the Tolerance Principle can be used to predict not only when a child learns a productive rule but also when a child exhibits "U-shaped" learning.

Introduction

Even the most experienced writers may pause to decide whether they need to use “its” or “it’s.” While both sound identical, “its” and “it’s” have different meanings in English. The contraption, “it’s” represents a combining of two words: “it” and “is;” however, “its” indicates possession. This is a peculiarity in English that leads to a larger question of how children begin to understand and produce singular possessives using an (‘s) in utterances such as “Daddy’s car.” Specifically, how do children form a rule to use (‘s) for producing singular noun possessives when contraptions such as “let’s and it’s” don’t indicate any form of possession?

To be clear, singular possessives are different from plural possessives(i.e. Bess’ truck) because they require the addition of an “s.” It is widely known that children form utterances of possessive intention by using structure before adding (‘s). For example, children will say “give mommy hand” to mean “give me mommy’s hand.” Nevertheless, children begin to systemically adopt (‘s) to indicate possession. Using this intuition, child utterances can be examined to analyze the underlying process of learning the rule of using (‘s) for singular possessives in English.

First, similar to Xu & Pinker (1995), a mean error rate of possessives can be developed to demonstrate a child’s learning process. In particular, this error rate could demonstrate the overarching pattern of “U-shaped” learning that permeates many rules in language acquisition such as past tense.(Jackson & Cottrell, 1997) This “U-shaped” pattern, high performance going down to low performance only to come back to high performance, is thought to occur due to item-specific memory giving way to a general rule of language production. At first, children may strictly memorize instances on an item-by-item basis, which produces high performance. Then, after enough instances are collected, children able to form a generalized rule, which leads to low performance. After a period of curtailing the generalization, the rule begins to lead to higher performance. The question is whether a rule for using (‘s) to produce singular possessives in English will demonstrate “U-shaped” learning.

Second, if patterns of “U-shaped” learning emerge from the process of creating generalized rules, will the Tolerance Principle outlined by Yang(2005) be able to offer any prediction on when the “U-shape” will appear in a child’s learning? By treating instances of contraptions such as “let’s” and “it’s” as exceptions to the rule for using (‘s) for singular possessives in English, the Tolerance Principle will be able to show when children have enough correct instances to form such a productive rule.

Method

Subjects

In total, three English-NA subjects were selected from the CHILDES database: **Eve**[1;6 - 2;3], **Peter**[1;9 - 3;1], and **Naomi**[1;2 - 4;9]. Eve's data was pulled from the Brown Corpus.(Brown 1973) Peter's data was pulled from the Bloom70 Corpus.(Bloom 1975) Naomi's data was pulled from the Sachs Corpus.(Sachs 1983) In general, these subjects were selected not only for corpus variation and focus on naturalistic conversation but also for age range. For age range, it was critical to find rich longitudinal data that had early speech production as well as more developed speech production to capture the process of learning a productive rule for singular possessives.

In terms of background, Peter was an upper-middle-class white male with college-educated parents. Naomi was an upper-middle-class white female that was the daughter of Professor Sachs. Eve was an upper-middle-class white female that was particularly linguistically precocious. Although not a factor in selection criteria, it should be noted that all children were raised in an upper-middle-class setting.

Procedure

A python script[Appendix A] was written to analyze the children's transcript data. Every line that contained a child utterance was parsed to collect three components using Stanford's NLTK Part-of-Speech Tagger and regular expressions.[Appendix C-E] All unique cases of **(1)**, **(2)**, and **(3)** were stored for calculations.

(1) A correct instance of a singular possessive(i.e. Daddy's car)
(2) A missing instance of a singular possessive(i.e. Daddy car)
(3) A correct instance of a contraction usage with 's(i.e. It's Daddy)

For every file in a child's dataset, ONLY that file's count of **(1)** and **(2)** was used to calculate a file "error rate" of possessive usage according to the following formula:

$$Error\ rate\ =\ \frac{(2)}{(1)+(2)}$$

Each file's (1 - error rate) was then graphed to show learning of a singular possessive rule across the entirety of the child's data.

For every file in a child's dataset, TOTAL counts of **(1)** and **(3)** were used in calculating the Tolerance Principle to determine if a child had a productive rule for singular possessives. Assume that E is the number of exceptions and N is the number of instances of a rule. That is, E := **(3)** because using an ('s) as a contraption is

an exception to a rule that (‘s) denotes singular possession. N := **(1)** + **(3)** to count all instances in which (‘s) was used in the child’s speech production. Then we arrive at determining if a child has a productive rule if and only if:

$$E \leq \frac{N}{ln\, N} \equiv (3) \leq \frac{(1)+(3)}{ln[(1)+(3)]}$$

Then age periods in which the child had a productive rule were overlaid in a shaded blue region on the child’s graphs. This allows easy discernment of when the child would have a productive rule for using (‘s) for singular possession.

Finally, the count of **(1)** was graphed after parsing each file to visualize the number of unique singular possessive utterances the child produced over the dataset. Again, periods of a productive rule were overlaid in blue to visualize the relationship between a singular possessive productive rule and generating unique singular possessive utterances.

Results

Eve, Peter, and Naomi managed to reduce their error rate to below 15% by the age of 2;0.[Figures 1,3,5] This shows that, although the children did not produce many singular possessives using an (‘s) before 1;0, they ultimately learned a rule for using (‘s) to indicate singular proper noun possession. Furthermore, the proposition that the children learned a rule is validated because after 1;1 each child’s rate of unique singular possessives increased significantly.[Figures 2,4,6]

Ultimately each child went through a few periods in which they could have formed a productive rule for singular possessives according to the Tolerance Principle. In general, the early periods when the child had a productive rule were based on very few examples and exceptions. This shows how the Tolerance Principle can be effective despite a small number of instances. In addition, the earliest periods in which the child could have formed a productive rule corresponded to an increase in the “error rate”. The increase in error rate, due to an increase in the number of instances in which the child omitted an (‘s) for possession, demonstrates that the child may have been under generalizing the formed rule. Furthermore, this pattern demonstrates U-shaped learning that may be a result of forming such an under generalized rule after using item-specific memory.

Nevertheless, each child’s last productive period corresponded to not only a drop in error rate but also a rapid increase in unique singular possessive productions. This further reinforces that the children learned a productive rule for singular possessives using (‘s) and that the timing of when this rule becomes productive can be estimated using the Tolerance Principle.

Figure 1 - Eve - Graph of (1 - error rate) across data - ■ - Productive Rule

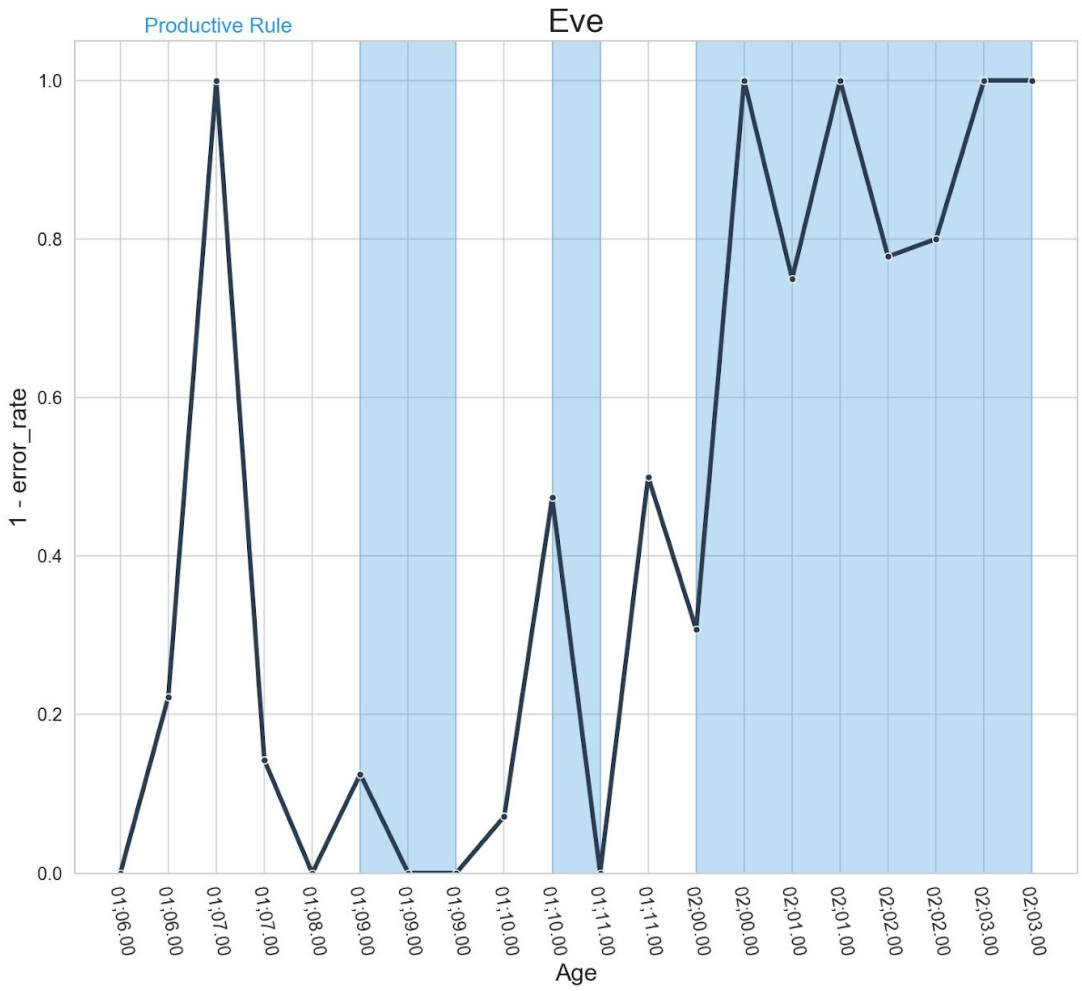


Figure 2 - Eve - Graph of unique singular possessive production across data

■ - Productive Rule

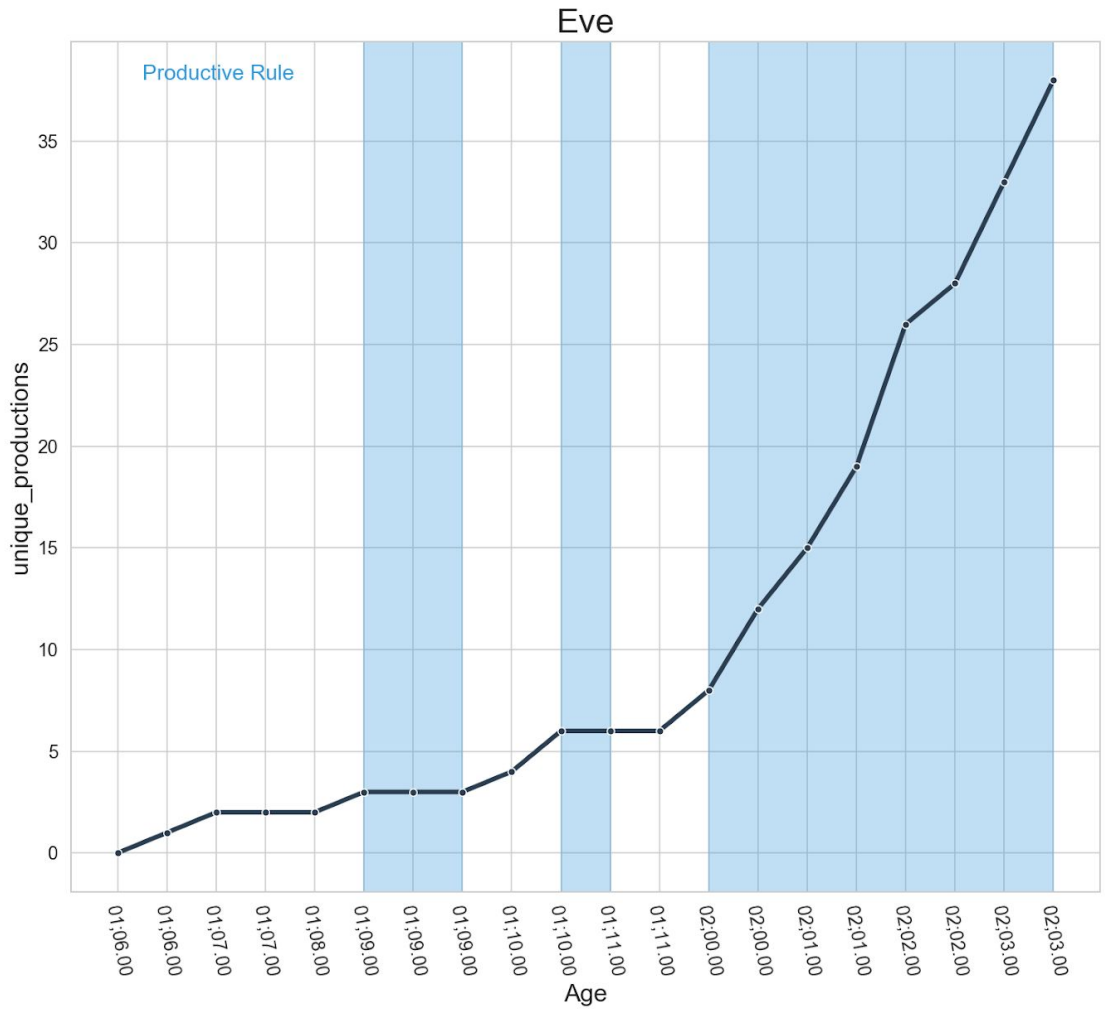


Figure 3 - Peter - Graph of (1 - error rate) across data - ■ - Productive Rule

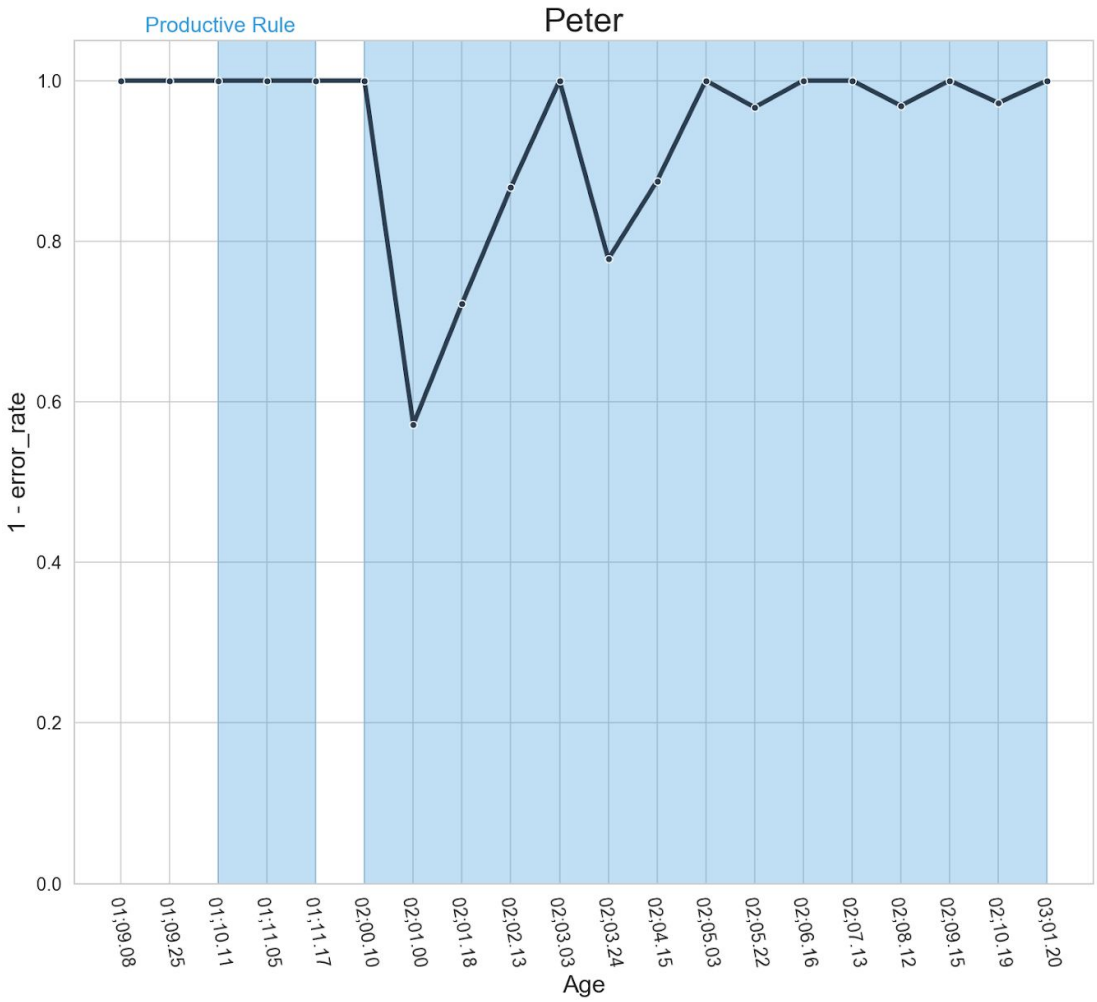


Figure 4 - Peter - Graph of unique singular possessive production across data

■ - Productive Rule

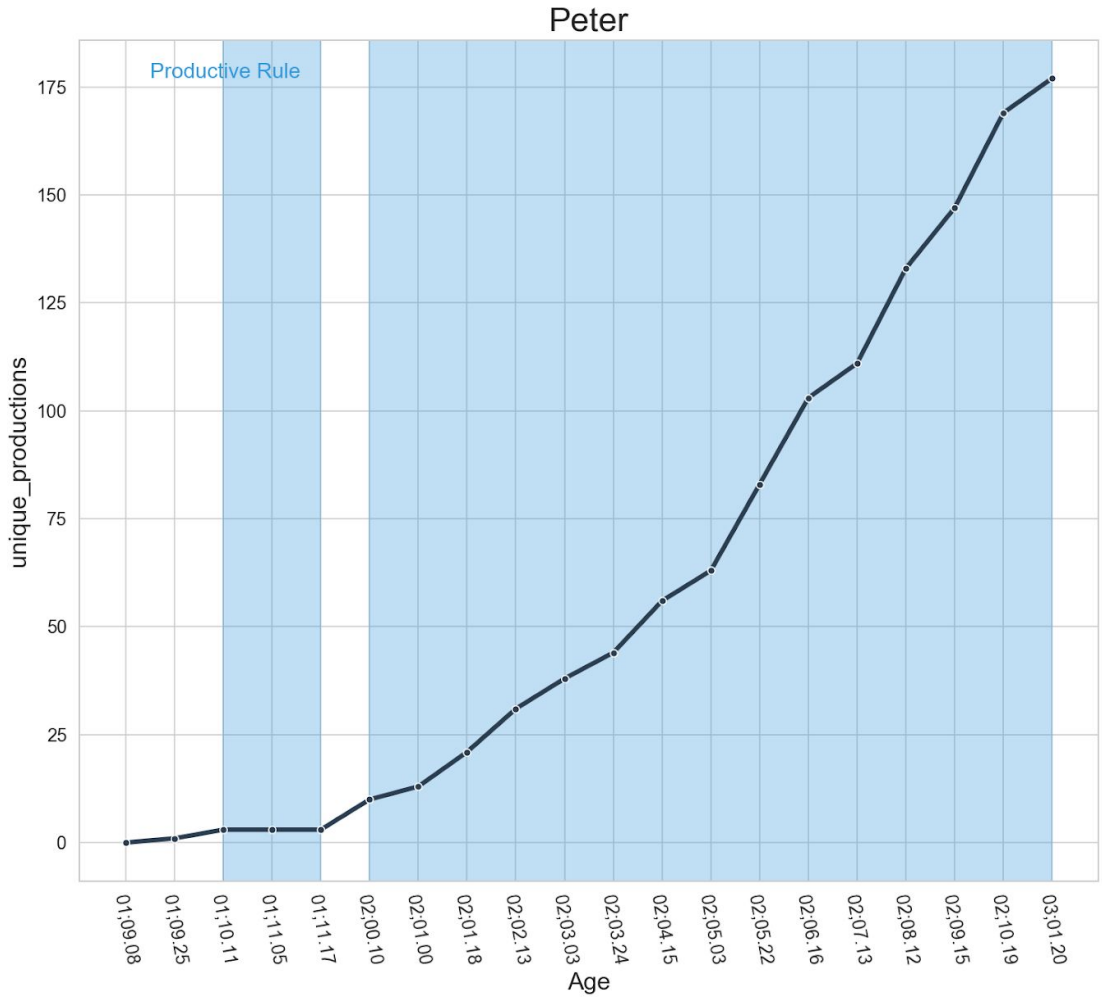


Figure 5 - Naomi - Graph of (1 - error rate) across data - ■ - Productive Rule

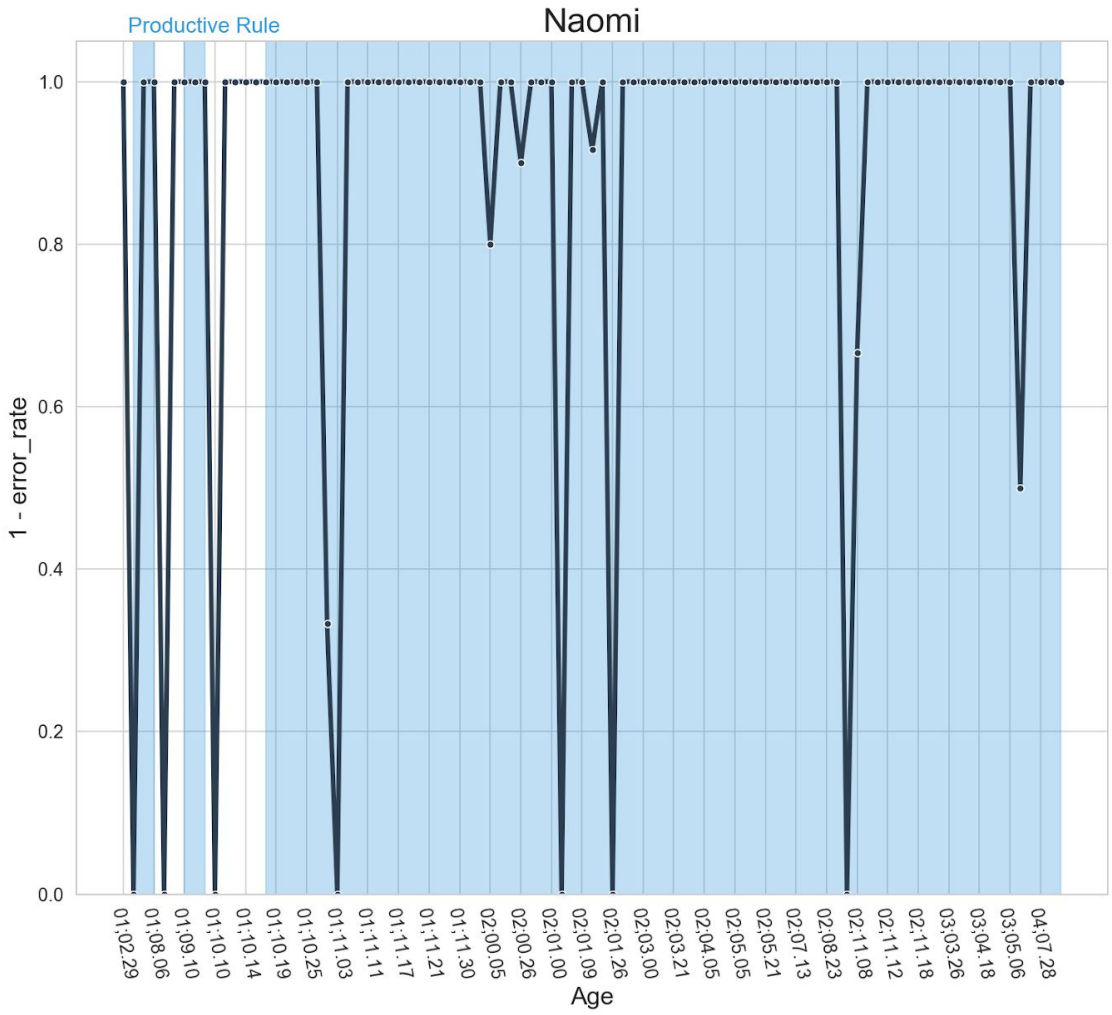
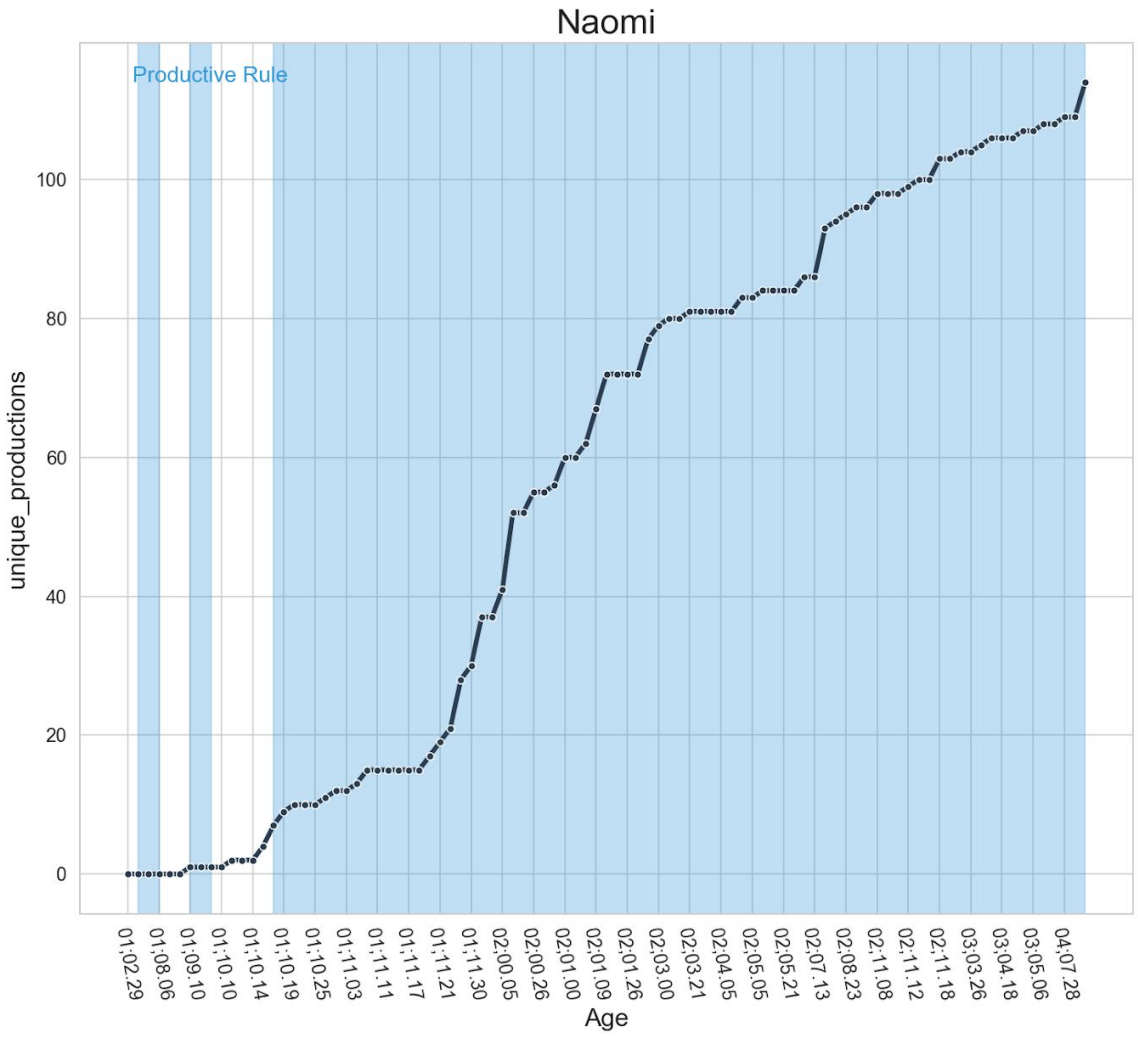


Figure 6 - Naomi - Graph of unique singular possessive production across data

■ - Productive Rule



Discussion

"U-shaped" Learning

Both Eve's and Peter's graphs of error rate produced patterns of learning that strongly resembled the characteristic "U-shape."[Figures 1,3] For Eve, this pattern began around 1;07 and Peter's began around 1;11. A possible explanation for the apparent "U-shape" is that Eve and Peter used item-specific memory to store instances of ('s) possessives at first. This would mean that (1 - error rate) started high because Eve and Peter were able to replicate memorized instances without flaw. However, when they formed a general rule, (1 - error rate) dropped because the children omitted ('s) for possessive instances. In the end, (1 - error rate) rose high again perhaps due to their rule becoming more productive and less generalized.

In contrast, an overall pattern of "U-shaped" learning is hard to find in Naomi's graph of error rate.[Figure 5] Although brief periods exhibit a "U-shape," the pattern is neither predominant nor discernable. Yet, notice how Naomi's error rate bounces between 1 and 0. This hints to a possibility that because Naomi's utterance files were one-third shorter than Eve's or Peter's that each file was less likely to contain correct instances or lacking instances. In order to fix this problem, Naomi's utterance files could be chunked into age periods to make the error rate more salient.

Tolerance Principle for Productivity

For all children, the Tolerance Principle produced periods in which a productive rule was formed. Eve and Naomi had three productive rule periods while Peter had two.[Figures 1,3,5] Most notably, the last productive rule period for each child corresponded with a rapid increase in the number of unique productions.[Figures 2,4,6] Combined with the fact that, during this last period, each child's error rate decreased, the Tolerance Principle provides significant predictability of rule productiveness for singular possessives in English using ('s).

Excluding the last productive rule period, the earlier productive rule periods for Eve and Naomi corresponded with a drop in error rate.[Figures 1,5] This further supports the underlying pattern of "U-shaped" learning may be due to the child forming rules as he/she learns. Peter's first productive rule period neither corresponded with a drop in error rate nor an increase in unique productions.[Figures 2,3] This may be due to small errors in the method of counting N and E. It is also possible that just because the Tolerance Principle predicts a rule becoming productive for a short period does not require that the child uses it productively in that short period.

Conclusion

In the end, this paper's goal was to investigate "U-shaped" learning and the Tolerance Principle through analyzing a specific rule in English, the rule for singular possessives. From analyzing the corpus results of Eve, Peter, and Naomi, the paper provides evidence to suggest not only that children's learning of this rule follows the pattern of "U-shaped" learning but also that the Tolerance Principle predicts the "U-shaped" pattern.

For this paper, two improvements can be made. As suggested before, to deal with smaller file sizes of Naomi's data, her utterances could have been combined to produce more meaningful error rates that captured her learning process. Second, in order to further confirm the item-specific component hypothesized to be behind the beginning of "U-shaped" learning, the parent's correct instances of singular possessives could have been collected. Then the child's first correct instances of singular possessives could be compared to the parent's correct instances of singular possessives to show if the child was memorizing item-specific instances of what the parents said.

To conclude, this paper also intended to highlight key areas that require more work from the linguistic community. Specifically, although extremely difficult to conduct, longitudinal studies of children's speech are critical for investigating long term development of rules. A focus to produce data in this manner should become more of a target for researchers in order to enhance scholarship on the topic of language rule learning. Also, many rules similar to past tense and singular possessive remain to be studied not only in English but also in other languages. Work needs to be conducted to analyze "U-shaped" learning and the Tolerance Principle for these unexplored rules.

References

1. Bloom, L., Lightbown, P., & Hood, L. (1975). *Structure and variation in child language.* Monographs of the Society for Research in Child Development, 40, (Serial No. 160).

2. Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

3. Jackson, D., & Cottrell, G. W. (1997). *Attention and U-shaped learning in the acquisition of the past tense.* In Proceedings of the Nineteenth Cognitive Science Conference (pp. 325-30).

4. Sachs, J. (1983). *Talking about the there and then: The emergence of displaced reference in parent–child discourse.* In K. E. Nelson (Ed.), Children's language, Vol. 4, Hillsdale, NJ: Lawrence Erlbaum Associates.

5. Xu, F., & Pinker, S. (1995). *Weird past tense forms.* Journal of child language, 22(3), 531-556.

6. Yang, C. (2005). *On productivity.* Linguistic Variation Yearbook., 5(1), 265–302. https://doi.org/10.1075/livy.5.09yan

Appendix

A. Link to Python Script:

   https://github.com/Henri93/Possessive_Rule/blob/master/possesive.py

B. Link to Figures:

   https://github.com/Henri93/Possessive_Rule/tree/master/Figures

C. Regular expression for (1) A correct instance of a singular possessive(i.e. Daddy's car) "`dn-POSS\s((adj)*(n)*)+`"

   ```
   Ex. Possessive Noun-space-adjective-OR-noun

   In practice:

   *CHI:  Granny's house .

   %mor:  adj|Granny&dn-POSS n|house .

   Parsed: Granny's house
   ```

D. Regular expression for (2) A missing instance of a singular possessive(i.e. Daddy car) "`n:prop\|(\w*)\sn\|(\+n\|)?(\w*)(\+n\|)?(\w*)`"

   ```
   Ex. Proper Noun|-any word-space-compound_noun-OR-noun

   In practice:

   *CHI:  Mommy telephone .

   %mor:  n:prop|Mommy n|telephone.

   Parsed: n:prop|Mommy n|telephone
   ```

E. Regular expression for (3) A correct instance of a contraction usage with 's(i.e. It's Daddy) "`\w*'s`"

   ```
   Ex. word-'-s

   In practice:

   *CHI:  here's your rattle .

   %mor:  pro:exist|here~cop|be&3S det:poss|your n|rattle .

   Parsed: here's
   ```