Parents' Response Times Provide Implicit Negative Evidence for Grammar Learning

Emilia Russo[1] & Daniel Yurovsky[1]

[1] University of Chicago

Author Note

This work was supported by a James S McDonnell Scholars Foundation grant to DY

Correspondence concerning this article should be addressed to Daniel Yurovsky, 5848 S University Ave, Chicago, IL 60637. E-mail: yurovsky@uchicago.edu

Abstract

A key debate in language development is how children learn an infinitely generative language from a finite amount of evidence. Although children can reasonably take the production of an utterance from an adult as evidence for its grammaticality, this positive evidence may not be sufficient to constrain the learning of an infinitely generative grammar with complex rules and numerous, subtle exceptions. The problem would be easier if children consistently received negative evidence after producing incorrect utterances. However, while parents sometimes correct children's semantic errors, they rarely correct syntactic errors. Parents' reformulations of children's utterances (e.g. "I runned yesterday" with "Yes, you ran yesterday") could be useful for learning correct grammar, but knowing when a response is a reformulation is non-trivial without knowledge that allows the two forms to be aligned. We hypothesized that children may rely on a lower-level signal in conjunction with or even instead of reformulations: response time. We analyzed response times from three dense corpora to examine how parent response times vary with the grammaticality of the child utterance. This analysis revealed that parents were significantly slower to respond to ungrammatical utterances than grammatical utterances. These results indicate that response time may be one implicit learning cue for language. Additionally, we employed a self-paced reading experiment and found that adults are significantly slower to process overregularized utterances compared to their grammatical counterparts, indicating that parents may take longer to respond due to a processing delay.

Keywords: language acquisition, learning, cognitive development

Word count: X

Parents' Response Times Provide Implicit Negative Evidence for Grammar Learning

MAYBE START WITH SOME KIND OF OVERREGULARIZATION EXAMPLE?

Why is it hard to learn language? One fundamental property of language is that it is infite–there are infinitely many true, grammatically correct utterances of any given natural language. However, because a child's experience with language is finite, language learning is necessarily an inference problem. Just because a child has not heard a particular word or syntactic construction, for instance, is no guarantee that is not a part of the language they are learning. But any finite set of data is consistent with infinitely many hypotheses, only one of which the "right" language (Gold, 1967). What is a child to do?

One possibility is that even though many hypothetical language are consistent with the child's input, there are reasons to prefer some to others. Children could bring a variety of inductive biases–whether learned or innate–to bear on their inferences about language, the way people use language, or both. For instance, children might preferentially favor languages that are simpler, or in which relationships between sounds and meanings–or meanings and other meanings–are systematic (Clark, 1987, Markman (1990), Smith, Jones, Landau, Gershkoff-Stowe, and Samuelson (2002), Tenenbaum (1999)). Or they might assume that speakers are cooperative, and that they choose their words efficiently, not wasting effort. (Grice, 1969, Zipf (1949), Frank and Goodman (2014)). If language learners' inductive biases are well matched to the language they are learning, and the environment they are learning in, these biases can allow them to correctly the right generating language from among the infinite set of consistent languages (Gold, 1967).

Children could have a second source of infomation at their disposal: negative evidence. In addition to the positive evidence of grammatical utterances they have heard, children could potential get evidence from their communicative partners about the grammaticality of their own utterances. When children make semantic errors, i.e. calling a horse a "dog," parents often respond with correction–"that's not a dog, that's a horse" However, these kinds of explicitly corrective signals are rare for syntactic errors (Newport, Gleitman, &

Gleitman, 1977). When a child produces an overregularized paste tense (like “runned” above), they are unlikely to get explicit corrective feedback and appear to be resistant to the rare explicit corrections they receive (Brown & Hanlon, 1970). Children could, however, benefit from less explicit forms of feedback. For instance, in response to an overregularized past ense, parents will sometimes reformulate their child’s utterance (i.e. a child might produce “I catched the ball” and her father might respond, “Yeah, you caught the ball yesterday at tee-ball practice.”; Hirsh-Pasek, Treiman, and Schneiderman (1984), Chouinard and Clark (2003)] For example, However, such reformulations are infrequent, and do not trivially indicate an error (Marcus, 1993, Morgan and Travis (1989)).

Prior work has found that adults are slower to process an unpredicted utterance (Jurafsky, 1999; Fine & Jaeger, 2013). Given this work, we investigate the possibility that parents may take longer to process their child’s unexpected (i.e., ungrammatical) utterances and thus respond later to their child’s utterance. If parents demonstrate a delay in response to a child’s overregularized utterance, the delay may provide an implicit cue to the child that they produced an incorrect utterance. We propose and investigate a novel form of negative evidence for grammar learning - parent response time.

<!– Sometimes, however, the rules that infants generate from positive evidence are incorrect or insufficient. It would certainly be helpful if these infants could get feedback when their rules fail, i.e., when they produce ungrammatical utterances. This negative evidence about what is incorrect to produce could greatly simplify the grammar learning problem by providing correction. If children also receive feedback when they produce something ungrammatical, they would learn grammar through supervised learning, i.e., learning in which the learner receives both positive and negative evidence. Some aspects of language learning, namely category labelling, occur in a supervised context. # Corpus Analyses ## Method ### Child data Analyses were performed on the Eleanor, Fraser, and Thomas corpora publicly available from the CHILDES database (MacWhinney, 2000). The corpora were chosen due to capturing the age of interest, high recording frequency, and the

availability of timing data. See Appendix A for resources regarding recording and transcription procedures for each corpus. To be included in analysis, the data had to meet the following requirements: the speaker must be the child, the responder must be a parent, and there must be valid time codes for both the utterance and response. Since overregularizations, by definition, may only occur in the context of past-tense or plural utterances, overregularized utterances were compared to error-free utterances that contained a past-tense or plural.

Thomas. The Thomas corpus consists of 379 60-minute recordings and transcriptions of natural speech interaction primarily between Thomas and his mother (**???**). This corpus was chosen due to its density, age range of 2 years; 0 months through 4 years; 11 months that captured several initial and overregularized forms, as well as the presence of timing data. Additionally, Thomas was used in a prior analysis concerning overregularization which indicated the presence of a number of well-recorded overregularizations and served as a check for our methods of retrieving Thomas's utterances and information about them (see Maslen, Theakston, Lieven, & Tomasello, 2004). Analysis concerned all child-produced past-tense and plural utterances with a direct response from the mother (n=1,540). We limited analysis to utterances to which the mother responded as the proportion of utterances to which the father responded was negligible (n=321; .23% of parent responses).

Eleanor. The Eleanor corpus consists of 194 60-minute transcriptions between Eleanor and her parents (Lieven et al., 2009). Although also recorded at a high frequency, the Eleanor and Fraser corpora cover a smaller age range from 2 years, 0 months through 3 years, 0 months. Analysis was limited to all child-produced past-tense and plural utterances with a direct response from either parent (n=550). While most of her exchanges were with her mother (n=402), there was a substantial number (n=148; 26.91%) of responses from Eleanor's father as well. As the responses of Eleanor's father did not vary significantly from the responses of the mother, we considered the responses of both parents in our analysis.

Fraser. The Fraser corpus features 216 60-minute transcriptions between Fraser and his parents (Lieven et al., 2009). As noted above, the Fraser corpus covers a smaller age range from 2 years, 0 months, through 3 years, 0 month. Analysis was performed on the set of child-produced past tense or plural utterances for which there was a direct response from the mother (n=870). Although there were responses from Fraser's father and mother, only his mother's responses were analyzed because the proportion of father responses was much smaller (n=50; 5.75%) than that constituted by the mother and because the responses by the father varied significantly from those of the mother.

Pre-Processing. Data for all children were extracted from the XML versions of their transcripts in the CHILDES database (MacWhinney, 2000). Utterances were collected from the corpus by using a Python script to parse the XML transcriptions. An utterance was defined as an entry in the transcript and a response to an utterance Ui was defined as the utterance, Ui+1, immediately following Ui. For each file, the script recorded information about each utterance entry which included the time data, utterance content, response content, speaker, responder, error presence, and the presence of a past-tense or plural in the utterance. This data was then further filtered and analyzed using R. See Appendix A for references regarding source code and data.

Identification of overregularized utterances. Potential overregularizations were identified as utterances where an error occurred and the child used past tense, past participle, or plural in the utterance. Included in this were errors that occurred in the utterance but were not due to an overregularization of the past tense or plural. For example, the utterance "it did go elephants sleep" is an ungrammatical utterance and includes a plural ("elephants") but does not feature an incorrectly regularized form. As such, the data that fulfilled the above criteria were selected and then filtered as true overregularizations through hand coding. Utterances that featured an error but were not overregularizations were tagged as "other error." As such, all overregularizations and all other errors were hand identified.

Results

Parents took significantly longer to respond to ungrammatical utterances compared to error-free utterances across all three children. When comparing overregularized utterances and error-free utterances that feature a past tense or plural, parents were significantly more likely to overlap their child's speech when there was no error present in the child's utterance and least likely to overlap their child's speech when there was an overregularization error in their child's utterance.

Response Time Distributions. Generally, the distribution of response times varied qualitatively in the same way across corpora. Parent response time to grammatical utterances tended to be more negative and much more variable than response time to overregularized and non-regularized ungrammatical utterances. Additionally, response times to overregularized utterances in the Thomas and Eleanor corpora tended to more positive and less variable than response times to ungrammatical utterances that did not contain overregularizations.

This qualitative difference was reflected in the mean response time and variance calculations for error types. For all three corpora, there was an extremely significant ($p < .001$) positive difference between response times to overregularized utterances and response times to grammatical utterances. Reports detailing child-specific results can be found in Appendix B.

Thomas and Fraser showed very similar ratios of overregularization production compared to grammatical utterances (Thomas = 23.4%, Fraser = 22.4%) and a similar difference between mean response times of overregularized utterances and grammatical utterances (Thomas = 264ms, Fraser = 292ms). Response times to Eleanor's utterances generally tended to be much more extreme (Mno error = -278ms, Moverregularization = 312ms) resulting in a much larger difference between response times to grammatical and overregularized utterances (547ms). Additionally, Eleanor's ratio of overregularizations to grammatical utterances (4.9%) was much smaller than those of Thomas and Fraser.

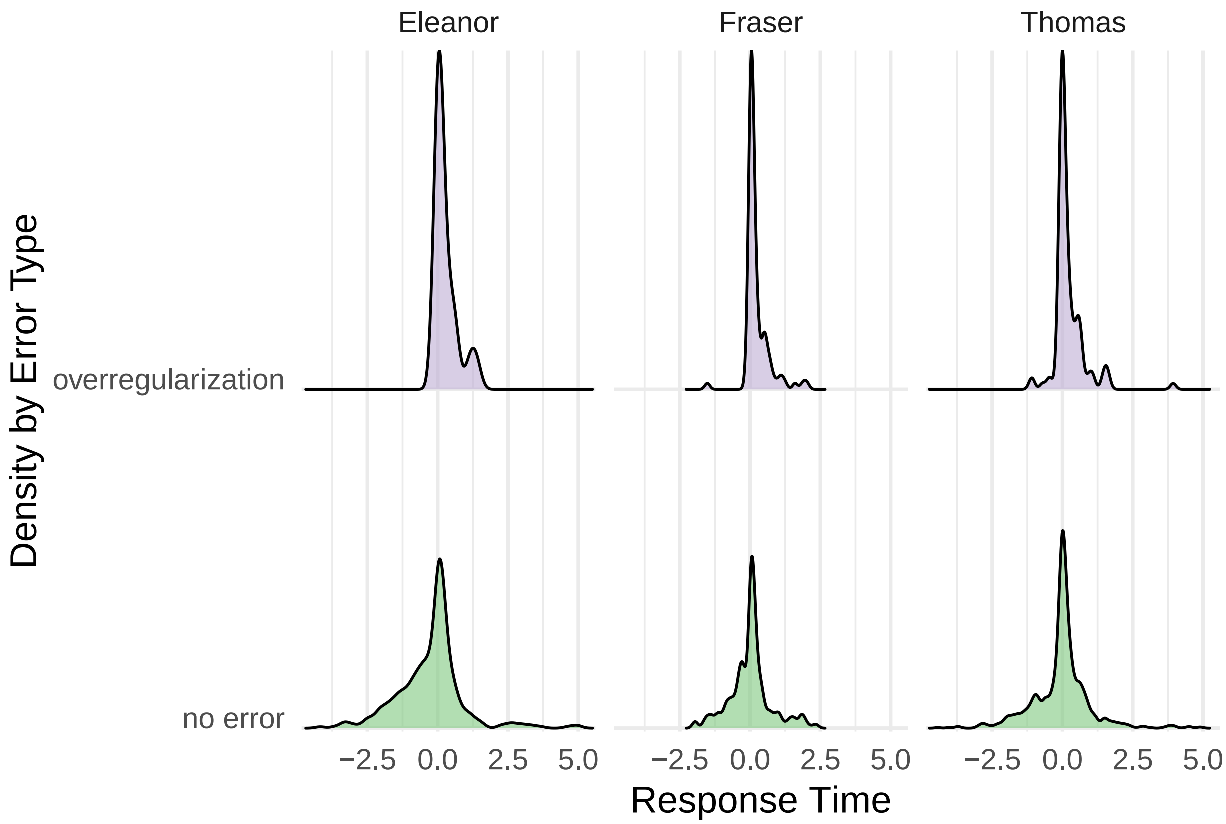

Figure 1

Predicting Response Time by Error Type. We fit a mixed-effects linear model to the timing data predicting response times from fixed effects of error type, age, and response length and a random effect of error_type nested within child. In addition, because the scales of response times varied across children (see above), response time was centered and scaled to be on the same normal scale so we could compare across children: response_time ~ error_type + age + response_length + (error_type | kid)

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: response_time ~ error_type + age + rl + (error_type | kid)
##    Data: .
##
## REML criterion at convergence: 7963.1
```

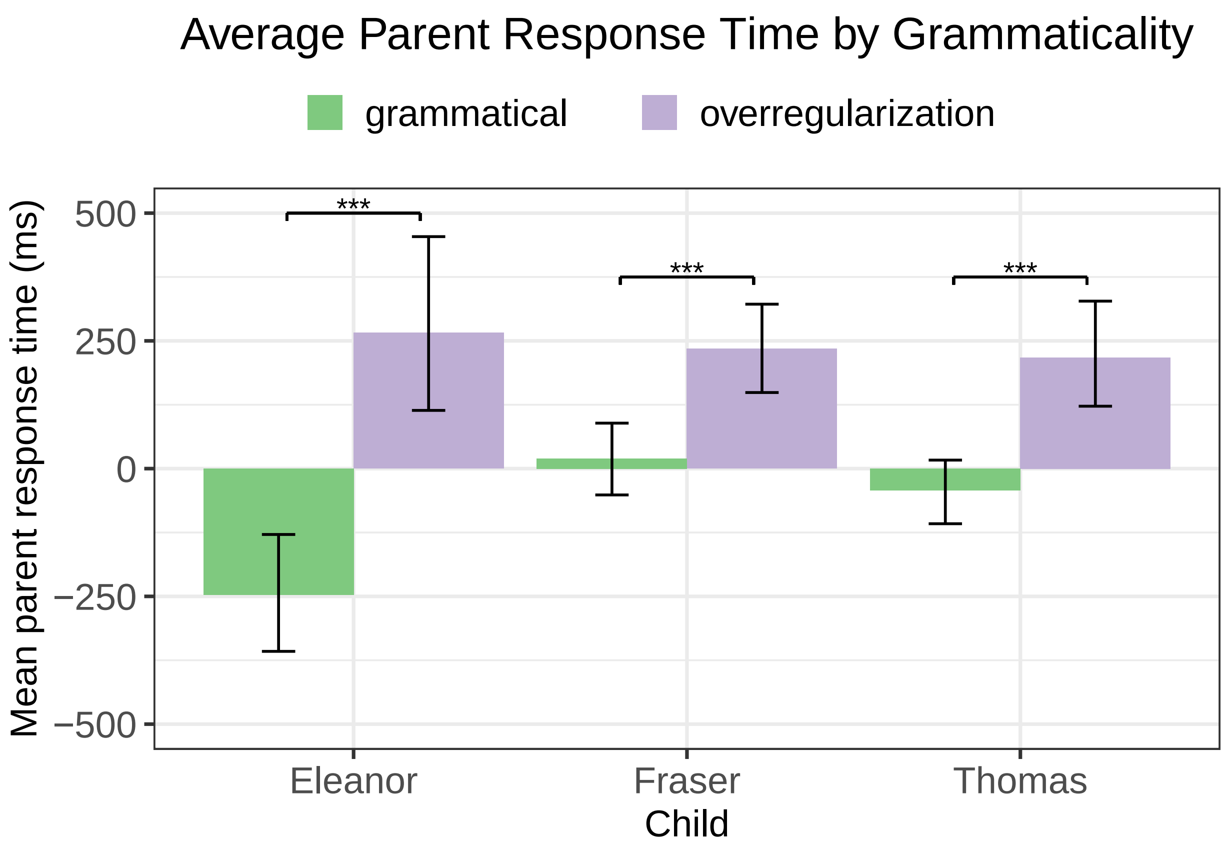

Figure 2

```
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.0003 -0.3367  0.0141  0.3489  5.2937
##
## Random effects:
##  Groups   Name                          Variance Std.Dev. Corr
##  kid      (Intercept)                   0.014158 0.11899
##           error_typeother error         0.006021 0.07759  -1.00
##           error_typeoverregularization 0.004278 0.06541  -1.00  1.00
##  Residual                               0.963069 0.98136
## Number of obs: 2832, groups:  kid, 3
```

```
##
## Fixed effects:
##                                Estimate Std. Error         df t value
## (Intercept)                   3.368e-01  1.051e-01  8.318e+00   3.203
## error_typeother error         1.909e-01  6.273e-02  2.819e+00   3.044
## error_typeoverregularization  3.452e-01  7.923e-02  5.739e+00   4.358
## age                          -1.520e-02  2.133e-03  7.801e+02  -7.125
## rl                            2.591e-02  5.099e-03  2.775e+03   5.081
##                              Pr(>|t|)
## (Intercept)                   0.01193 *
## error_typeother error         0.06036 .
## error_typeoverregularization  0.00531 **
## age                          2.38e-12 ***
## rl                           3.99e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) errr_e errr_t age
## errr_typthe -0.560
## errr_typvrr -0.338  0.461
## age         -0.694  0.003 -0.061
## rl          -0.181 -0.019 -0.010 -0.007
```

Error type (“other error” and “overregularization”) was a significant independent predictor of response time although the p-value for “other error” was slightly above the traditional value. Both overregularizations and errors not due to overregularization correlated with an increase in response time. That is, parents took slightly longer to respond to non-overregularized errors and even longer to respond to overregularized errors. The difference was particularly pronounced for overregularizations, suggesting that there may be something particularly noticeable about overregularizations that could relate to longer response times from parents. However, since the "other error" error type is a catchall for all ungrammatical utterances not containing overregularizations, the particular differences between this error type and overregularizations is unclear. Additionally, age and response length were strongly significant predictors independent of the increase in utterance length that comes with age. The slight decrease with age indicated that parents are faster to respond to their child's utterances as the child ages irrespective of the MLU increase over development. Not surprisingly, the increase with response length indicated that parents took slightly longer to respond when their response was longer. From these results, we can conclude that parents took longer to respond to their child's ungrammatical utterances and took a particularly long time to respond to their child's overregularized utterances. Importantly, this relationship between ungrammaticality and delayed response time occurred independently of age and length of utterance and response.

Discussion

The strikingly similar change in response time across error types indicated a reliable and consistent relationship between parent response time and grammaticality. Generally, parents tended to overlap their children's grammatical utterances quite frequently and in a highly variable way. However, when children produced grammatical errors, particularly overregularizations, parents took longer to respond and did so in a less variable way. These effects remained significant factors when controlling for age, utterance time, and utterance and response length. Given that parents appear to respond in a systematically different ways to ungrammatical utterances, it is possible that children are able to make use of this information in the same way they are able to make use of other statistical regularities in their language environment. Even if children are not sensitive to fine-grained response delays, the manifestation of this delay as a decreased likeliness to overlap the child may be a more salient cue children could take as negative evidence. This analysis demonstrated the presence of a reliable statistical difference in response to grammaticality and thus presents a possible form of negative evidence. The availability of a low-level response cue could allow or help children to draw conclusions about the grammaticality of their utterances.

Although parents clearly took longer to respond to their child's ungrammatical utterances, the nature of parent-child interactions meant that it was impossible to determine if the increased response time was due to the grammaticality of the utterance or some confounding feature of the utterances. To investigate this, we employed an experimental self-paced reading task consisting of utterances from the Thomas and Fraser corpora which were corrected to control for grammaticality.

Experiments

The goal of our experiments was to confirm the correlational results of our corpus analyses with experimental manipulations that would allow us to make causal inferences about our hypothesis that overregularizations lead to slower processing and consequently slower responses. We wanted to isolate the effect of grammaticality from the surrounding context of the utterance. That is, it is possible that parents responded differently not to the grammatical error in the utterance but rather some other feature that exists in concert with overregularized utterances. As such, the stimuli consisted of overregularized utterances from the Thomas and Fraser corpora. We constructed yoked control trials by correcting the overregularization in each critical utterance to its grammatical irregular form (see Fig. 3). Thus, each test trial varied from its control only in grammaticality of the overrergularized verb. These stimuli were used in a self-paced reading task, a standard paradigm for assessing linguistic processing time that is also appropriate to administer remotely (Keller, Gunasekharan, Mayo, & Corley, 2009).

We revised the procedure over three iterations – a pilot of 10 trials with 9 participants, a full experiment of 20 trials with 40 participants, and a replication with 100 participants. See Appendix A for resources detailing the specific changes in each iteration.

Method

Stimuli. Exchanges each containing one verb overregularization between children and their parents were collected. Control trials were created by manually correcting the overregularization. Acceptable exchanges were those with a minimum of four cogent utterances on the same topic. Test trials consisted of the uncorrected exchanges. A trial consisted of one of the exchanges with each utterance shown on the screen as a moving window task one at a time.

speakers	testExchange	controlExchange
CHILD	are these shorts?	are these shorts?
MOTHER	yeah, they're shorts, aren't they?	yeah, they're shorts, aren't they?
CHILD	they're like shorts.	they're like shorts.
MOTHER	they are shorts.	they are shorts.
CHILD	we got some when we goed to the beach, didn't we?	we got some when we went to the be
MOTHER	that's right.	that's right.

Attention check questions consisted of two full utterances both displayed on the screen. One utterance was an utterance that was shown and attributed to the mother. The other utterance was an utterance that was not shown. The false utterance was a randomly selected utterance from the same file that was spoken by the mother and roughly ($\pm$ 1) the same number of words as the correct utterance.

Procedure. Upon accepting the task, participants saw a screen which provided brief instructions about the experiment and consent information. Participants completed a moving-window self-paced reading task (Just, Carpenter, & Woolley, 1982) for each utterance of a trial followed by an attention check. Each participant viewed 5 ungrammatical test trials and 5 grammatical control trials for each child. Participants viewed instructions about the attention check in the same self-paced manner and completed two practice attention check questions before proceeding to the first trial.

Participants were exposed in random order to 5 randomly selected ungrammatical trials and 5 grammatical control trials for each child. The self-paced reading task consisted of the speaker label (CHILD: or MOTHER:) followed by a set of underlined spaces that represented the words of a single utterance. Participants were instructed to press the spacebar to advance each word. When the spacebar was pressed, the previous word that was displayed disappeared and the next word became visible above its respective underline. A timer began at the press of the spacebar and recorded the time until the spacebar was pressed again. This time was defined as the response time for a particular stimulus.

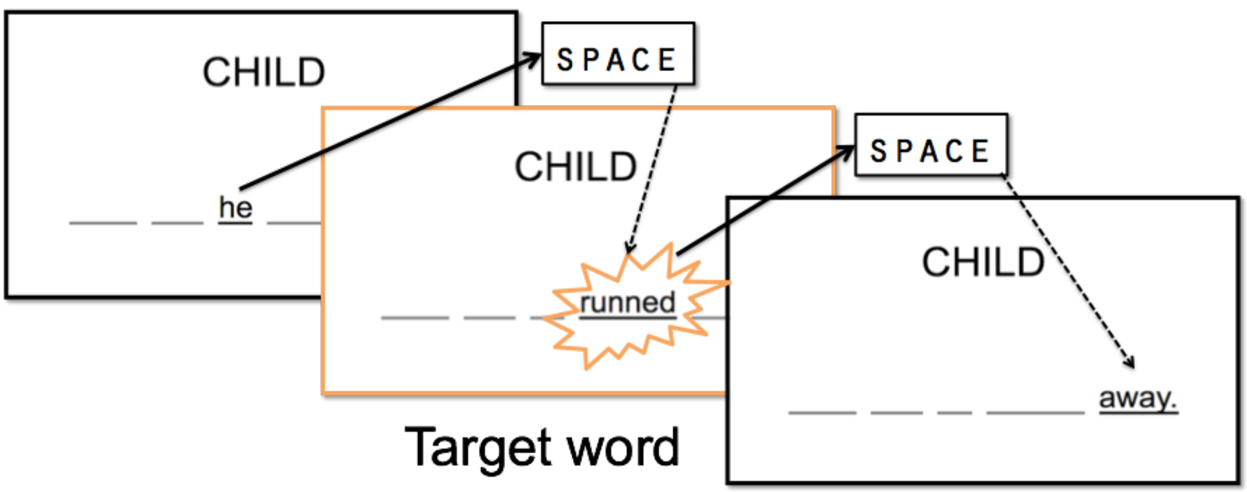

Figure 3. self-paced reading task

After participants had advanced through a full exchange, they were shown an attention check question. Each question displayed two utterances - one that had been seen in the exchange and one that had not. The utterance displayed on the left of the screen was labeled with F and the utterance displayed on the right of the screen was labeled with J. Which side the seen utterance was on was randomized. Participants were instructed to press the key on their keyboard which corresponded to the utterance they had seen.

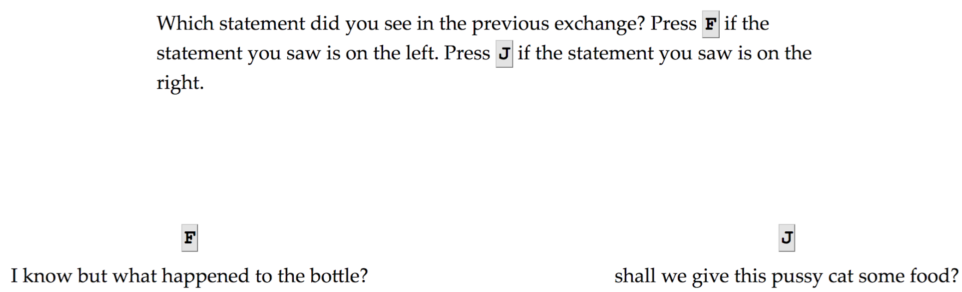

Figure 4. attention check example

Pilot Experiment

Stimuli. Ten exchanges, each containing one verb overregularization between Thomas and his mother were collected. As such, there were a total of 20 possible trials (10 test trials, 10 respective control trials).

Participants. Participants (n = 9) were recruited and tested remotely through Amazon Mechanical Turk.

Results and discussion. We found some basic predicted effects that provided validity for our procedure. These included an exponential decrease in response time from about 450ms across 2 or 3 words which then stayed around 250 - 300ms until a slight increase nearing the end of utterances. Additionally, all but three attention checks were passed (~4% failure) which indicated that the attention checks were of an appropriate difficulty. It was clear, however, that we were underpowered to find the predicted effect of overregularization. Thus, in Experiment 2, we used a larger set of child utterances and increased the length of the experiment.

Experiment I

Participants. Participants (n = 40) were recruited and tested remotely through Amazon Mechanical Turk.

Stimuli. Following the pilot, we added an additional 10 trials by drawing utterances from the Fraser corpus following the same outline described above. There were ten exchanges found for each child that where then manually corrected to create 40 trials for 20 possible test trials that contained an overregularization and 20 possible control trials. An individual trial consisted of an exchange of several utterances between the relevant child and his mother.

Procedure. Each participant viewed 5 ungrammatical test trials and 5 grammatical control for each child for a total of 10 test and 10 control trials. Participants saw all Thomas trials before they saw the Fraser trials. Additionally, participants viewed instructions about the attention check in the same self-paced manner and completed two practice attention check questions before proceeding to the Thomas test trials. After completing all the Thomas trials, participants viewed additional instructions that they would see exchanges between a different child and his mother.

The self-paced reading task followed the same procedure as in the pilot for each child. That is, participants were exposed in random order to 5 randomly selected test trials and 5 control trials for the Thomas utterances and then saw 5 randomly selected test trials and 5 control trials for the Fraser utterances. An attention check question followed every trial for each child.

While each test trial differed from the corresponding control trial by the grammaticality of the test word, the content of each trial may vary a large amount from any other trial because the stimuli were pulled from real parent-child interactions data. Exchanges ranged from 4 to 8 utterances in length but the majority (18 trials of 20) were 5 to 7 utterances long. The total number of utterances labelled as the child's for Thomas was 26 and for Fraser was 30 and the total number of utterances labelled as the mother's for Thomas was 31 and for Fraser was 34.

Results and discussion. We removed all participants that failed more than one attention check (n = 4; 10%) for a total sample of n = 36. Due to the nature of trial content, there were several confounding factors. To control for such factors, we fit a mixed-effect model of the following form: response_time ~ condition * word_position + display_order + utterance_in_exchange + (word_position + condition | subject) + (word_position + condition | trial)

```
##
##  Pearson's product-moment correlation
##
## data:  predicted_data$predicted and predicted_data$rt_per_char
## t = 16.989, df = 2096, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.3097143 0.3849613
## sample estimates:
##       cor
## 0.3478979
```

We predicted that if overregularizations cause a processing delay then participants would have increased response time to and following the test word in ungrammatical trials. This effect was slight but significant in response times to words following the test word but not to the test word. That is, independent of training effects and general position effects, participants took longer to respond to words after reading an overregularization than they did after reading a correct irregular. This indicates that overregularizations independent of other characteristics of a given utterance can result in a processing delay.

Experiment II

The analysis that was conducted on Experiment 1 yielded successful results but was constructed post-hoc. We wanted to replicate the results found knowing in advance the analysis that we would employ so we repeated the experiment using an identical procedure with a larger sample size.

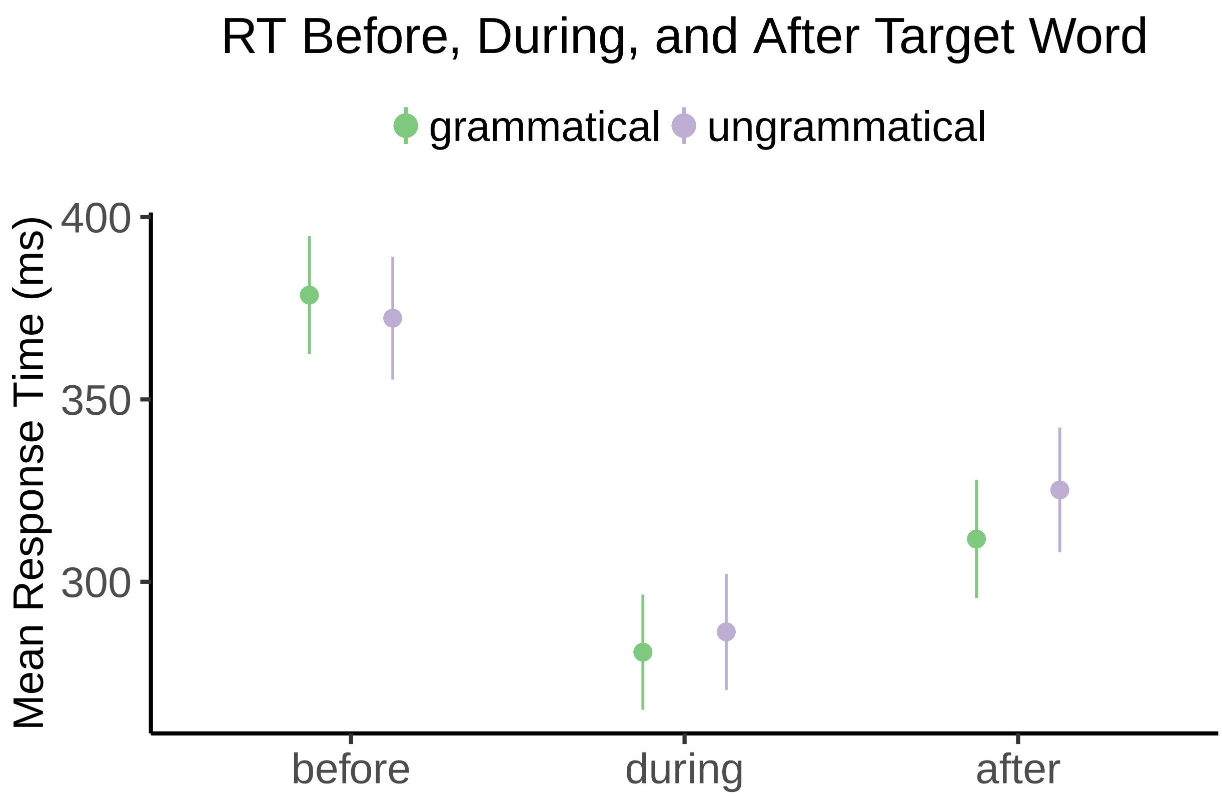

Figure 5

Participants. We followed Simonsohn's recommendation (2015) of a replication sample size that was 2.5 times the size of the original study. As such, we recruited n = 100 participants on Amazon Mechanical Turk.

Stimuli. The trials used were identical to those in Experiment 1 with the small exception of correcting the label on a single utterance in one trial. In the prior experiment, the utterance was labelled "MOM" and in this iteration we corrected it to be labelled "MOTHER" as in all other mother-produced utterances.

Procedure. The procedure was identical to that employed in Experiment I.

Results and discussion. We again removed all data from participants that did not pass at least 95% (19/20) of the attention checks. This resulted in 11 participants being removed from analysis for a total of n = 89. Once again we fit a mixed-effects model to control for subject, training, and position effects.

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
```

```
## lmerModLmerTest]
## Formula:
## word_response_time ~ version * pos + true_display_order + utterance_in_exchange +
##     (pos + version | subj) + (pos * version | trial)
##    Data: aggregate_data
##
## REML criterion at convergence: 61170.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.0440 -0.5618 -0.1255  0.3669  7.0731
##
## Random effects:
##  Groups   Name              Variance Std.Dev. Corr
##  subj     (Intercept)       7003.37  83.686
##           posat             1064.30  32.624   -0.47
##           posafter          1358.58  36.859   -0.07  0.63
##           versionu            92.72   9.629    0.35  0.17  0.41
##  trial    (Intercept)       2695.10  51.914
##           posat             2645.53  51.435   -0.95
##           posafter          6520.43  80.749   -0.92  0.98
##           versionu           110.48  10.511    0.44 -0.36 -0.27
##           posat:versionu     452.15  21.264   -0.43  0.31  0.29 -0.43
##           posafter:versionu  380.45  19.505   -0.39  0.27  0.21 -0.80
##  Residual                   6929.86  83.246
##
##
```

```
##
##
##
##
##
##
##
##
##   0.87
##
## Number of obs: 5182, groups:  subj, 89; trial, 20
##
## Fixed effects:
##                        Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)            395.0061    17.7249  55.1586  22.285  < 2e-16 ***
## versionu                 3.6210     4.7973  25.1425   0.755  0.45738
## posat                  -84.4832    12.6672  22.1413  -6.669 1.02e-06 ***
## posafter               -64.7394    18.9109  20.7293  -3.423  0.00259 **
## true_display_order      -5.4684     0.3733 609.4375 -14.649  < 2e-16 ***
## utterance_in_exchange    7.7607     2.8909  19.0141   2.685  0.01466 *
## versionu:posat          -2.4832     7.4168  22.5282  -0.335  0.74087
## versionu:posafter       22.8674     7.1925  18.5921   3.179  0.00504 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) versin posat  posftr tr_ds_ uttr__ vrsn:pst
```

```
## versionu     0.080
## posat       -0.664 -0.014
## posafter    -0.607 -0.016  0.917
## tr_dsply_rd -0.078  0.007  0.001  0.000
## uttrnc_n_xc -0.498  0.002 -0.001  0.001 -0.260
## version:pst -0.122 -0.597  0.009  0.121  0.000  0.006
## versn:psftr -0.089 -0.715  0.060  0.000  0.000 -0.001  0.646
```

```
##
##  Pearson's product-moment correlation
##
## data:  predicted_data$predicted and predicted_data$word_response_time
## t = 24.971, df = 5180, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.3032632 0.3518724
## sample estimates:
##       cor
## 0.3277848
```

```
##
##  Welch Two Sample t-test
##
## data:  word_response_time by version
## t = -3.9757, df = 1674.1, p-value = 7.315e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -36.84274 -12.50001
## sample estimates:
```

```
## mean in group g mean in group u
##        299.5091        324.1804
```

We successfully replicated our results from the prior experiment. There was an interaction effect of condition and word position in test utterance such that participants took significantly longer to respond to words after seeing an overregularized word than after seeing a grammatical word. According to the mixed-effects model, this effect was significant independent from general position and training effects.

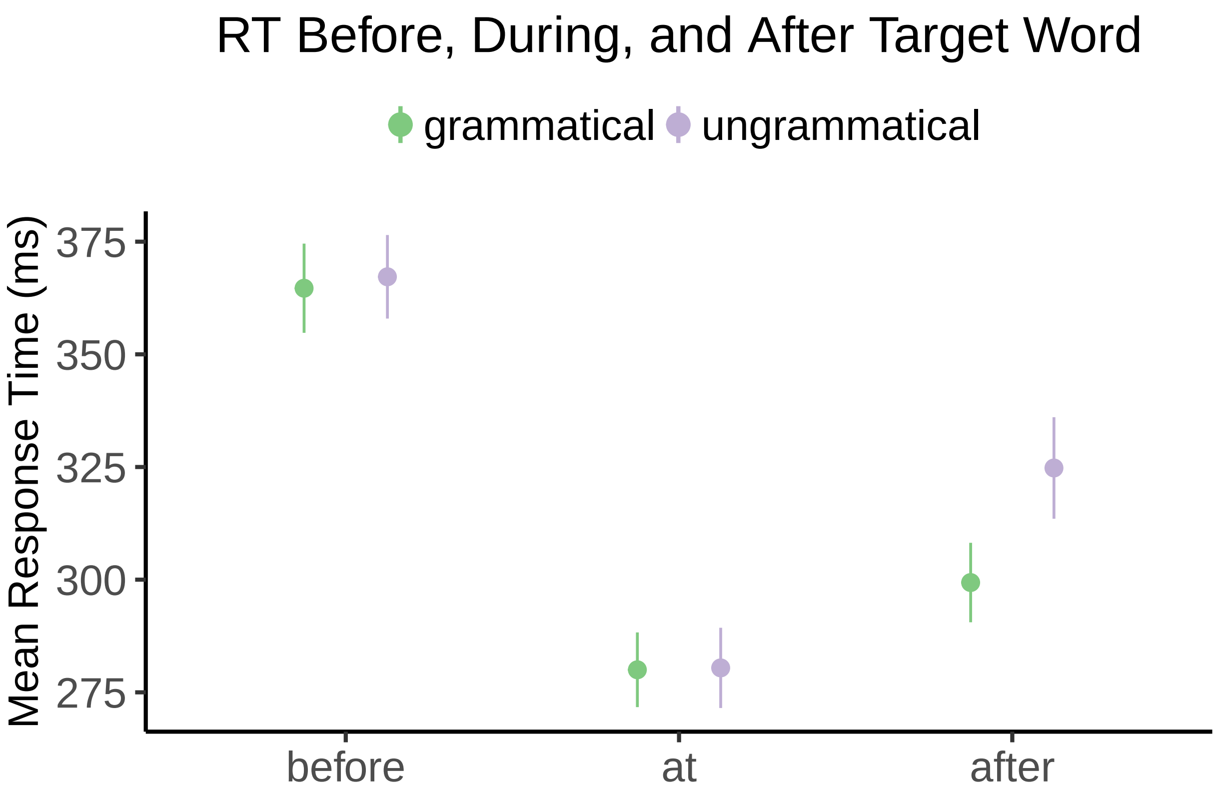

Figure 6

General Discussion

We hypothesized that there exists a lower-level form of negative evidence - parent response time. By examining three CHILDES corpora we found that parents do, in fact, respond reliably later to ungrammatical past tense and plural utterances. Additionally, parents appear to take particularly long to respond to overregularizations than to utterances that are ungrammatical for other reasons. We were also able to conclude from experiments 1 and 2 that adults take longer to respond utterances due to the presence of overregularizations specifically rather than some confounding feature of the utterance. By using a self-paced reading procedure, we claimed that this increase in response time is due to processing difficulty rather than another factor such as desire to correct.

Future Directions

Given the consistency of the change in response times across distinct corpora and even research projects, further work might consider employing additional child-parent interaction corpora to complete a large-scale analysis of parent response times. To better understand the driving forces behind parent response time latencies, several experiments should be conducted which more closely match the environment in which parents are responding to their children. A clear extension of study 2 would be an online self-paced listening task. Additionally, future work should consider how parents may adapt to their child's particular language errors over time. For instance, perhaps it is more surprising and subsequently more difficult to process when their child does not overregularize something that they frequently overregularize.

Also important to consider is how surprising certain overregularizations appear both when listening - in the case of parents, and when reading - in the case of our online self-paced reading task. One reasonable way we've considered investigating this is to compute the Levenshtein distance between a given overregularization and the correct regularization on the IPA transcription of the overregularization. We anticipate that response time will increase with increased Levenshtein distance.

Finally, an experimental procedure which attempts to teach children an artificial grammar by varying the response time latency of conversational responses would be an excellent next step for examining whether children are sensitive to and can learn from response time delays of the scale we found in natural parent-child interactions.

Conclusions

The presence of a reliable delay in response time from parents across age in response to ungrammatical utterances indicates that there is a lower level signal of grammaticality that could act as negative evidence for grammar learning. Significantly, the presentation of new forms of negative evidence could address the "no negative evidence problem" that has led many researchers to conclude, perhaps prematurely, that there must exist innate constraints on grammar learning (Bowerman, 1988). This examination of a little-researched form of negative evidence raises the question of what other cues might provide informative feedback for language learning.

References

Brown, R., & Hanlon, C. (1970). Derivational complexity and order of acquisition in child speech. In J. Hayes (Ed.), *Cognition and the developmenf of language*. New York: Wiley.

Chouinard, M. M., & Clark, E. V. (2003). Adult reformulations of child errors as negative evidence. *Journal of Child Language*, *30*(3), 637–669.

Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition.

Frank, M. C., & Goodman, N. D. (2014). Inferring word meanings by assuming that speakers are informative. *Cognitive Psychology*, *75*(C), 80–96.

Gold, E. M. (1967). Language identification in the limit. *Information and Control*, *10*, 447–474.

Grice, H. P. (1969). Utterer's meaning and intention. *The Philosophical Review*, *78*, 147–177.

Hirsh-Pasek, K., Treiman, R., & Schneiderman, M. (1984). Brown & Hanlon revisited: Mothers' sensitivity to ungrammatical forms. *Journal of Child :Anguage*, *11*, 81–88.

MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk. Third Edition.* Mahwah, NJ: Lawrence Erlbaum Associates.

Marcus, G. F. (1993). Negative evidence in language acquisition. *Cognition*, *46*(1), 53–85.

Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, *14*(1), 57–77.

Morgan, J. L., & Travis, L. L. (1989). Limits on negative information in language input. *Journal of Child Language*, *16*(3), 531–552.

Newport, E., Gleitman, H., & Gleitman, L. R. (1977). Mother i'd rather do it myself: Some effects and non-effects of maternal speech style. In C. E. Snow & C. A. Ferguson (Eds.), *Talking ro talking to children: Language inpuf and acquisition* (pp. 109–149). Cambridge, UK: Cambridge University Press.

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science*, *13*(1), 13–19.

Tenenbaum, J. B. (1999). *A bayesian framework for concept learning* (PhD thesis). Massachusetts Institute of Technology.

Zipf, G. K. (1949). *Human behavior and the principle of least effort. Addison-Welsey: Reading, Mass.* Addison-Welsey: Reading, Mass.