How Children Learn to Disappear Causative Errors

Ava Irani

1. Introduction

In English, some verbs like break can occur both as an intransitive (e.g., the vase broke) or as a causative transitive (e.g., John broke the vase). Many verbs display this property, known as the causative alternation, but not all verbs that occur as an intransitive have a causative transitive counterpart (e.g., fall, disappear). Adults are well-aware of this restriction, but children, on the other hand, make overgeneralization errors (Bowerman 1982; Pinker 1989) by extending the causative form to intransitive only verbs. Some of these errors can be seen in (1).

(1) Child causative errors (from Bowerman 1982):
   a. He’s gonna die you David (4+)
   b. Kendall fall that toy (2;3)
   c. You ached me (4;1)
   d. She came it over there (3;4)

These errors indicate that there is a time when the learner has acquired a grammar that is a superset of the adult grammar. The question then is, how does the learner retreat from the superset grammar, and arrive at the more conservative adult grammar? Previous approaches have argued that this retreat from overgeneralization occurs through the use of indirect negative evidence (e.g., Ambridge 2013). In contrast, I argue in this paper that this developmental trajectory is a result of learning from positive evidence.

This paper is structured as follows: first, I describe previous approaches that use indirect negative evidence to explain how children retreat from overgeneralization errors. Next, Section 3 proposes an alternative analysis that uses learning from positive evidence, and describes how children overgeneralize from the linguistic input. Section 4 proposes an account of how children retreat from overgeneralization errors using positive evidence.
2. Previous Approaches

In this section, I describe approaches invoking preemption and entrenchment to account for the overgeneralization facts, and show why they are inadequate in accounting for the problem.

2.1. Evaluating Preemption

Preemption accounts claim that children retreat from overgeneralization errors when the ungrammatical form is overridden by a more frequent grammatical form that is evidenced by the learner in the linguistic input (Bowerman and Croft 2008; Pinker 1989). Verbs like die and come are said to have suppletive transitive causative counterparts kill and bring that the learner eventually acquires. Pragmatic situations will call for the use of the more frequent grammatical forms of verbs and the ungrammatical forms are simply never used in later stages of acquisition, as they never occur in the input. For verbs that do not have a phonologically distinct lexical causative form (like stay, but unlike die/kill), make-causatives are argued to preempt the lexical causative form. However, the make-causative systematically carries a different meaning from the lexical causative (Ammon 1980; Bowerman and Croft 2008; Fodor 1970), and is its own productive construction that occurs independently. Hence, the make-causative is said to be a weaker cue in the input than a phonologically distinct lexical causative form. Errors with verbs like disappear, in that case, are predicted to persist longer than those with verbs like die, as there is no competing form to preempt it. However, Bowerman and Croft (2008) show that this prediction is not borne out in examining the causative errors of two children C and E.

![Figure 1: Child C’s causative errors over time (from Bowerman and Croft 2008:297, Figure 13.2).](image-url)
As illustrated by Figures 1 and 2, there is no visible difference in the rate of abating errors between verbs that have a suppletive counterpart and verbs that do not for either child. Additionally, for verbs that do not have a suppletive counterpart, it is unclear whether make-causatives do indeed preempt the causative forms of overgeneralized verbs. Child production data indicates that children are aware of make-causative early on, even for verbs that they overgeneralize. The following examples of make-causatives are found in child speech:

(2) Make-causatives in child data:
   a. someone made me go away (2;9, Ross, MacWhinney corpus)
   b. I’m going to make it disappear (3;2, Ross, MacWhinney corpus)
   c. Shoes may make you jump higher (3;9, Adam, Brown corpus)
   d. you stand up to make it go fast (3;10, Adam, Brown corpus)
   e. She’s gonna make it die (5;0, Christy, Bowerman 1982)

As seen from the examples above, the verbs in (2) are exactly those verbs that children are known to make errors with. Moreover, the use of the make-causative overlaps with the time at which the overgeneralization of the causative rule occurs, indicating that children are aware of this alternative construction at the time when the errors are produced. Some examples of the causative errors made by Adam (Brown corpus) and Ross (MacWhinney corpus) are provided in (3).

(3) Causative errors produced by Adam and Ross:
   a. And my mommy might break this and fall this . (3;0, Ross, MacWhinney corpus)
b. I want to disappear it. (3;3, Ross, MacWhinney corpus)
c. Gon(na) fall him to pieces. (3;7, Adam, Brown corpus)
d. How to go it? (4;7, Adam, Brown corpus)

The examples in (3) show that both Adam and Ross know the make-causative construction when they produce the lexical causative errors. The make-causative examples produced with verbs like *go* and *disappear* also illustrate that both children are aware of an alternative construction in place of a potential lexical causative form of these verbs. Early use of the make-causatives here indicate that even if make-causatives could preempt lexical causatives in theory, they do not do so in practice.

2.2. Evaluating Entrenchment

Another solution that has been proposed to account for the problem of overgeneralization is that of entrenchment (Ambridge 2013; Brooks and Tomasello 1999, i.a.). Entrenchment accounts claim that if a verb is heard as an intransitive, but not as a transitive, enough times then the learner assumes that the transitive frame is impossible altogether. Brooks and Tomasello (1999) argue, for instance, that children are more likely to use an intransitive frame for novel verbs heard in an intransitive frame.

An entrenchment approach, which claims that children cease to make causative errors once they have heard a verb used in an intransitive frame a number of times, predicts a difference in the production of errors between verbs of varying frequencies. The child is expected to retreat from using a pure unaccusative verb in a causative frame earlier if the verb occurs with a relatively high frequency in the input. In addition to the aspect of conservative learning implied by the entrenchment account, the prediction of errors made by this approach is not borne out.

This prediction can be tested by examining the causative errors made by the children in CHILDES. Ross from the MacWhinney corpus (MacWhinney 2000), for instance, makes several causative errors. Some of these errors are shown below.

(4) a. Maybe I can fall it down the stairs (3;3)
b. And are you going to stay me at my new school at Pittsburgh (3;5)
c. To go it down my tummy (3;11)

As the examples above illustrate, Ross makes causative errors with a variety of verb types, and these verbs vary in terms of their frequency. A search through all the caretakers’ speech in CHILDES, i.e., the combined input data from all the corpora, allows us to obtain the total number of times these verbs occurred in the input. Out of the six million words of combined data, *disappear* occurs 153 times overall. *Stay* occurs a total of 2,662 times, *fall* a total of 2,819 times, and *go* occurs
55,689 in all. However, in spite of the differences in their frequency, the errors we find for Ross all cluster together around the age of 3-4 years, and there is no evidence that the child stops producing errors from these verbs purely on the basis of hearing them intransitively a number of times. If verb properties were indeed susceptible to entrenchment, we would not expect children to be making errors with verbs like *go* that occur over fifty thousand times in the input data. If it was true that a verb needed to be heard that many times in order to learn its argument structure, we would expect children to be making causative errors with verbs like *disappear*, which occur a fraction of the time, well until their adult years. As we are aware, this is not the case.

Instead, I argue below that children learn verb properties from positive evidence. They produce overgeneralization errors when there is sufficient evidence for a rule in the input, and they cease to generalize when there is insufficient positive evidence.

3. Learning Causativization

In order to learn from positive evidence, the learner must make use of structural information along with situational cues. This information has been argued to be relevant for learning verb argument structure properties, and more specifically, in learning the causative construction (Naigles 1990, 1996; Naigles and Hoff-Ginsberg 1995; Pinker 1989). In this section, I describe how structural and situational cues aid the child in acquiring the different kinds of intransitive verbs.

We delve into how the learner uses notions such as intentionality to distinguish between unaccusative and unergative verbs, which is a relevant distinction for the causative alternation rule. Once the learner has identified the two types of unaccusatives, I show how they generalize the causative rule from the evidence available to them in the input. I also demonstrate how as the learner acquires more verbs, the proportion of verbs that undergo the causative alternation rule decreases, and the learner no longer generalizes. Furthermore, I argue that the motivation behind subdividing verbs into smaller classes may arise from the need to form productive rules in language (cf. Yang 2016), i.e., the learner searches for subclasses of verbs to identify a class where the causative alternation rule is in fact generalizable.

3.1. Learning Verb Classes

This section describes the situational concepts and structural information used by the learner in acquiring verb classes. I first discuss how unaccusative and unergative verbs are differentiated, as this distinction results in the overgeneralization of the causative rule. The adult grammar shows a productive causative rule for some classes; e.g., the manner of motion class consisting of verbs like *roll* (see also Brooks and Tomasello 1999), but not all kinds of verbs show this rule productively. For example, verbs of inherently directed motion do not generally
display this property (e.g. *rise*). The learner must, therefore, acquire verb classes that productively employ the causative alternation rule, while at the same time, learn that not all verbs undergo this alternation.

Languages are known to exhibit two kinds of intransitive verbs (e.g. Perlmutter and Postal 1984): unergatives and unaccusatives. Unergative verbs introduce an external argument, and optionally allow certain kinds of objects, while unaccusative verbs do not introduce an external argument. Given these two types of verbs, the learner is faced with the task of distinguishing between the two classes of verbs, as a host of properties are associated with each type. Learning the two classes of verbs is also relevant to the acquisition of causatives as it is primarily the class of unaccusative verbs that show the property of the causative alternation; unergative verbs generally do not display this property. Consequently, the generalization of the causative rule that children make cannot stem from considering the entire class of unergative verbs as well. In order for the rule to become a candidate for generalizing, the learner must have identified the distinction between them. Below, I describe how children make this differentiation between unaccusative and unergative verbs.

3.2. Structural and Situational Cues

I argue that children use intentionality as a differentiating factor between unaccusative and unergative verbs. Unergatives verbs in their intransitive frame are known to typically show intentionality (e.g., *Jamie ate*), whereas unaccusative verbs in their intransitive frame lack intentionality (e.g., *Jamie fell*). An unergative verb like *eat*, thus, shows intention, whereas an unaccusative verb like *fall* does not. There is some evidence in the literature that children are attuned to notions like intentionality early on (e.g., Tomasello and Barton 1994; Woodward 1998). For instance, Woodward (1998) uses a visual habituation paradigm to show that children distinguish between animate and inanimate agents. In her studies, infants between the ages of 4 to 11 months were habituated to the motion of a hand grasping or a rod touching one of the two objects placed in front of them. In the testing phase, the infants were shown either a hand or rod grasping or touching a different object. The results showed that infants in the hand condition, but not the rod condition looked longer at the action when there was a new goal with an old path versus in trials that showed a new path with an old goal. These results indicate that infants as young as a few months old are able to encode the goal-oriented nature of actions. As children are attuned to intentionality early on, I will use intentionality to model the distinction between unergative and unaccusative verbs.

Structural information also comes into play in learning causative verbs. For some abstract verbs, if a causative form exists in the language, it must be learned from the linguistic input. This fact is clear from cross-linguistic differences where a verb like *arrossire* ‘blush’ has been argued to have a direct causative form in Italian (e.g., Levin and Hovav 1992), but not in English where sentences like *the sun blushed my cheeks* are ungrammatical. Similarly, the Hindi verb *khilaanaa* ‘to
make bloom’ or ‘to make blossom’ has a lexical causative form, which is a pure intransitive in English. This verb in English is said to be internally caused, and therefore, does not undergo external causation (e.g. Pinker 1989), but this claim does not hold for this verb in Hindi. However, as discussed earlier, the possibility of direct causation need not be learned from the input for all verbs, especially not verbs that involve physical contact such as fold. In these cases, there is situational evidence for a verb to be learned as involving direct causation, but we cannot do away with the structural information that supplements the learner’s knowledge for more abstract verbs denoting a direct causative relationship. Thus, both situational and structural linguistic cues are indispensible to the learner in acquiring the finer primitive meanings of unaccusative verbs.

The importance of structural cues is also evident in learning motion verbs. There are structural correspondences in the syntax that indicate motion in the meaning of a verb. For instance, verbs followed by locative prepositions often indicate a motion verb like roll (Fisher, Gleitman, and Gleitman 1991; Talmy 1975). This structural cue is available for the class of inherently-directed motion verbs and manner of motion verbs. Thus, we have seen so far how the combination of structural and semantic cues allows the learner to distinguish between unaccusative and unergative verbs, externally caused verbs and internally caused verbs, and motion verbs from other unaccusative verbs.

In sum, studies on children’s knowledge of linguistic and situational cues indicate that children are attuned to concepts such as intentionality and causality as early as four months (Carpenter, Akhtar, and Tomasello 1998; Woodward 1998; Woodward, Sommerville, and Guajardo 2001). Leaving aside the question of whether children are innately endowed with the ability to identify cues such as causality and intentionality, we can at least use the findings that children are aware of these concepts early on to understand how they learn to causativize verbs. Now that we have seen which structural and semantic cues are relevant to the learner in acquiring the semantic and syntactic properties of verbs, we can move on to the question of how children overgeneralize the causative alternation rule from positive evidence in the input, and how they eventually retreat from this overgeneralization.

3.3. Overgeneralization

In this section, I present an analysis of the developmental trajectory seen in the acquisition of the causative alternation rule that uses learning from positive evidence. I show that children overgeneralize the causative alternation rule when there is sufficient positive evidence, and they retreat from this overgeneralization when there is insufficient positive evidence. Sufficient positive evidence here is determined by the Sufficiency Principle proposed by Yang (2016). The Sufficiency Principle is provided in (5).
(5) “The Principle of Sufficiency: Let \( R \) be a generalization over \( N \) items, of which \( M \) items are attested to follow \( R \). \( R \) can be extended to all \( N \) items iff:

\[
N - M < \theta_N \text{ where } \theta_N := \frac{N}{\ln N}\]

(Yang 2016:140)

The causative overgeneralization errors only arise when there is sufficient evidence in the input as determined by the Sufficiency Principle. If there is sufficient positive evidence, the learner then generalizes the rule over other members in a class. We can test this hypothesis by investigating the causative errors made by children in corpus data (CHILDES, MacWhinney 2000). I examined Adam (Brown 1973) and Ross (MacWhinney 2000) as two case studies of children who go through a stage of overgeneralizing the causative alternation. These two children are good candidates to test the claims made so far as there is more data available for them than most individual children in CHILDES, and both of these children make causative errors.

3.3.1. Test Case 1: Adam

The errors noted for Adam include unaccusative verbs such as come and go, among others. Examples of the errors he produces are illustrated in (6).

(6) a. Don’t fall my head (3;2)
   b. Gon(na) fall him to pieces (3;7)
   c. How to go it? (4;7)
   d. If you can go it fast the pictures might run (4;7)

The causative errors for Adam extend until age 4;7; however, since the data for Adam ends at age 5;2, we cannot say for sure when these errors stopped occurring. The errors we do find indicate that the stage of overgeneralization lasts for as long as Adam’s corpus.

To establish whether the causative errors arose in the input from sufficient positive evidence, I examined Adam’s production data for each verb used until the first unambiguous error was produced at age 3;2. Up to age 3;2, Adam produces a total of 208 verbs in CHILDES. Out of these 208 verbs, 110 of them occur as a plain intransitive. From these verbs, using intentionality as a guideline for distinguishing most unergative verbs from unaccusative verbs, we were left with 51 verbs that lacked intentionality in an intransitive frame. These verbs included two verbs, laugh and cry, which are generally classified as unergatives in the adult grammar (Perlmutter and Postal 1984). The remaining 59 intransitive verbs were intentional intransitives. Adam’s verbs and the productivity of the causative alternation rule is summarized in Table 1.

---

1 The reader is referred to Yang (2016) for details regarding the derivation of the Tolerance Principle and the Sufficiency Principle.
Table 1: Total number of Adam’s transitive and intransitive verb types.

<table>
<thead>
<tr>
<th>Verb Type</th>
<th># of Verbs</th>
<th>Causative Verbs</th>
<th>Productive?</th>
</tr>
</thead>
<tbody>
<tr>
<td>transitive</td>
<td>98</td>
<td>9</td>
<td>NO</td>
</tr>
<tr>
<td>intentional intrans.</td>
<td>59</td>
<td>9</td>
<td>NO</td>
</tr>
<tr>
<td>unintentional intrans.</td>
<td>50</td>
<td>37</td>
<td>YES</td>
</tr>
<tr>
<td>Total</td>
<td>206</td>
<td>48</td>
<td></td>
</tr>
</tbody>
</table>

The child’s production data within a single corpus can be used to determine whether the child knows a verb, but not to gauge whether the child can use the verb in all possible frames. In order to estimate the learner’s vocabulary size, data from each caregiver in the North American English CHILDES corpora were examined, rather than just the single child’s caregivers’ speech. Words in the input follow a Zipfian distribution; i.e., many verbs only occur a few times, or even just once in the input data. Therefore, it is unexpected that we would find each verb occurring in all possible syntactic structures in a single caregiver’s datafile. In contrast, examining the caregiver’s production data for all of CHILDES provides us with a reasonable guess as to whether a verb occurs as a causative in the linguistic input in early stages of acquisition. Combining all the corpora from CHILDES, I checked to see whether the above verbs used by Adam occurred as a causative in the input.

One unintentional plain intransitive, `fold`, had to be excluded from the analysis as it did not occur in an inchoative frame; i.e., it only occurred as infinitives. Out of the remaining 50 verbs, 12 of them did not occur with a causative frame in the input data in CHILDES. Many of these verbs like `fall` do not have a causative form in the adult grammar. Following the sufficiency principle, a class with 50 members requires at least 37 verbs to show a property in order for the property to be generalizable. Here, 38 verbs show the causative alternation, and hence, the rule is generalizable to the other members of the class, and Adam makes causative errors resulting from overgeneralization of the rule.

Additionally, we would not expect Adam to generalize the causative alternation if all intransitive verbs were considered. In total, Adam used 59 verbs that were classified as intentional intransitives. Out of these verbs, only 9 of them causativize. As only a fraction of these verbs have a causative form, it would be unexpected that the child would generalize the causative alternation rule if they were taking into account the entire set of intransitive verbs. The total number falls well below 50%, and most models would not make predictions for the learner to generalize so aggressively. Thus, we see here that Adam generalizes in the presence of sufficient positive evidence.

---

2Note that `fold` also occurred in the causative frame in the input, and including it in the analysis would add one more verb that follows the alternation rule, which further supports the analysis presented in this paper.
3.3.2. Test Case 2: Ross

Another case study of the causative overgeneralization was obtained from Ross’s data in the Macwhinney corpus. Like Adam, Ross also produced causative errors around the ages of 3-4 years, as shown in (7).

(7) a. And are you going to stay me at my new school at Pittsburgh (3;5)
   b. And my mommy might break this and fall this (3;0)
   c. I want to disappear it (3;3)
   d. How did it disappear this air out of here (4;2)
   e. To go it down my tummy (3;11)

The last error we see for Ross is around age 4;1. The errors here are clustered around the ages of 3 and 4, and hence, we can check to see if there was sufficient evidence in the input for these errors. This can be achieved by looking through the child’s production data.

Similar to what was done for Adam, I examined Ross’s data to find the list of verbs that he knew up to the time the first error was produced. For Ross, I obtained a total of 121 verbs were obtained. Out of 121, 66 verbs occur as plain intransitives in CHILDES. 42 of these intransitives showed intentionality. The remaining 24 verbs were intransitives lacking intentionality, which also included *cry*, an unergative verb. The different verb types are summarized in Table 11. 18 of the 24 verbs occurred with a causative form in the input in CHILDES. A class with 24 verbs requires 16 for the rule to be generalizable, and thus, the causative rule is overgeneralized to the other verbs.

<table>
<thead>
<tr>
<th>Verb Type</th>
<th># of Verbs</th>
<th>Causative Verbs</th>
<th>Productive?</th>
</tr>
</thead>
<tbody>
<tr>
<td>transitive</td>
<td>55</td>
<td>7</td>
<td>NO</td>
</tr>
<tr>
<td>intentional intrans.</td>
<td>42</td>
<td>18</td>
<td>YES</td>
</tr>
<tr>
<td>unintentional intrans.</td>
<td>24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>121</td>
<td>66</td>
<td></td>
</tr>
</tbody>
</table>

Additionally, causative errors would be unexpected for Ross if the intransitive verbs were not divided along the lines of intentionality. Only 7 of the 42 verbs used by Ross occur in a transitive causative frame in the input. If the learner were to not categorize the verb classes this way, we would not expect any errors of overgeneralization. Adding the 7 intentional intransitives that causativize, there would only be 25 out of 66 intransitives in all that follow the causative alternation rule, which is well below 50%. Thus, overgeneralization errors would be unexpected under any model of generalization that requires a majority to follow a rule.
4. Retreat from Overgeneralization

Now that we have seen how overgeneralization occurs, we can discuss how children eventually retreat from their overarching hypothesis to arrive at the adult grammar. I claim here that the errors cease to occur when there is insufficient positive evidence in the input. To find evidence for the unproductivity of the rule in later stages of acquisition, we must estimate the growing vocabulary of the learner. To test whether the number of verbs that undergo the rule indeed decreases with increase in vocabulary size, I examined the properties of verbs listed in Levin (1993). I then found the frequencies for these verbs in CELEX, and found the verb with the lowest frequency that occurred in CHILDES. This way, we can estimate the number of verbs the child comes to know later on. The verb with the lowest frequency that occurred in CHILDES is *germinate*, which has a frequency of 1 per million. Hence, only unintentional intransitive verbs that occur with a frequency of once per million or higher were considered. This resulted in a total of 261 verbs. Out of the 261 verbs, 65 occur only as intransitives; i.e., only 196 of the verbs have a causative form. According to the Sufficiency Principle, a class with 261 members requires at least \( N - N/\ln(N) = 214 \) verbs to follow the rule. 196 is insufficient evidence for the learner to generalize, and therefore, they cease to apply the rule to verbs for which they have not encountered the causative form.

It should be noted here that learning verb properties cannot stop at this point. For instance, Maratsos et al. (1987) provide evidence that both adults and children use novel verbs that indicate external causation in a causative frame. These results show that the grammar does have productive rules for causativization. The approach I have taken here allows us to address how children end up with a productive rule for some verb classes, but not for all intransitive verbs overall. Recall that the goal of the learner is to find rules and patterns within their language to facilitate mastery of the grammar. Once the learner reaches the stage where the rule for forming causatives is not generalizable, they are motivated to look for further subclasses where the rule might in fact be productive. Two subclasses that are known to allow the causative alternation productively are the manner of motion verb class and verbs that indicate external causation. I claim that children then learn these subclasses when the rule is not generalizable to the larger set of unintentional intransitive verbs.\(^3\) The overall learning mechanism is illustrated in Figure 3.

As seen in Figure 3, the learner first posits a linguistic class that has certain properties associated with it. In this case, the learner posits a class of unintentional or intentional intransitives. The rule under consideration for this class is the causative alternation rule. The learner then computes, according to the Sufficiency Principle...
Principle, whether or not the rule observed is a productive or unproductive one. If the rule is productive, the learner generalizes. If the rule is unproductive, the learner seeks subclasses where the rule may be productive.

Figure 3: Summary of the learning mechanism employed by children. Rules are determined based on sufficient positive evidence in the input as ascertained by the Sufficiency Principle. If the rule is unproductive, the learner seeks subclasses where a productive rule may be found.

5. Conclusion

In this paper, I have shown how the child can learn the causativization rule in their language from positive evidence in the input. Crucially, it is not merely possible for us to assume learning from positive evidence, but absolutely necessary. We have seen here that evidence for indirect negative evidence in the input is sparse, and not useful for the learner in acquiring verb argument structure properties in their language.

References


