If They Had Been More Transparent, the Child Would Have Discovered Them More Easily: How Counterfactuals Develop

Lina Rouvoli, Vina Tsakali, and Nina Kazanina

1. Introduction

Language is a powerful device that allows humans to communicate about concepts and events that are *irrealis*, i.e., not part of the real world. From a developmental perspective, an important question is when and how children develop the ability to understand and talk about hypothetical worlds, and as a special case, counterfactual worlds. This study focused on *Past Counterfactual conditionals* (PastCFs), a linguistic construction that refers to events in the non-actual world, and investigated when and how children become aware that these constructions are counterfactual. We argue that the understanding of the grammatical construction that expresses counterfactuality in the past, i.e. if + Pluperfect, is separable (although not independent) from counterfactual reasoning abilities.

2. The phenomenon of counterfactuality

In language, the term ‘counterfactual’ refers to expressions that concern events that did not happen in the real world, hence are counter to the factual state of the real world. Past Counterfactual conditionals (PastCFs) such as (1) are a clear example of a counterfactual structure.1

(1) If Amy had eaten the apple, she would have won the medal.

(1) conveys prominently that Amy did *not* eat the apple in the past and she did *not* win the medal in the past, i.e., it considers a conditional relationship between *eating an apple* and *winning a medal* events that was not attained in the

---

* Lina Rouvoli, University of Bristol, (email: er15301.2015@my.bristol.ac.uk), Vina Tsakali, University of Crete, (email: tsakali@uoc.gr), Nina Kazanina, University of Bristol, (email: nina.kazanina@bristol.ac.uk).

1 Besides conditionals, counterfactual meaning can also be found in the context of *wishes*, as in ‘I wish I had/had had a million dollars’ which exhibit a significant number of similar properties to counterfactual conditionals (see Iatridou 2000 for discussion); however, these constructions are beyond the scope of the present study.

real world. It has a hidden negation in its meaning in that the events described in each clause contradict the reality.²

The hidden negation present in the meaning of (1) is absent in Simple Conditionals such as (2) that refer to the future:

(2) If Amy eats an apple, she will win the medal.

Indeed, (2) is compatible with a scenario where Amy eats an apple and Amy gets a medal in the future and is evaluated on the basis of the events in the real world.

The hidden negation in (1) (also available with Present Counterfactuals, see footnote 2) has given rise to a number of theoretical questions regarding the inference of this negation despite the absence of a negative morpheme in the construction. The main theoretical issues that have attracted a significant interest in the literature are related to the nature of this inference, namely whether the inference is a presupposition, an entailment or an implicature, and the role of morphosyntactic elements which contribute to the counterfactual meaning.

Iatridou (2000, 2009) argued that counterfactuality is conveyed via certain morphosyntactic forms and more specifically, that the combinations of tense, mood and aspect morphemes introduce the meaning of events in the non-actual world. According to Iatridou (2000, 2009) the past tense morphology which appears in many languages in PastCFs is a ‘fake’ past in the sense that the past morpheme is not used in order to denote an event that happened in the past as in declarative sentences (after all the event did not take place); instead the past tense morphology functions as a particular relation (exclusion) between a topic time/world and the time/world of the utterance. Simplifying, this exclusion feature intervenes between the worlds we can talk about (Topic Worlds) and the worlds of the speaker (Actual worlds) by making the Topic world not being part of the Actual world.

From the acquisition perspective, PastCFs pose several interesting challenges. First, being counterfactual, their correct understanding requires a general cognitive ability to entertain counterfactual worlds. As discussed in Section 4, this ability appears to be subject to developmental maturation. Second, given that the past perfect form by itself (i.e., when not embedded under a conditional) asserts that an event has happened in the past in the actual world, the question under investigation in our study is whether children, have acquired that if + past perfect is counterfactual.

² Together with PastCFs, counterfactual meaning is also available in Present Counterfactuals (PresCF) such as ‘If Amy ate the apple, she would win the medal’. More generally, the meaning of PresCFs extends to the present time, leaving open the window to an agnostic view, i.e. that Amy has not (maybe not yet) eaten the apple, but she may still have the option to do so at the present time. The current study focusses specifically on PastCFs, leaving PresCFs for future research.
3. Previous studies on development of counterfactual reasoning and counterfactual language

In this section we review existing developmental psychology literature on the development of counterfactual reasoning in children. This literature is relevant in two ways. First, we want to know whether children are able to entertain counterfactual worlds from birth or whether this ability develops later, and if so, when. In this regards, counterfactual reasoning abilities are often compared to the ability to think conditionally about future events. Second, developmental psychology studies on counterfactual reasoning sometimes used PastCF structures as a measure of the child’s counterfactual reasoning. To this end, we can get some potentially useful data on how children perform on PastCFs specifically.

Counterfactual reasoning ability in young children has attracted much attention from developmental psychologists. Questions that have been asked most commonly are (i) when children become fully adultlike in their ability to think counterfactually, and (ii) which cognitive processes are involved in counterfactual reasoning. To date, both questions have not been fully resolved. Some studies (Harris, German & Mills 1996, German & Nichols 2003) have argued that children can think counterfactually as early as at the age of 3. Other, mostly later studies, argued that counterfactual thinking does not seem to be adultlike in 3- and even 4-year olds (Beck & Guthrie 2011, Rafetseder & Perner 2010, Beck, Riggs & Gorniak 2009, Riggs, Peterson, Robinson & Mitchell 1998). For example, in Riggs et al.’s (1998) study, the children watched an act-out story in which Sally and Peter were in their house. Peter wasn’t feeling very well and so he went to bed. Sally then went to the shops to get some medicine and while she was there, Peter got a call from the post office and was asked to come and help put out a fire. Peter got out of bed and went to the post office to help. Children were asked “If there had been no fire, where would Peter be?” Of 32 children aged between 3;5 and 4;4, only 17 children responded to the question correctly. Similarly, Beck, Riggs and Gorniak (2009) used two similar act-out stories with a sample of 93 children aged 3;3–4;5. In the Piglet story, for example, Piglet was drawing a picture at a table when the wind blew the picture into a tree. Children then had to answer the past counterfactual question “What if the wind hadn’t blown, where would the picture be?” Children gave an incorrect response on 55% (102/186) trials.

It has also been shown that the difficulty of past counterfactual judgements is not linked to the conditional aspect: 3- and 4-year olds who succeed on future conditionals may nevertheless fail on past counterfactuals (Robinson & Beck, 2000; Perner, Sprung & Steinkogler 2004; Beck, Robinson, Carroll & Apperly 2006). For example, Robinson and Beck (2000) acted out a story in which a toy car was positioned on a road with a blue garage at one end, and a red garage at the other end. The car drove along a road to a blue garage. Children’s performance on future conditionals “What if next time he drives the other way, where will he be?” was better than on past counterfactual trials “What if he had driven the other way, where would he be?” Beck (2016) proposes that the asymmetry is due to the fact that only in the past counterfactual scenario the child needs to simultaneously
hold in mind both the state of things in the actual world and the way things could have turned out but did not, which leads high general executive demands (see also Beck & Riggs 2014). Rafetseder and Perner (2010) and also Rafetseder et al. (2013) draw attention on another significant aspect of past counterfactual reasoning, namely, the 'nearest possible world' constraint. Applied to the case of past counterfactuals, the constraint requires that only those possible worlds that differ from the actual one in the features that are causally dependent on a counterfactual antecedent are evaluated, whereas more distant possible worlds are not. According to Rafetseder and Perner (2010), this feature of counterfactuals makes them problematic for children.

4. Our study

Our starting point was the idea that adultlike understanding of PastCFs requires (at least) the following three pieces of cognitive and/or linguistic knowledge:

a) The understanding of conditionals, i.e. that the form ‘if A \(\rightarrow\) then B’ embeds two events A and B into a conditional antecedent – consequence relation.

b) The ability to represent counterfactual worlds, i.e. ability to entertain alternative worlds in which the course of some past events was different to that in the actual world, and

c) knowing that if + past perfect construction denotes counterfactuality, i.e. that if + past perfect, then should be understood as conveying a conditional relationship between the antecedent and the consequent events that did not take place in the actual past.

Whereas points (a) and (b) have been discussed in the previous literature, to the best of our knowledge, point (c) that concerns the linguistic knowledge, has not been raised.3 This is potentially problematic for prior research on the development of conditional reasoning, as children’s failure on past counterfactuals may have been due to their incorrect treatment of the linguistic form rather than their inability to reason counterfactually.

4.1. Participants

All participants spoke English as their first language and were recruited from four fee-paying Kindergarten and Primary schools and one state Primary School

---

3 It is worth mentioning that a general of language in the development of past counterfactuals has been discussed. Beck et al. (2009) reported verbal ability in 3- and 4-year olds (assessed by British Picture Vocabulary Scale Second edition (BPVS); Dunn, Dunn, Whetton & Burley 1997) to be among the strongest predictors of their performance on past counterfactuals. As a possible explanation for this finding, Beck et al (2009, p. 349) mention that “language is likely to be important to cue one to imagine what might have been.”
in London (England). The sample consisted of 74 children (34 girls) aged between 3;1 and 6;11 (mean age = 5;0), of whom 34 children attended a kindergarten and/or a pre-school, 19 children attended Year 1, and 21 children attended Year 2.

4.2. Design and procedure

The main aim of the study was to examine young children’s understanding of PastCFs such as (1). We tested our hypothesis that the linguistic understanding of a grammatical construction if + present perfect is additional to the child’s counterfactual reasoning abilities by investigating the child performance on PastCFs in relation to their performance on future conditionals and to their ability to entertain counterfactual worlds in a false belief task.

Earlier studies on children's counterfactual reasoning (Beck, Riggs, & Burns 2011, Beck et al. 2009, German & Nichols 2003, Riggs et al. 1998) that used if + past perfect forms in a context of a question and required generation of alternative consequences by asking participants to speculate about a counterfactual or to think about an alternative possibility. The present study used if + past perfect as part of a counterfactual statement in order to reduce linguistic and executive functioning demands. More generally, children may be more familiar with or may find it easier to make inferences as a response to a counterfactual statement (Roese 1994).

Throughout the testing session, the child was watching video stories on the experimenter’s laptop together with a hand puppet, manipulated by the experimenter. The overall theme was that of a food contest in which different animals eat healthy or not-so-healthy foods and subsequently are rewarded differentially by the judge according to their food choice. The child’s task was to watch the videos attentively and to answer the puppet’s questions about the story after each video.

Future conditional pretest and training

The study started with 6 trials based on stories that were acted out with animal and hand puppets. The purpose of these trials was twofold: (i) to familiarize the participants with the rules of the food-contest game that was used in the main experiment and the participating animal characters, and (ii) to evaluate children’s understanding of future if-then conditionals.

The children were told that they will be watching videos of a food contest in which different animals will have a choice to eat a healthy snack (e.g. fruit) or a not-so-healthy one (e.g. a sweet). The judge of the contest will award prizes to the animals according to the rules stated in (3) and (4).

(3) If the animal eats a fruit, they will get a medal.
(4) If the animal eats a sweet, they will get a cross.
Note that (3) and (4) are both future conditional statements. In order to assess whether children have understood the rules in (3)-(4), they were asked to respond to six questions, of which three tapped into the child’s understanding of the antecedent [antecedent verification, see (5)] and the other three into the child’s understanding of the consequence [consequence verification, see (6)]:

(5) We want Mr Wolf to get a medal, what does he need to eat?
Correct response: fruit

(6) Mr Wolf is about to eat an apple, what will the judge give him for that?
Correct response: medal

Given our interest in understanding of PastCFs, we considered the understanding of future conditionals a prerequisite. Hence, a score of at least 4 correct trials out of 6 trials was required for the child to proceed to the next stage of the study.

**Past counterfactual task**

As in the pretest, the critical experiments that tested the child’s understanding of PastCFs used the context of an eating contest with two animal contestants and a judge. In each video story, the animals were of the same gender as indicated by them wearing either a black top hat and a bow tie (male) or a pink head piece (female). The male animal pairs were Mr Dog and Mr Pig, Mr Penguin and Mr Frog, Mr Hedgehog and Mr Rabbit, and Mr Monkey and Mr Swan or the female pairs were Mrs Bear and Mrs Cat, Mrs Goose and Mrs Mouse, Mrs Owl and Mrs Fox, and Mrs Turtle and Mrs Hen. Figure 1 presents a sample story with Mr Penguin and Mr Frog as contestants. The story started by showing the choices of food for the animals (slices of watermelon and ice cream cones), the judge introducing the contestants (“In this story we have Mr Dog and Mr Pig”) and reminding them the rule of the game (“If you eat a watermelon you will get a medal. If you eat an ice cream you will get a cross.”; Figure 1a).

The animals then in turn considered both foods and chose their preferred food, i.e. Mr Penguin ate an ice cream and Mr Frog ate a watermelon. They each then received an award from the judge (Figure 1b). This last scene was paused as a visual reminder of the last event that took place in the story.
Figure 1. Stills from the video story from the Past Counterfactual task showing (a) the opening scene, including competitors Mr Penguin and Mr Frog, the foods and the judge with awards, (b) the final scene in which Mr Penguin is wearing a cross for having eaten an ice cream and Mr Frog is wearing a medal for having eaten a watermelon.

After the video, the puppet described what happened in the story using a PastCF such as (7a) or (7b), and the child had to respond or point which animal the sentence was about (only one of (7a,b) was asked on each trial, in a counterbalanced order):

(7) a. If he had eaten a watermelon, he would have got the medal.  
   Correct response: Mr Penguin  
   b. If he had eaten an ice cream, he would have got the cross.  
   Correct response: Mr Frog

Thus, the correct response required the child to choose the animal other than the one matched the surface sentence form, i.e., in (7a) NOT the animal who ate the watermelon and got the medal. The child’s response thus made it possible to infer whether the child was aware of the counterfactuality of the PastCF constructions.

False belief task

The false belief reasoning task was based on a classic false belief task (Wimmer & Perner, 1983) and involved an unexpected transfer of a physical object. The task was administered in order to assess the child ability to entertain possible worlds that contradict the actual world.

In order to provide continuity with the preceding stories, two animals (e.g. Mr Dog and Mr Pig) were chosen as the protagonists. The child and the puppet watched Mr Dog place the chocolate in the basket and then leave the scene. Then Mr Pig moved the chocolate from the basket into the tin. Subsequently, Mr Dog came back to the scene and the child was asked the false belief question in (8):

(8) False Belief Question: Where will Mr Dog look for the chocolate?
The correct answer (‘in the basket’) requires ability to entertain possible worlds that are distinct from the actual world in which the chocolate is in the tin. To ensure that children had followed the story, they were asked two control questions (9)-(10):

(9) **Current Reality Question**: Where is the chocolate now?
(10) **Memory Question**: Where was the chocolate in the beginning?

It was assumed that if children were able to remember the initial or current location, they had understood the story, hence their performance on the false belief question reflects their abilities to entertain possible worlds.

4.3. Results and discussion

**Future conditional pretest**

From the initial sample of 74 children, six children aged 3;1 – 4;4 did not pass the correct response threshold of 4/6 trials on future counterfactuals and were excluded from the analysis. The remaining 68 children (30 girls and 38 boys) gave accurate responses on at least 4 out of 6 trials; their performance on antecedent verification trials and consequence verification trials was similar (99% correct and 95% correct respectively). These 68 children therefore correctly understand future conditionals; their performance on PastCFs is discussed next.

**Past counterfactual task**

Of 68 children who qualified in the task above, 10 children did not complete all 8 trials of the counterfactual task and were excluded from the analysis. The results from the remaining 58 children are shown in Figure 2, split by age.

![Figure 2. Performance on Past Counterfactual task (n = 58), split by age.](image)

As Figure 2 suggests, the 3-year olds (n = 15) had clear difficulties in judging PastCFs; they were correct in only 23% of trials. The correct response rate increases to 86% in the 4-year-old and to 92% and 99% in 5- and 6-year olds (n=17 and n=18) respectively.
A closer look at the individual level performance reveals a strong bimodal distribution whereby each child either shows a very high or a very low performance on PastCFs. This is shown in Table 1 in which all children are classified into 3 groups: those giving at most 1/8 correct answer (i.e. non-adultlike), those giving at least 7/8 correct answers (i.e. adultlike) and others.

Table 1. Individual subject performance on past counterfactual task. Number of children in each age group is indicated in brackets.

<table>
<thead>
<tr>
<th>Age group</th>
<th>Number of correct response (out of 8)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-1</td>
</tr>
<tr>
<td>3-year olds</td>
<td>11</td>
</tr>
<tr>
<td>(n = 15)</td>
<td></td>
</tr>
<tr>
<td>4-year olds</td>
<td>2</td>
</tr>
<tr>
<td>(n = 8)</td>
<td></td>
</tr>
<tr>
<td>5-year olds</td>
<td>0</td>
</tr>
<tr>
<td>(n = 17)</td>
<td></td>
</tr>
<tr>
<td>6-year olds</td>
<td>0</td>
</tr>
<tr>
<td>(n = 18)</td>
<td></td>
</tr>
</tbody>
</table>

Several points are noteworthy. As Table 1 shows, children demonstrate predominantly non-adultlike performance (0-1 correct trials) or predominantly adultlike performance (7-8 correct trials), while ‘other’ category corresponding to midway performance is limited to one child. Thus, 57 out of 58 children demonstrated a strongly consistent non-adultlike or adultlike performance on PastCFs; the consistency of the child’s behavior clear points to a step-like (as opposed to continuous) learning trajectory for past counterfactuals. The non-adultlike pattern was restricted to 13 children in younger age groups, i.e. 3- and 4-year olds (although other children from the same age groups have showed a clear adultlike pattern). This finding demonstrates that PastCF constructions are not correctly understood from the outset; their meaning is subject to discovery as part of linguistic (and general cognitive) development. However, it is important to know whether the failure on PastCFs in these 13 children is specifically due to their unawareness of the linguistic meaning of PastCFs or due to a more general cognitive limitation on entertaining possible worlds other than the actual world. We thus turn to the results of the False Belief task.

**False belief task**

As shown in Table 2, the correct response rates were 87% for the 3-year olds, 100% for the 4-year olds, 76% of the 5-year olds and 100% for the 6-year olds. All children were able to remember the initial or current location of the object.
and pass the memory control questions, including most of the 3-year old children who answered these control questions correctly.

**Table 2. Performance on the False Belief task: number of children providing a correct response (‘basket’) or incorrect response (‘tin’) to the false belief question.**

<table>
<thead>
<tr>
<th>Age group</th>
<th>Correct</th>
<th>Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-year olds</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td>(n = 15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-year olds</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>(n = 8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-year olds</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>(n = 17)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-year olds</td>
<td>18</td>
<td>0</td>
</tr>
<tr>
<td>(n = 18)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The critical question was whether there is a correlation between child’s performance on counterfactual vs. false belief task. Point biserial analysis was conducted and showed no significant interaction in the performance on the two tasks ($r=0.189$). Cumulatively, our findings suggest that children’s correct performance on the false belief tasks (which we interpret ability to entertain non-actual worlds) is not necessarily accompanied by an adultlike performance on PastCFs. We conclude that younger children’s errors on PastCFs stem from the children’s lack of linguistic awareness that these constructions are counterfactual rather than difficulty with counterfactual reasoning per se.

5. General Discussion

The main finding of the study comes from a group of 13 younger children who understand simple *if-then* conditionals (as shown by their success on the pre-test), can represent counterfactual possible worlds (as suggested by their successful performance on the false belief task), yet fail on counterfactual conditionals. Given that conditional reasoning and ability to represent counterfactual possible worlds are two conceptual pre-requisites for counterfactual conditionals, the children’s failure cannot be due to cognitive immaturity. Instead, we take this pattern to indicate that these children have not yet learnt that the specific grammatical construction, i.e. *if* + *past perfect*, encodes counterfactuality.

The relevant issue is then to elucidate the trajectory for discovering that *if* + *past perfect* in English encodes counterfactuality. We propose that this grammatical construction, alongside other verbs and constructions that encode modal meanings (the term ‘modal’ is used in a broad sense and encompasses any
aspects of semantics that goes beyond the actual world, including counterfactualitiy), are subject to the Actuality bias in language learning. According to the Actuality bias, children initially construct the semantics of a lexical or grammatical category without appealing to non-actual worlds, while modal meanings emerge only if the original hypothesis mismatches the evidence (Kazanina, Seddon & Baker, accepted). The Actuality Bias is essential in that it makes it possible to learn the semantics for verbs or categories that require evaluation in a non-actual world on the basis of positive evidence, without reliance on negative evidence. For example, in the case of counterfactuals the child initially treats expressions such as “if she/they had eaten dinner” to express conditional meaning in the (distant) past. However, once the child is exposed to a sentence like “If you had eaten your dinner, I would have bought you an ice cream” in a scenario where the child did not eat the dinner and did not receive an ice cream (and is perfectly aware of both facts), that experience would point them towards realizing that “if you had eaten a dinner” must encode a hidden NOT and refers to the situation that is opposite to the actual state of events. Note that once a grammatical construction – here if + past perfect – is concluded to encode modality, it is considered as such across the board. This prototypical all-or-nothing behavior that is typical of rule learning is clearly demonstrated in our results: most children either consistently fail to treat if + past perfect as counterfactual or consistently succeed, with almost no midway performers.

The idea that child language development may be subject to a Reality bias was discussed in the previous literature by Jill de Villiers (2005) in relation to verbs such as want, say or think that take a syntactic complement clause. These verbs all have complements of an ‘irrealis’ type in that the sentence with any of these verbs as a main verb may be true even though the proposition of the complement clause is false. For example, John says that it is sunny outside can be true even if it was sunny outside is false. De Villiers argues that children start off with all verb meanings being realis, i.e. connected to ongoing events in the actual world, with cases of irrealis ‘branching off’ as evidence for the latter becomes available. Kazanina et al. (accepted) proposed a similar developmental trajectory for sublexicaly modal ditransitive transfer verbs such as send and throw. As mentioned above, the benefit of such an approach is that it provides an account for how irrealis/modal meanings develop while ensuring the lack of spurious overgeneralization in the lexicon and grammar: all realis meanings remain such as there is no positive evidence to the opposite.

References

Beck, Sarah, & Riggs, Kevin (2014). Developing thoughts about what might have been. Child Development Perspectives, 8(3), 175–179.


Iatridou, Sabine (2009). About the Imperfective in Counterfactuals, handout from Yale Imperfective Workshop, April 10-11, 2009


