The influence of linguistic temporal organization on children’s understanding of temporal terms and concepts

Laura Wagner
Ohio State University

The Typological Prevalence Hypothesis argues that the cognitive biases that lead to common cross-linguistic patterns also guide language development. Two studies investigated English learning children’s understanding of temporal terms and concepts and how they might be influenced by three typologically common patterns: the use of a deictic center, the asymmetry between past and future times, and the use of temporal remoteness. Using a timeline task, five- and eight-year-old children were asked to locate events with a range of temporal expressions as well as with general event descriptions. The results provided only limited support for the TPH for the temporal domain but do demonstrate the potential for the hypothesis as a research approach.

Keywords: tense, time, timeline, temporal semantics, typology

1. Introduction

This paper investigates children’s developing understanding of temporal terms in English and the concepts that organize linguistic temporal systems. The approach adopted is inspired by the Typological Prevalence Hypothesis (TPH) of Gentner & Bowerman (2009, see also Slobin, 1985). This hypothesis states that, all things being equal, if a particular semantic element is commonly grammaticized across a diverse set of the world’s languages (i.e., it is typologically prevalent) then it will be comparatively easy for children to learn. The intuition is that those concepts which humans are cognitively pre-disposed to learn will not only shape the local learning processes of individual children, but will also shape larger linguistic patterns around the globe. Of course, in the real world, not all things are equal, and so the hypothesis is not couched in terms of absolute universals or necessities. Nevertheless, the TPH serves as a useful tool for researchers by suggesting what the learning biases may be
Laura Wagner

for children acquiring a language and therefore, the kinds of meaning components that might influence children’s interpretations – even components that are not directly encoded in the child’s target grammar.

Temporality is one of the handful of conceptual domains that is reliably grammaticized by languages, although the specific encodings vary dramatically from language to language (see Dahl, 1985). Moreover, temporality is a multi-layered domain involving event structure, perspective, aspectual information, and linear ordering among other things (see Binnick, 2012). The breadth of this domain is too large for the scope of this paper, and the current experiments will focus exclusively on tense, that is, on the specification of the past, present, and future. More specifically, they will focus on three dimensions of tense meaning that are commonly grammaticized within a range of languages and therefore good candidates for the TPH. These dimensions are the deictic center, an asymmetry between the past and future, and the encoding of temporal remoteness (or distance) from the deictic center.

Most formal accounts organize tenses around a deictic center (cf. Reichenbach, 1947; Binnick, 1991; Klein, 1994, inter alia); usually that center is the speaker’s time of utterance, which is also often linked to the present moment. Past tenses refer to times occurring before the time of utterance and future tenses refer to times occurring after the time of utterance. This type of organization is not the only way one can define time. For example, many philosophers (and some physicists) have argued that time is better understood without the deictic reference point, as a simple linear sequence of moments with no special status for the speaker’s role within it (see, for example, Ludlow, 1999). Thus, while using a deictic center is not a necessary way of encoding time, it is a common one across languages. In addition, given that the current work will focus on children acquiring English, it is worth noting that the deictic center is part of how English grammaticizes tense.

An asymmetry between the past and future manifests itself across the world’s languages in a variety of ways. Most notably future markers in many languages, including English, are homophonous with markers of present tense and modality (Comrie, 1985; Deo, 2012). In English, there are no verb auxiliaries or morphological markers that exclusively express future meaning: all markers used for the future are also, synchronically, used for other functions such as present tense (the progressive go) or modality (will, shall). By contrast the English past tense has distinctive forms (was, -ed). And, while it is true that English’s simple past tense form conveys perfective aspectual meaning in addition to tense meaning, not all past tense forms do so (e.g., was does not).

Temporal remoteness, where remoteness is calculated relative to the deictic center, is also a common feature of many linguistic tense systems (Dahl, 1985; Botne, 2012). For example, Washo (a native American language in the Hokan
family) grammatically marks not only whether an event happened in the past, but whether it happened within the past day (hodiernal -leg), previous to yesterday but not too far away (-ayʔ), within the speaker’s memory (-gul), or beyond the speaker’s memory (-lul) (Botne, 2012). Remoteness distinctions are found in both the past and future tenses, although they are more common in the past tenses (Deo, 2012). Note that for languages with remoteness marking, the specification is obligatory. While a language like English certainly allows one to provide a precise temporal distance (e.g., I wrote this sentence at 4:23 on October 18th), those distances are not grammaticized in the language and are wholly optional; English does not directly code remoteness within the grammatical system.

These three temporal elements, therefore, pass the first hurdle for the TPH: they are all well described as being part of the grammatical systems in many of the world’s languages. The next step is to consider whether any corroborating evidence exists within the cognitive domain. By hypothesis, these elements are prevalent because they reflect something “natural” about how we build these concepts and so they should matter for children’s learning trajectories (and potentially also for adults’ representations).

The strongest support for cognitive naturalness for any of these three temporal concepts can be found for the idea that the past and future are asymmetric and further, that the future is less well specified. For example, McCormack and Hanley (2011) found that pre-school aged children were better at reasoning about the past than the future, and Busby Grant and Suddendorf (2009) similarly found that three- and four-year-olds are more accurate at locating events on a timeline for past times than future times. There is more limited support for the cognitive underpinnings of the other typological elements considered here. Suddendorf (2010) reviews ERP research that found that adult brains show more activation when reasoning about distal (i.e., remote) events than close ones. He argues that it takes more effort to think about things that are not close to the immediate moment. Moreover, there is also a long tradition going back to Piaget (1969) suggesting that children initially reason better about the here-and-now (the present) than they do about more distant time periods. With respect to the deictic center, there is some evidence that children as old as nine years may have difficulties using it when reasoning about times. In various studies by Friedman (2002; Friedman, Garder & Zubin, 1995), children located events along a road that corresponded to a timeline. A subset of the children – including some 9-year-olds – interpreted the road in a non-deictic fashion even after being explicitly told not to. These data, however, are difficult to interpret as the critical cases involved cyclical daily events such as having breakfast. The non-deictic responses arose when children put breakfast at the beginning of the road (since it starts the day) instead of much further on (since children were tested mid-day and their next breakfast was in the future). Cyclic reasoning may
be especially difficult to do, but it is something that children must connect to their deictically organized tense system; the fact that this integration takes years to accomplish suggests some dimension of this integration is quite difficult.

Overall, the existing work on temporal cognition does not provide particularly strong evidence for the TPH and raises the possibility that the causal direction between cognitive organization and language development goes the other way. An alternative to the TPH is Slobin’s “thinking for speaking” hypothesis (e.g., Slobin, 1996). From this point of view, children are initially more open-minded about how to conceptualize domains such as time, but through the process of learning their native language they develop habits for how to think about the domain. This position predicts that influences should be restricted to elements children are exposed to (i.e., that are grammaticized in their native language). Moreover, it predicts that these deep organizational influences will take time to be felt: it could take years before a child’s target language organization has a strong influence on their habits of cognitive organization.

One final literature to consider is work looking explicitly at children’s acquisition of tense meanings. Temporal morphology is among the earliest produced by children (e.g., Brown, 1973) and its use appears to be coordinated with systematic grammatical rules, as has been discussed in the Optional Infinitive debates (see Poeppe & Wexler, 1993). Studies examining children’s ability to assign meaning to tense markers suggest that children begin to make language specific mappings before the age of 3 years (Valian, 2006) and that they have largely completed their acquisition of grammatical tense by the age of 5 years across a variety of different languages (Wagner, 2012; Maastricht & Hollebrandse, 2011; Weist, Atanassova, Wysocka, & Pawlak, 1999). All of the languages examined in detail have been organized around a deictic center, and so these data support the idea that children can learn such tense systems in this time period. Interestingly, a recurring theme of the literature has been that children’s performance is often facilitated by the inclusion of a future tense form in the task (Weist, Wysocka & Lyytenin, 1991; Weist et al., 1999; Wagner, 2001; Valian, 2006) suggesting not only that the future may have a distinctive status in children’s early grammatical systems, but that it may be easier to acquire; early facility with future forms is not obviously predicted by the typological or cognitive results noted above but it could be a consequence of the less differentiated nature of the future – if there’s less specificity to learn, it may be easier to accomplish. No studies have explicitly considered how temporal remoteness might influence children’s early tense understanding but this is hardly surprising given that none of the languages closely investigated have grammaticized remoteness.

The current pair of studies investigated the TPH by asking whether children’s temporal understanding was influenced by typologically common patterns even when those patterns are not actually grammaticized in their own language. A
The timeline task was used in which children were asked to locate events along a line with a pre-established deictic center. This task allows children to demonstrate their understanding of three typologically prevalent temporal properties. The use of a deictic center to organize temporal interpretations (a property that is grammaticized in the participants’ native language of English) would be shown by correctly placing events relative to the established center. An asymmetry between the past and future (a property that is debatably grammaticized in English) would be shown by more correct event placements, either in isolation or in ordered pairs, in one timezone (past or future) than the other. An influence of remoteness (a property that is not grammaticized in English) would be shown by a differentiation in performance as a function of how distant from the present time the event was/will be.

The timeline task is not without limitations. Children in the younger age group (5-year-olds) have relatively little experience with timelines and this necessitated a brief training session. Children in the older age group (8-year-olds) by contrast have likely received some amount of explicit training with timelines in a school setting. As shall be shown, these issues are of little practical importance: the younger children did perform coherently on the task in general, making performance differences between ages meaningful. One additional issue concerns the radically different size of the temporal intervals being mapped onto lines of identical length: for example, across items, children located events from the recent past (e.g., eating breakfast), the distant past (e.g., being a baby), and the mythic past (e.g., when dinosaurs were on earth). Children were not provided with any instructions about how to handle the range problem across items, however, a very conservative coding strategy that simply considered regions of the line (rather than absolute distances) was adopted to help counteract differences in range interpretations across items.

To assess children’s temporal interpretations with elements outside of the grammaticized tense system of English, children were tested with event descriptions that required the child to make knowledge-based inferences about when the events occurred, and they were also tested on a range of temporal adverbials. The adverbial forms were all checked in the CHILDES database and all appeared in the speech of multiple children. All the events used were chosen to reflect events children would be reasonably familiar with (e.g., driving, being a baby) or had stereotypical associations with (e.g., science fiction scenarios, famous historical periods). Informal piloting confirmed that children had relevant associations to the events used.

A possible concern is that the distal future elements all had a science-fiction aspect to them and suggested it might invoke a fantasy element for these items. And indeed, all the pictures were chosen to encourage children to see the events as very unlike the current day (e.g., the future robot cleaning one’s home was not depicted as a Roomba, but more like Rosie of the Jetsons). The future is inherently irrealis – by definition, it does involve possible worlds and an imaginative sense the
past does not. To that extent, this problem is inherent in what it means to be part of the future. However, in the analyses, the more mythic times were collapsed with more “realistic” future/past times (being a baby, being a mom/dad) and inspection showed that they did not pattern substantially differently from each other.

The dimension of language that the TPH is generally thought to depend on is grammatical: grammars functionally organize a language to a much greater extent than open class words within a lexicon. Grammatical tense systems clearly fall within this dimension, and general descriptions of events clearly do not. Temporal adverbs, however, potentially occupy an intermediate space. Investigations of children’s understanding of temporal adverbs suggests that preschoolers understand that they convey some kind of temporal information – but that it takes until children are around 7 to 8 years old before most individual items are correctly mapped onto specific temporal concepts (Shatz, Tare, Nguyen, & Young, 2010; Tillman & Barner, 2015; but see Busby Grant & Suddendorf, 2011). This shift from a more general temporal grouping to specific temporal understandings may explain why temporal adverbs have inconsistently influenced how pre-school aged children interpret grammaticized tense marking (Wagner, 2001; Valian, 2006). Within the current context, the shift may also reflect development in children’s understanding about exactly which temporal concepts are in fact being grammaticized in their language. Adverbial expressions do permit a much wider range of temporal points to be targeted (including points not grammaticized by the children’s target language); however, given their intermediate status, adverbs will be analyzed separately from the event descriptions whose temporal status depends on knowledge-based inferences.

If the TPH is grounded in cognition itself, then any effects of common patterns should arise regardless of whether children are tested with explicit linguistic expressions about time (the adverbials) or with general event descriptions. However, it is also possible that the effects of linguistic typology are restricted to linguistic items; if so, then the effects of the more common patterns may be seen more strongly (or perhaps only) when children are tested with linguistic encodings of time. By contrast, a “thinking for speaking” approach would more easily predict that the presence of explicit linguistic elements would better foster habits that are grounded in speech. More generally, the “thinking for speaking” position predicts that children should not be influenced by features not grammaticized in their target language (like remoteness) and in general, any effects of language structure should be stronger as children get older and have more developed habits.
2. Experiment 1

2.1 Participants

Participants were tested in a local science center. The Younger group consisted of 24 children ($M = 5;7, 9$ girls) and the Older group consisted of 10 children ($M = 8;5, 8$ girls). All children had English as their sole or dominant language.

2.2 Stimuli

The task was performed using a timeline consisting of a line approximately 7 inches long with a large star in the middle (approximately 1 inch wide). Each trial used a different line, so children always started with a fresh, blank timeline. Children marked times by placing pictures (~1 inch wide) on the line. The pictures were described in different ways depending on the phase of the task – in the Explicit Linguistic phase, the picture depicted a potentially recurring event (e.g., having a birthday party, playing with Lego blocks) combined with a specific temporal descriptor (such as last year). In the General Event Description phase, the pictures depicted events that were naturally anchored in time. The events were described in terms of their content (e.g., “imagine this is a picture of you as a baby”) but no specific temporal descriptors were added. That said, it frequently happened that the experimenter added appropriate tense forms on the verb (“when you were a baby”) so there was a limited amount of linguistic tense that occurred as well. The full set of items and their descriptors are shown in Figure 1.

The target items ranged over the past, present and future times, reflecting the three possible relations to a deictic center. Moreover, for the past and future times, the items also ranged over levels of remoteness. Three remoteness divisions were used. Very Close times (within a matter of minutes or hours) were used to target a potential “vast now” period that might include quite loose boundaries of the present time; Close times (within a matter of days or the week), and Far times (including the span of a person’s lifetime as well as the mythic past and future) were chosen to link to distances that are grammaticized in other languages (but not in English). Note that typologically, many languages that grammaticize remoteness also combine those levels of remoteness into a single marker. However, because the span of a lifetime and mythic times are conceptually different, the data was inspected to insure that the effects were not being driven primarily by one or the other of these distances. In addition, one Unspecified level of remoteness was used for a subset of items in the Explicit Linguistic phase that were described with markers that signaled a deictic relationship but no specific temporal remoteness (e.g., plain tenses).
Explicit Linguistic Adverbs/Tenses

<table>
<thead>
<tr>
<th>Past Items</th>
<th>Present Items</th>
<th>Future Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did (unspecified)</td>
<td>E1</td>
<td>Is doing</td>
</tr>
<tr>
<td>Before (unspecified)</td>
<td>E2</td>
<td>Today</td>
</tr>
<tr>
<td>Just (very close)</td>
<td>E1</td>
<td>This week</td>
</tr>
<tr>
<td>A little bit ago (very close)</td>
<td>E2</td>
<td>This year</td>
</tr>
<tr>
<td>Yesterday (close)</td>
<td>E1</td>
<td>Tomorrow (close)</td>
</tr>
<tr>
<td>Last week (close)</td>
<td>E2</td>
<td>Next week (close)</td>
</tr>
<tr>
<td>Last year (far)</td>
<td>E1</td>
<td>Next year (far)</td>
</tr>
</tbody>
</table>

Naturally anchored general event descriptions

Note that only a single Present Time item was used in both experiments: When you are sitting with me.

<table>
<thead>
<tr>
<th>Past Items</th>
<th>Future Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>When you came to &lt;museum space&gt; (very close)</td>
<td>When you leave &lt;museum space&gt; (very close)</td>
</tr>
<tr>
<td>When you drove here (close)</td>
<td>When you drive home (close)</td>
</tr>
<tr>
<td>When you ate breakfast (close)</td>
<td>When you go to bed (close)</td>
</tr>
<tr>
<td>When you wore diapers (far)</td>
<td>When you are a mom or dad yourself (far)</td>
</tr>
<tr>
<td>When people wore togas (far)</td>
<td>When people drive spaceships to work (far)</td>
</tr>
<tr>
<td>When dinosaurs were on earth (far)</td>
<td>When robots clean our homes (far)</td>
</tr>
</tbody>
</table>

Figure 1. Stimuli used across both experiments. The temporal remoteness classification for past and future items is indicated in parentheses; all present items were treated as equivalently remote from the current time. The experiment in which the item was used is indicated after each item. For the naturally anchored event descriptions, some items were placed singly (single) and others were placed as parts of pairs (pair).

2.3 Procedures

All sessions began with a brief practice session about how to use the timeline. Children were told that the star represented right now, the region to the left of the star was where things have happened already and the region after the star was later than now (note, specific adverbs were chosen based on Wagner, 2001, as well as informal piloting). Children placed three pictures and received corrective feedback. No additional feedback was given after the practice phase.
Testing began with the Explicit Linguistic phase, in which children were provided with explicit temporal specifications to describe recurring events. Children were told explicitly that the events recurred (e.g., “This is something the boy does a lot”) but the event itself was not described. Children also received one recurring item with plain tense marking, but the tense bearing element received strong prosodic stress (e.g., “Show me where it [the picture] goes for when the boy IS doing it”) to encourage attention to the tense per se. Children located the recurring events multiple times (e.g., “Show me where it goes for last year/this year/next year”), but they received a new timeline for each picture placement. Children placed pictures for 12 items in this phase.

Testing continued with the General Event Description phase which had two parts. First, children were provided with a single picture and a description of it that implied a particular time (e.g., people don’t currently wear togas, nor can the children being tested currently drive their own cars). Children were asked to place the picture on the timeline where it was supposed to go. Next, children were given pairs of pictures and were provided with descriptions for each. They were asked to place both pictures on the timeline. Children placed 5 single pictures and 4 pairs of pictures.

2.4 Coding

For all experimental phases, the experimenter marked on the line where children located the picture – the center of the picture was used as the chosen location regardless of orientation. To code the time markings, a clear plastic template was made that broke the line down into one-inch regions; these regions were given numerical ranks with –3 being the left-most (or most distantly past) region, 0 being the center star (or present time) and +3 being the right-most (or most distantly future) region. The template was laid over the child’s line and the location of the child’s picture placement was given a numerical score corresponding to its region. Placements on the border between two regions were given intermediate credit (e.g., 1.5 indicated that the picture’s midpoint was on the line between regions 1 and 2).

In effect, this coding scheme gives a basic temporal rank ordering to children’s temporal placements, where the size of the number signals remoteness from the present, and the valence signals general time (negative numbers are past, positive numbers are future, zero is present).
2.5 Results

2.5.1 Ability to use the timeline appropriately

The first analysis consisted of a validity check: do children understand how to use a timeline in an appropriate fashion at all? The essential feature of a timeline is that spatial distance should correspond to temporal distance. To check whether children were sensitive to this overall fact, all the pictures used in the task were ordered temporally from past to future; events happening at the same time (e.g., at the present time) were given the same rank. Temporally unspecified items (e.g., plain tenses) were ranked as very close to the present time. The ordering was established by discussion among the author and the two coders, but there was little disagreement. This ordering was treated as the “objectively true” times for each event. Children’s temporal choices for each picture were determined by where they were located on the timeline. Significant positive correlations were found between the objectively true ordering and the ordering on the timeline for all the items placed by the Younger age group (Pearson’s $r = .26$, $N = 572$, $p < .001$) and by the Older age group (Pearson’s $r = .71$, $N = 237$, $p < .001$). Figure 2 shows the correlations for the children in both Experiments 1 and 2. On the whole, these data show that children are indeed capable of translating their temporal understanding into this spatial representation.

![Figure 2. Children’s use of the timeline](image-url)

The Objective Time Ordering is the true temporal ordering of the items used with lower scores being further in the past. The Average Timeline Region Score is the location on the timeline where children placed that item with lower scores being further in the past (zero represents the present timezone). All groups of children showed significant positive correlations, reflecting veridical use of the timeline. Filled-in symbols/solid lines represent children from Experiment 1 and Open symbols/dashed lines represent children from Experiment 2; Squares/grey lines represent Older children (8-year-olds) and Triangles/black lines represent Younger children (5-year-olds).
2.5.2 Analysis of effects
The next set of analyses examined whether children understood time as organized around a deictic center. Each item was categorized as picking out a time in the past, the present, or the future (for the paired event descriptions, each half of the pair was treated as a separate item in this analysis); children’s answers were coded as being on the central star (associated with the present), to the left of the star (associated with the past) or to the right of the star (associated with the future). Overall rates of correctness for all three timezones were very high for both age groups, and most scores were above chance levels (see Table 1). A repeated measures ANOVA was conducted using age group as a between-subjects variable, cue phase (Explicit Linguistic vs. General Event Descriptions) and target timezone (Past, Present, and Future) as the independent variables, and proportion correct as the dependent variable. A main effect was found for age group ($F(1, 32) = 27.12, p < .001$, partial $\eta^2 = .46$) as the Older children out-performed Younger children in general, but there were no significant interactions with age group as the overall patterns of ease and difficulty were the same across ages. No main effects were found between the Explicit Linguistic items and General Event Descriptions ($F(1, 32) = 1.56, p = .22$) or for the three timezones ($F(2, 31) = 0.84, p = 0.49$), but there was a significant interaction between those two variables ($F(2, 31) = 4.11, p = .026$, partial $\eta^2 = .21$). The interaction stemmed from the fact that while there was a slight (non-significant) advantage favoring the General Event Descriptions over the Explicit Linguistic cues for items in the past and present, but for the future times, there was a significant advantage favoring the items with Explicit Linguistic cues ($M = .69$) over those with General Event Descriptions ($M = .56$, $t(33) = 2.35, p < .025$). Thus, overall children are highly accurate in this task, and their ability to reason temporally about the future was facilitated by having explicit linguistic direction.

The paired Event Descriptions permitted two further analyses: First, did children do better when the two items within the pair were in different timezones? To the extent that the deictic center organizes temporal organization, it was predicted that children would do better with pairs containing one item from the past and one from the future relative to two past items or two future items. A repeated measures ANOVA was conducted with the independent variables of age group and time-pairings (Both Past, Both Future, One Past & One Future). Results showed that there was improvement with age ($F(1, 32) = 12.69, p < .001$, partial $\eta^2 = .28$) but no interaction between age group and time pairings. However, there was a significant main effect of type of time pairing ($F(2, 32) = 9.98, p < .001$, partial $\eta^2 = .24$). Post-hoc paired t-tests showed that children were more likely to be correct when the two items crossed the deictic center with one in the past and one in the future than when both items were in the past ($t(33) = 3.4$, $p = .002$) or both items were in the future ($t(33) = 5.2$, $p < .001$). There was no difference in overall
deictic accuracy (that is, getting the pictures on the correct side of the timeline) between the all-past and all future items ($t(33) = 1.7, p = .10$). Also note that the paired items crossing the deictic center did not all create a larger objective temporal distance between the two items and children performed virtually identically with items that were maximally distal from each other (“people dressed in togas” vs. “when people have robots clean their homes”) and minimally distal from each other (“when you drove here” vs. “when you drove home”). What facilitates children’s performance is having the deictic center temporally located in between the events.

The paired event descriptions also allowed for the examination of children’s ability to correctly order the two events in a linear sequence. For the pairs where one item was in the past and the other was in the future, the timezone accuracy score is equivalent to linear ordering success: if children correctly placed the past item in the past timezone and the future item in the future timezone, then they

Table 1. Summary of accuracy scores for both experiments
The overall accuracy for correct timezone placement for the different conditions across both age groups and both experiments (standard errors in parentheses). Asterisks mark values that are significantly different from chance ($p < .05$), which was set at .33 reflecting the three major timezones of past, present and future. All non-significant results are highlighted. For paired event descriptions, each picture was coded separately.

<table>
<thead>
<tr>
<th>Mean Timezone Accuracy</th>
<th>5-year-olds</th>
<th>8-year-olds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exp 1</td>
<td>Exp 2</td>
</tr>
<tr>
<td><strong>Linguistic Items</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Far</td>
<td>.74 (.09)*</td>
<td>n/a</td>
</tr>
<tr>
<td>Close</td>
<td>.58 (.1)*</td>
<td>.68 (.1)*</td>
</tr>
<tr>
<td>Very Close</td>
<td>.29 (.09)</td>
<td>.36 (.1)</td>
</tr>
<tr>
<td>Unspecified</td>
<td>.67 (.09)*</td>
<td>.55 (.1)</td>
</tr>
<tr>
<td>Present</td>
<td>.51 (.07)*</td>
<td>.45 (.1)</td>
</tr>
<tr>
<td>Future</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very Close</td>
<td>.57 (.1)*</td>
<td>.32 (.1)</td>
</tr>
<tr>
<td>Close</td>
<td>.61 (.1)*</td>
<td>.45 (.1)</td>
</tr>
<tr>
<td>Far</td>
<td>.78 (.09)*</td>
<td>n/a</td>
</tr>
<tr>
<td>Unspecified</td>
<td>.5 (.1)</td>
<td>.45 (.1)</td>
</tr>
<tr>
<td><strong>Event Descriptions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Far</td>
<td>.60 (.06)*</td>
<td>.61 (.07)*</td>
</tr>
<tr>
<td>Close</td>
<td>.71 (.09)*</td>
<td>.72 (.1)*</td>
</tr>
<tr>
<td>Very Close</td>
<td>.52 (.08)*</td>
<td>.72 (.1)*</td>
</tr>
<tr>
<td>Present</td>
<td>.63 (.1)*</td>
<td>.55 (.1)</td>
</tr>
<tr>
<td>Future</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very Close</td>
<td>.52 (.08)*</td>
<td>.95 (.05)*</td>
</tr>
<tr>
<td>Close</td>
<td>.25 (.09)</td>
<td>.55 (.1)</td>
</tr>
<tr>
<td>Far</td>
<td>.56 (.08)*</td>
<td>.57 (.06)*</td>
</tr>
</tbody>
</table>
also by definition provided the correct linear order of the items. This analysis asks whether children differentiated pairs with respect to their linear ordering within a timezone: were children better (or worse) able to sequence two past items compared to two future items? The linear sequence was coded as being correct based on whether the linear order of the items was appropriate (e.g., “When dinosaurs were on earth” should precede “When you wore diapers”), but did not consider the actual distance between the items. A repeated measures ANOVA was conducted with the independent variables of age group and time-pairings (Both Past vs Both Future) and the dependent measure of linear sequencing accuracy. Results showed no effect or interaction with age group, but a significant effect of the time-pairing ($F(1,32) = 7.72, p < .01$, partial $\eta^2 = .194$). Children were more accurate in ordering two past items ($M = .76$) than two future items (.41). Children were better able to make precise judgments about the past than the future.

The final analysis asked how the temporal remoteness of the items influenced children’s performance. A modified version of the initial omnibus ANOVA was conducted. This version also used the independent variables of age group and type of cue (Explicit Linguistic vs. General Event Description) and the dependent variable of timezone accuracy. However, the three remoteness levels were also used as independent variables (Very Close, Close, and Far); moreover, as these levels were only defined for the past and future times, only those times were included as levels in the analysis. In addition, plain tenses (which lack any remoteness value), were excluded from the analysis. Not surprisingly, this analysis reproduced the previously noted main effects of age group (Older children out-performed Younger children) and the interaction between cue type and timezone resulting from the difference between the cue types in the future timezone. In addition, this analysis found a main effect for temporal remoteness ($F(2, 31) = 7.67, p = .002$ partial $\eta^2 = .33$) as well as an interaction between remoteness and timezone ($F(2, 31) = 5.0, p = .013$ partial $\eta^2 = .24$) and a three-way interaction between remoteness, timezone, and cue type ($F(2, 31) = 4.07 p = .027$ partial $\eta^2 = .21$). Overall, the most distal (Far) items were the most accurate, and children performed significantly better with them than either the Close items ($t(33) = 2.3, p = .026$) or the Very Close items ($t(33) = 4.7, p < .001$). However, as the interactions suggest, this effect was driven in large part by the fact that two specific cases were significantly lower than the others (paired $t$-test $p$-values ranged from .001 to 032): children were significantly

---

1. Paired $t$-tests were conducted to see how the plain tense items compared to the linguistic items with more precise temporal specification. The unspecified items received an accuracy score of .68 which was not statistically different for either age group from any level of specified linguistic remoteness.
less accurate with the General Event Descriptions for Close Future times (e.g., when you go to bed) and with Explicit Linguistic items for Very Close Past times (e.g., She just did it) although those two items were not significantly different from each other ($t(33) = 0.3, p = .7$). Although these conditions were the primary drivers of significance, it is worth noting that the overall pattern within both timezones and with both cue types is for increasing accuracy with increasing distance from the present time.

2.6 Discussion

These children showed significant improvement with age: Relative to the 5-year-olds, the 8-year-olds showed more precise use of the timeline (as seen in the higher correlation between their picture placements and the objectively true times), and they were more likely to place items in the correct overall timezone (past, present, or future). This improvement, however, should be seen in the context of 5-year-olds’ overall high rates of success in using the correct overall timezone and significant positive correlation with the objective times of the events. More importantly, it should be seen in the context of the complete absence of interactions of age with the other factors: the influences of the deictic center, event remoteness, and differential ordering in the past and future were qualitatively the same across the age groups. Consistent with the TPH, children may get better at the task, but the influences of the temporal components are already in place for the younger age group.

The data also support several other predictions of the TPH. The children were influenced by the deictic center, as seen by their ability to correctly locate events with respect to it, and by the fact that they were better able to locate pairs of events when the items crossed the deictic center compared to when both were on the same side of it. Children also demonstrated an asymmetry in their past and future understanding, as seen by the fact that they were more likely to get the linear order correct for two paired items in the past than in the future timezone. Most strikingly, even though the children were acquiring a language (English) that does not grammaticize remoteness, the children nevertheless were influenced by how remote the event was from the deictic center. Specifically, the children were more accurate in locating items in the correct overall timezone when the item was linked to a time far from the deictic center than when it was linked to times close or very close to the center. Thus, even though the English grammatical tense system treats near and far past events as equivalent (Mary wore a hat yesterday/1000 years ago) these children – like many other languages of the world – did not.

One piece of data that suggests that temporal organization might be influenced by something more along the lines of the “thinking for speaking” approach is the fact that children’s accuracy in locating future events in the correct general timezone
was better with explicit linguistic temporal adverbials than without them. It is interesting to note, though, that language appears to help in organizing a dimension of time which the TPH suggests would be less well defined through cognitive biases.

3. **Experiment 2**

One potential concern with Experiment 1 is that a limited number of items were used and the effects may depend on those specific words and events rather than the more general features they embodied. Experiment 2 is a conceptual replication of Experiment 1. It uses the same task and procedures, and even a few of the same items. However, it also uses all new Explicit Linguistic items, which will help determine if children’s abilities are linked to the more abstract features of the items. Moreover, in the General Event Description phase, in addition to using a few new items, it also re-arranges how the paired event descriptions were matched up. Experiment 1 found an advantage for pairs in which the items were on different sides of the deictic center, but it is also possible that those pairs simply involved particularly easy items to locate. Every paired item that was in a highly successful configuration from Experiment 1 has been switched to be in a configuration that was found to be more difficult in that experiment. Again, the intention is to insure that it is the abstract temporal configurations that matter and not the specific items.

3.1 **Participants**

Participants were tested in a local science center. The Younger group consisted of 22 children ($M = 5;6, 13$ girls) and the Older group consisted of 13 children ($M = 8;5, 9$ girls). All children had English as their sole or dominant language.

3.2 **Stimuli and procedures**

The procedures used in this study were identical to those used in Experiment 1. The stimuli consisted of the same types of items, but as noted, new specific linguistic terms were used (see Figure 1) and new arrangements of the events pairs were used.

3.3 **Results**

The same analyses were conducted with these data. Figure 2 shows the correlation between children’s placements and the objectively true temporal ordering of the events. As with Experiment 1, significant positive correlations were found between
these two measures for both the Younger children (Pearson’s $r = .23$, $N = 440$, $p < .001$) and the Older children (Pearson’s $r = .65$, $N = 260$, $p < .001$).

A repeated measures ANOVA was conducted using age group as a between-subjects variable, cue phase (Explicit Linguistic vs. General Event Descriptions) and timezone (Past, Present, and Future) as the independent variables, and proportion correct as the dependent variable. A main effect was found for age group ($F(1, 33) = 13.60$, $p < .001$, partial $\eta^2 = .29$) as the Older children continued to out-perform the Younger children in general; there was also a main effect for cue type ($F(1, 33) = 18.27$, $p < .001$, partial $\eta^2 = .36$) with children doing better overall with the General Event Descriptions than the Explicit Linguistic cues. The only significant interaction was the full 3-way interaction of all the variables ($F(1, 33) = 3.88$, $p = .031$, partial $\eta^2 = .20$) which was driven by the fact that the Younger children did unaccountably badly with the General Event Description item involving the Present time while the Older children achieved a perfect score with this item. This effect does not have a principled explanation (the identical item was used in Experiment 1 and did not cause problems) and appears to be a local anomaly. It is worth noting that when the present time items are omitted from this analysis, the main effect for cue phase remains (children overall did better with the General Event Descriptions) but the interaction does not.

The analysis of the paired event descriptions yielded somewhat different effects from Experiment 1. A repeated measures ANOVA with time pairing (Both Past, Both Future, and One Past & One Future) found a main effect for age group ($F(1, 33) = 13.00$, $p < .001$, partial $\eta^2 = .28$) and time pairing ($F(2, 32) = 5.21$, $p < .011$, partial $\eta^2 = .25$) and an interaction between the two ($F(2, 32) = 6.17$, $p < .005$, partial $\eta^2 = .28$). Follow-up analyses showed that Older children succeeded with all time pairings (means ranged from .85 to .88 correct) but that Younger children – similar to Experiment 1 – were more accurate with pairs that crossed the deictic center relative to pairs where both items were in the past ($t(21) = 5.1$, $p < .001$) or in the future ($t(21) = 4.1$, $p < .001$). However, unlike in the last study, there was no difference in how accurately the children linearly ordered the items in the past and future times for either the Younger children ($t(21) = 0.9$, $p = .38$) or the Older children ($t(12) = 0.81$, $p = .44$). Comparing across the two studies, it appears that children did equivalently well (or rather, badly) at ordering the all-future items ($t(67) = .14$, $p = .71$) but their performance dropped in Experiment 2 for ordering the all-past items ($t(67) = 8.89$, $p = .004$).

The final analysis incorporated remoteness information. As with Experiment 1, present time items were omitted as were linguistic items with no specified temporal distance (e.g., before, after). Because only two remoteness levels (Very Close

---

2. As with the previous study, these unspecified linguistic items received intermediate accuracy scores which did not statistically differentiate them from any of the specified remoteness levels.

All rights reserved
and Close) were used in both the cue conditions the omnibus repeated measures ANOVA used just those two levels of remoteness, along with the independent variables of cue type (Explicit Linguistic vs. General Event Descriptions, timezone (Past vs. Future) and age group; the dependent variable was timezone accuracy. This analysis found a main effect for age group \(F(1, 33) = 8.25, p = .007, \text{partial } \eta^2 = .20\) and several interactions with age group all of which stemmed from the fact that Older children performed almost at ceiling and all the condition differences depended on Younger children. Therefore, a second ANOVA looking just at the Younger age group found these children were more accurate with the General Event Descriptions than the Explicit Linguistic items \(F(1, 21) = 16.30, p = .017, \text{partial } \eta^2 = .43\). No other main effects were significant but there was a significant interaction between cue type and remoteness \(F(1, 21) = 12.05, p = .002, \text{partial } \eta^2 = .37\) and between time and remoteness \(F(1, 21) = 5.50, p = .03, \text{partial } \eta^2 = .21\). Inspection of the data suggested that these interactions arose because the Younger children did particularly well with the Very Close Future items in the General Event Descriptions, and this intuition was confirmed by post-hoc t-tests \(p’s \text{ ranged from .001 to .06}\). In addition, a further ANOVA conducted just within the General Event Descriptions found that the most remote (Far) items in this condition patterned consistently with the intermediately remote (Close) items across the board.

3.4 Discussion

With the change of specific items, Experiment 2 was able to replicate some – but not all – of the results from Experiment 1. One overall general difference between the studies was the fact that the 5-year-olds failed to perform above chance on many of the linguistic items and thus performed worse overall compared to Experiment 1; similar declines were not seen with the general event descriptions, however. With respect to the TPH, one of the new results provides extra support. In Experiment 2, 5-year-olds again located pairs of events better when the deictic center was between them, but the 8-year-olds no longer showed that effect (they were effectively at ceiling). The fact that the influence of a potential cognitive bias (the desire to organize time around a deictic center) can decline with age supports the notion that the bias guides the learning process and is not the outcome of the process. By contrast, unlike the previous experiment, Experiment 2 found no asymmetry between the past and future and no effects of remoteness. Given that remoteness is the best test of the TPH because it is the only temporal feature that is not grammaticized in English, the lack of effect speaks against the hypothesis. Indeed, these results suggest that children’s ability to interpret remoteness is more ad hoc and dependent on the specific arrangement of events and linguistic items used.
4. General discussion

The TPH states that common grammatical patterns across languages arise fundamentally from the same biases in cognition that guide early language development. Cross-linguistic studies of linguistic typology can thus provide working hypotheses about what concepts children may use, regardless of the specific elements in their native language. Three typologically prevalent temporal elements were investigated using a timeline task: the use of the deictic center to organize time into a past, present, and future; an asymmetry between the past and future times; and the use of temporal remoteness from the present moment as an organizing metric. The first of these elements is clearly used to organize the English temporal system, the second is potentially instantiated in English, and the third is not grammaticized in English.

The results showed that children as young as five years old can appropriately use a timeline and understand a range of temporal adverbs and general event descriptions. However, their performance on the task provided only limited support for the TPH. The children showed strong understanding that time was deictically organized – they correctly placed items in the past, present, and future times, and their ability to linearly order pairs of events was improved when the deictic center could be used as an intervening anchor point. However, such results are weak evidence for the TPH as the use of a deictic center is not only common typologically, but it is also a part of the grammatical system of the language the children are learning. For the typological patterns that are not as fully grammaticized in English (the asymmetry between past and present) or not grammaticized at all (remoteness), the data was far more ambiguous: Experiment 1 found support for the idea that children’s past and future representations are asymmetric (with past being better understood) and also that their ability to correctly place events into the past, present or future timezones was influenced by temporal remoteness (with more distal items being more accurately located); however, Experiment 2 failed to replicate these latter two findings. Thus, it is still possible that children’s understanding of English is influencing their performance, but less likely that non-English patterns of organization – however typologically prevalent – are guiding their understanding.

This failure to replicate for the non-English elements lends support to the alternative “thinking for speaking” position because it shows that children are being influenced by their own native language’s organization (and not by other common means of organizing time in language). Moreover, this point of view is also supported by another core dimension of these results, namely that specifics matter. The differences between the two experiments were all in the specific items – slightly different adverbials and different combinations of specific items within the task. While these small changes did not influence the importance of the English instantiated dimension of using a deictic center, they did change the influence of the
non-English related temporal dimensions on locating the events. The acquisition of specific adverbs is a protracted process (see Tillman & Barner, 2015) and as noted previously, these items have an intermediate status with respect to the grammar. For these items to develop the habits necessary to guide “thinking for speaking” would take time and even so, might develop on a word-by-word basis as the different adverbials differ in terms of complexity and frequency. Thus, the “thinking for speaking” alternative predicts that the specific adverbials should matter in this task.

However, the “thinking for speaking” position does not predict the pattern of results found in Experiment 1 nor does it explain the typological prevalence of an asymmetry between the past and future tenses nor the use of remoteness as a temporal organizer more generally. Perhaps a third alternative is in order. Cognitive biases can be conceptualized in many different ways and it is possible that a more computationally oriented perspective would be helpful. Temporal elements which are typologically common may not arise because of some a priori conceptual bias to organize time around those dimensions but may instead reflect local maxima in any system designed to express time. As such, these local maxima might influence children’s temporal organization but their effects would be expected to be fleeting and fragile, as they would need to compete with a strong external signal, namely the actual temporal organization of the child’s language. These ideas are speculative, of course, but they suggest fruitful lines of ongoing inquiry.

There is an old story about a drunk looking for her keys on a dark street underneath a streetlamp. When asked why she is looking there (of all places) she replies “Because that’s where the light is.” One of the hardest aspects of studying the relationship between cognitive biases and language development is identifying elements that are not in the direct spotlight of the child’s target language. The TPH is a way of adding a streetlamp that can widen our hypothesis space. The current data provided equivocal evidence that the keys to understanding the cognition that undergirds temporal semantics are actually lying in this particular patch of light, but the general approach has expanded the possible places we could look for them.

Acknowledgements

These data were collected at the Center of Science and Industry (COSI) in Columbus OH in the little kidspace® area and in the Buckeye Language Network’s Language Sciences Research Lab. My thanks to COSI for their support and enthusiasm for allowing me to conduct research on-site. Thanks go to the research assistants who created the stimuli, ran the participants, and coded the data – especially Kerrianne Morrison and Sonam Jindal.
References


### Appendix of supplementary resources

There are many resources available to learn more about the typological variety of the world’s languages – for temporal systems as well as other dimensions of language. The links below provide a good starting place:

- Typological Tools for Field Linguists: The Tense, Moods and Aspect Questionnaire
  From the Department of Linguistics at the Max Planck Institute for Evolutionary Anthropology: <https://www.eva.mpg.de/lingua/tools-at-lingboard/questionnaire/tma_description.php>

- The World Atlas of Language Structure
  (Search for “tense” as a name under the Features tab to see temporal language information) <http://wals.info/>

- Ethnologue Website
  (Search under the Languages tab to see the languages described) <https://www.ethnologue.com>