

BRIEF RESEARCH REPORT

**The communicative contexts of grammatical
aspect use in English***

PAUL IBBOTSON AND ELENA LIEVEN

University of Manchester, UK

AND

MICHAEL TOMASELLO

Max Planck Institute for Evolutionary Anthropology, Germany

*(Received 24 April 2012 – Revised 20 September 2012 – Accepted 7 February 2013 –
First published online 29 May 2013)*

ABSTRACT

In many of the world's languages grammatical aspect is used to indicate how events unfold over time. In English, activities that are ongoing can be distinguished from those that are completed using the morphological marker *-ing*. Using naturalistic observations of two children in their third year of life, we quantify the availability and reliability of the imperfective form in the communicative context of the child performing actions. On average, 30% of verbal descriptions refer to child actions that are grounded in the here-and-now. Of these utterances, there are two features of the communicative context that reliably map onto the functions of the imperfective, namely, that events are construed as ongoing and from within. The findings are discussed with reference to how the context in which a child hears aspectual language may limit the degrees of freedom on what these constructions mean.

INTRODUCTION

Many of the world's languages grammaticalize how events unfold over time, for example, by marking events that are viewed as ongoing from those that

[*] Thank you to Nikki Shepherd who coded and organized the data, to Laura Wagner, and to the anonymous reviewers who provided insightful comments on earlier drafts of this paper. Address for correspondence: Paul Ibbotson, School of Psychological Sciences, University of Manchester, Manchester M13 9PL, United Kingdom. e-mail: paul.ibbotson@open.ac.uk

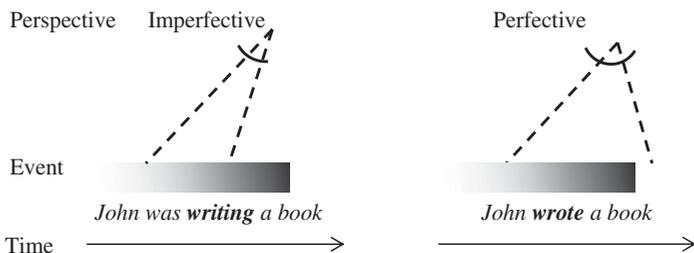


Fig. 1. A schematic of the imperfective–perfective distinction in English. By focusing on the event from within, imperfective aspect construes the scene as open and ongoing. In contrast, perfective aspect construes the scene as a completed whole, bringing into focus event boundaries.

are completed. In these languages the child’s task is to learn how the temporal contours of an event are expressed in their language. The social, pragmatic, and referential context in which a child hears aspectual language is a rich source of information that may limit the degrees of freedom on what these constructions mean. It is this relationship between context and aspectual input that we focus on in this study.

In English, one can say *John was writing a book, but he never finished it*, but it sounds odd to say *John wrote a book, but he never finished it* (note that they are both in past tense). The classic explanation for this is that the imperfective aspect, marked by *-ing*, coerces people into construing the activity ‘from within’, blurring the boundaries of the event so that there is no clearly defined beginning or end (Comrie, 1976). By contrast, unmarked non-progressive forms (verbs without the suffix *-ing*) bring into focus the boundaries of the event and thus activities are seen as accomplished. Lacking imperfective marking, the verb *wrote*, in the above example, is construed as a completed activity, which is why it sounds like a non sequitur to then add *but he never finished it* (see Figure 1). By mentally zooming in on the activity, the imperfective aspect can also turn actions into states, thus one has to say *he knows the answer* or *he wants a drink*, rather than *he is knowing the answer* or *he is wanting a drink*. Presumably, this is because the imperfective is redundant with verbs like *know* and *want*, which already denote states¹ (Pinker, 2007: 203).

There is a large amount of work documenting the comprehension, production, and emergence of aspect in child language (e.g., Clark, 1996; Johnson & Fey, 2006; Li & Shirai, 2000; Shirai & Andersen, 1995; Wagner, 2001). For instance, Li and Shirai (2000) showed that children’s tense–aspect use tends to fall into two prototypical clusters of features; one which groups

[1] This is not true for all languages. For example statives like *believe* and *know* can acceptably combine with the French *imparfait*.

atelic–durative–imperfective–present features together (e.g., *the teacher is carrying the box*) and one which groups telic–punctual–perfective–past features together (e.g., *the woman won the race*). Like children, adults find prototypical combinations easier to understand, and judge them as better sentences than those which draw on features from both groups, e.g., *John will have been walking* (Wagner, 2009). The relationship between the linguistic definition of grammatical aspect in English and children’s experience of different temporal events in a naturalistic setting is less well understood. We know that the progressive *-ing* is one of the earliest morphological markers children control (Brown, 1973). However, there seems to be relatively little, if any, research on the relationship between the communicative context in which aspectual language is used and the *meaning* of aspectual morphology such as *-ing*; including the role that the duration and timing of the event might play in this. We address this by investigating the mapping between the linguistic input (imperfective language in child-directed speech) and features of the child’s experience (duration of the activity they perform relative to verb) in a naturalistic setting.

In theory, grammatical aspect is independent of tense, lexical aspect, modality, and verbal semantics, but of course in practice it interacts with these features in complex ways (Croft, 2012; Li & Shirai, 2000; Shirai & Andersen, 1995). For example, one would predict that *break*, a verb that prototypically suggests a clear endpoint, would be incongruent with progressive aspect, as it is in *John was breaking his leg*. However, *break* can be construed in a relatively more progressive way if the semantics of other constituents in the phrase support this interpretation, as it does with *John was breaking the speed limit*. One foothold into this complexity is to pay attention to the different contexts in which adults use different chunks of language. Detecting the way adults reliably use language across different situations is a powerful learning mechanism children use for deducing the meaning of nouns, verbs, and other grammatical distinctions (Clark & de Marneffe, 2012; Scott & Fisher, 2012; Smith & Yu, 2008). To test whether there are such reliable associations between communicative context and aspectual language, we focus on a relatively simple pattern that two- to three-year-olds might be sensitive to; those verbs with the suffix *-ing* and those without. The reason for this is that English does not have a marked form for the perfective (unlike many Slavic languages²) so we can divide aspect between the marked imperfective – anything with the form *be V-ing* – and everything else with the form *be V*. The child must somehow learn to map this linguistic pattern to the conceptual distinction of imperfectivity – the ongoingness of the event to

[2] Researchers have suggested that the ‘ongoingness’ flavour of the English imperfective may come from its progressive nature and may not extend as thoroughly to other languages with non-progressive imperfectives (Comrie, 1976).

which the verb refers. A potential source of information that could constrain this mapping process is the child's own experience of performing actions, a hypothesis that to our knowledge has not been previously explored. One possibility is that the embodied experience of performing an action facilitates construing the scene from within and blurring the boundaries of the event. This would be a possibility if there were enough instances where imperfective morphology, rather than the unmarked form, is predominantly experienced as the child was performing an ongoing action.

We consider the wider social-pragmatic environment because we know different linguistic symbols fit into different social niches in different ways, and this has an effect on how they are learned. For example, it seems nouns are easier to learn in ostensive contexts while verbs are learnt best in non-ostensive contexts (Ambalu, Chiat & Pring, 1997; Kuczaj, Carter, Sherman & Borys, 1989; Tomasello & Kruger, 1992; Tomasello & Todd, 1983; Whitehurst, Kedesdy & White, 1982). Tomasello (1992) notes that despite the verb-general preference for non-ostensiveness, a number of his subjects' early verbs did seem to be learned ostensively, especially those words for distinctive child actions such as sweeping, waving, swinging, and crying – note their imperfective marking. Although we do not know the exact pragmatic contexts in which children hear either the perfective or the imperfective, we would argue that the child's task is basically the same as it with verbs and nouns: to determine adults' focus of attention and assign some function to a linguistic symbol, in this case *-ing*. The actions children perform naturally vary in how ongoing and durative they are. We investigate how FREQUENTLY and RELIABLY this variation maps onto the perfective/imperfective language children hear in child-directed speech (CDS; see MacWhinney, Bates & Kliegl, 1984, for similar approaches to linguistic cues). First, we establish to what extent caregivers are narrating their children's actions as they perform them, in the here-and-now. This provides us with a broad estimation of how available the aspectual cue is at the time the child is performing the action. Second, we assess how reliable the cue is when it is present, that is, how consistently does the ongoingness of the action as performed by the child map onto the perfective/imperfective aspect used by the caregiver to describe that event. The results of these analyses tell us whether the child could potentially capitalize on this source of information to infer certain facts about the way grammatical aspect works in English, essentially learning the function of *-ing*.

METHODS

The corpus

For obvious reasons, traditional text-based corpora make it difficult to explore the nature of the event that verbs refer to, therefore we examine

video-recordings. The data presented here are based on a total of six hours of mother–child interaction for two of the children studied in the Lieven, Tomasello, Behrens & Spears (2003) corpus. They are divided into three developmental periods approximately six months apart beginning at two years of age, with one hour’s worth of data from each age. By choosing windows at either end of the available corpus we attempted to maximize our chances of capturing any developmental change. The videos recorded naturalistic play between the mother and the child.

Coding procedure

Using the video-recordings, for every verb uttered by the caregiver we coded two features; the ‘groundedness’ of the event that the verb referred to and the aspectual morphology of the verb. First, we were interested in whether the verb referred to an event that was grounded in the immediate context or displaced from the here-and-now. Whether an event is grounded in the here-and-now is a matter of degree. For this study we defined the here-and-now as a 60-second window, 30 seconds either side of the onset of the verb; this was timed from the clock display on the video recording. For example, if the mother said *We baked a cake, didn’t we?* at T₁, we searched for a ‘baking event’ between T₁ minus 30 seconds and T₁ plus 30 seconds. If we found no referent associated with the verb within this window – because the mother was talking about when they baked a cake yesterday, for example – we coded this reference as displaced. If the mother said *Oh Thomas! You’re hitting it*, and the child was hitting something ± 30 seconds from when verb was uttered, then this was coded as a grounded event. This gives us a rough estimate of how recoverable the event is from the immediate context at the time the verb is uttered. If the event did occur within the window, we then coded for properties of the event, namely the duration of the action relative to the verb that described the action. This gives us an approximation of how reliably the imperfective verbal morphology maps onto the child’s action. Of course, the duration of actions naturally vary from the near instantaneous to those that extend over long periods of time. The lower limit was placed at one second for instantaneous events (e.g., *you snapped his head off!*), and the upper limit for ongoing events (e.g., *are you hiding from me?*) was the 60-second window we chose to search for the events. Thus the minimum an event could last was one second and the maximum was one minute. All verb forms analyzed in the study are provided in the ‘Appendix’.

Inter-rater reliability

We were interested in the reliability of two measures, the timing of the verb in relation to the event, and the duration of the event. We calculated the

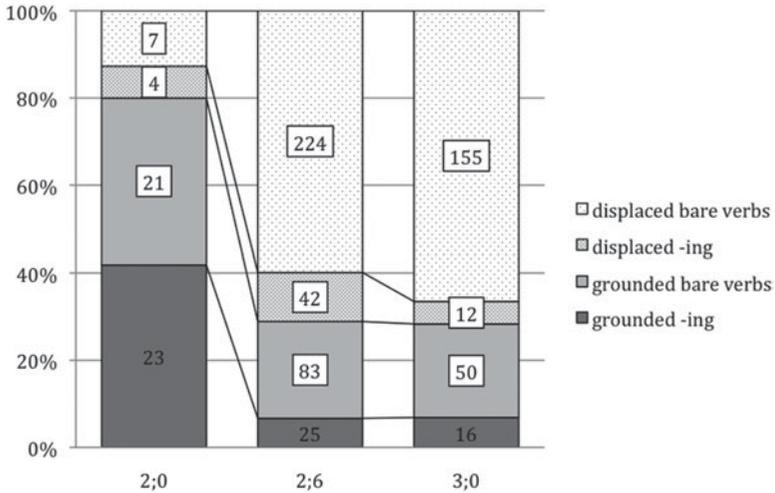


Fig. 2. Thomas' CDS broken down by grammatical aspect and groundedness.

coding reliabilities between two raters for a 1-hour sample of video (~18% of the data). Because both measures are continuous we could not use Cohen's kappa. The paired verb timing from the two raters showed a significant bivariate correlation coefficient (Pearson's $r=0.62$, $N=56$, $p<.001$) indicating agreement on the time at which a verb was uttered. The paired event duration timing between the two raters showed a non-significant difference (paired sample t -test ($t=-0.48$, $df=57$, $p=.627$)), indicating that there was no statistical difference between the two raters on the event duration.

RESULTS

Availability analysis : grounded versus displaced events and grammatical aspect

Figures 2 and 3 show the proportion of all verbs that are grounded versus displaced. These are further divided by those which carry imperfective morphology (-ing) and those which do not. The data are presented at the three developmental time-points for each child. Raw figures are displayed within the bars; proportion is displayed on the y-axis.

In total there were 1,099 adult verbs coded across six hours of video-recording, 662 for Thomas and 437 for Eleanor. For Thomas, there was a reduction from 80% to 30% in grounded events between 2;0 and 2;6 (and sustained until 3;0). Eleanor's CDS had a much lower level of grounded events to begin with (38% at 2;0), reducing to about 22% at 2;6 then recovering somewhat to 28% at age 2;11. In this data, looking at verbs, there was only one developmental period for one child where 'here-and-now'

GRAMMATICAL ASPECT USE IN ENGLISH

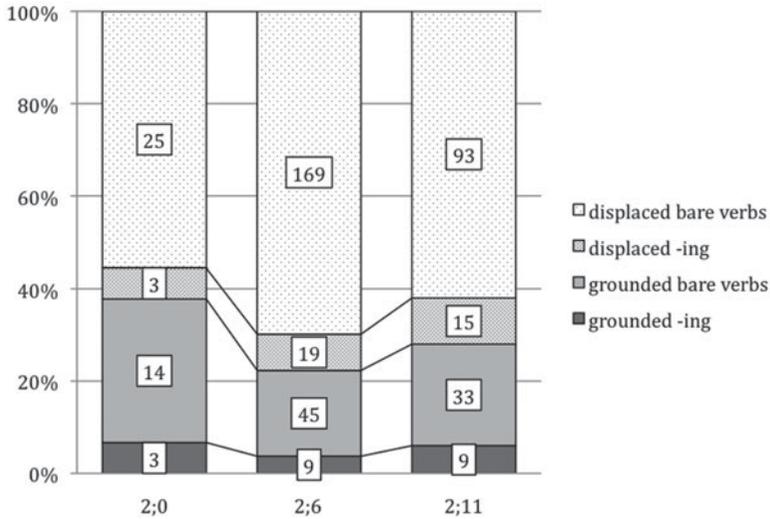


Fig. 3. Eleanor's CDS broken down by grammatical aspect and groundedness.

utterances represented the majority (2;0, for Thomas). As we might expect, the unmarked perfective form is more frequent overall and at most ages it outnumbers the imperfective form 4:1. The exception to this is in Thomas' CDS at 2;0 for grounded events when both forms are approximately equally frequent. Averaging over both children and all ages, approximately 30% of verbs were grounded, that is, when a verb was uttered a clear referent could be recovered from the 60-second window. The next analyses looks at the nature of the 30% of verbs which do have a grounded referent, to see if there is a relationship between the nature of the events and the pattern in the perfective/imperfective language used to describe them.

Reliability analysis: nature of the event and grammatical aspect

Figures 4 and 5 show the data from the 331 grounded verb–event pairs, codable from the video-recording. The horizontal bars represent the duration of the events. To compare one event with another we zeroed all verbal utterances, represented on the graph as the solid vertical line at 0. To do this, the beginning of the event was calculated as [beginning of event – time of verb utterance] and the end of event was calculated as the [end of event – time of verb utterance]. Events that appear to the right of this line are performed after the caregiver has uttered the verb, events to the left appeared before, and events that overlap the line are concurrent with the utterance. The dashed line divides all verbs into *-ing* verbs (above the line) and non-*ing* bare verb forms (below the line).

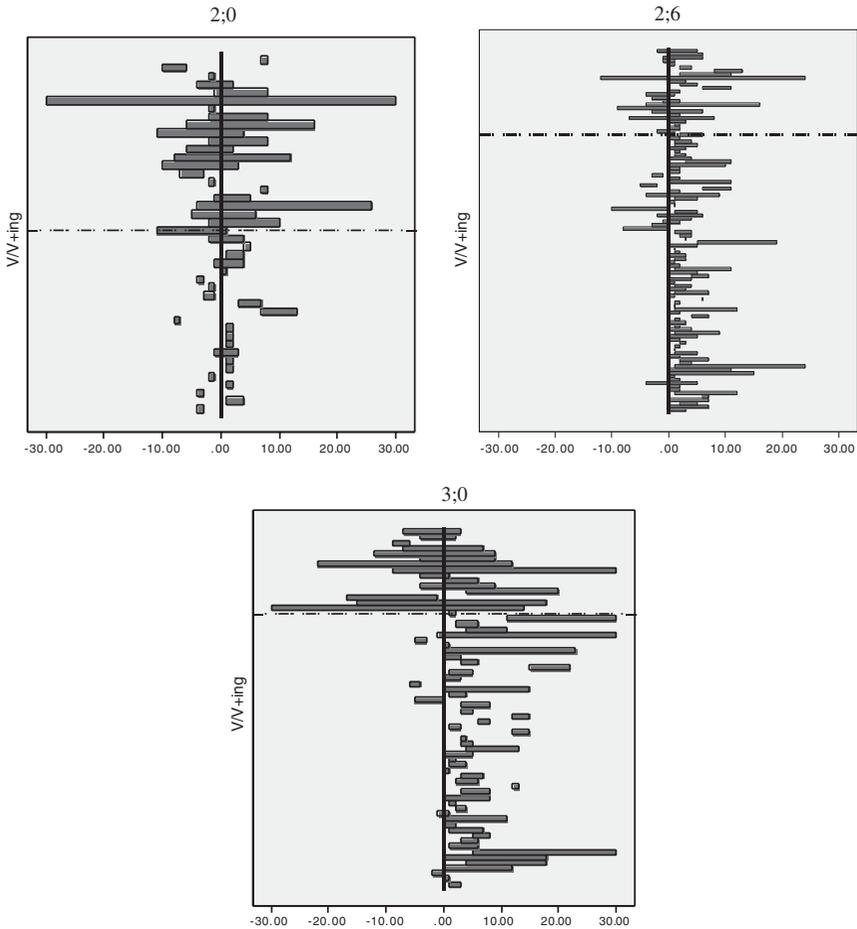


Fig. 4. Thomas: duration of event and grammatical aspect.

Event duration

We examined event duration because ongoingness is a crucial part of the perfective/imperfective function. A Mann–Whitney U test was conducted between the duration (in seconds) of the *-ing* events and the non-*-ing* events. Summing the data across all ages, for Thomas, this showed a significant difference ($U(218)$, $Z=4.92$, $p<.001$) between events labelled with *-ing* ($Med=6.5$ seconds) versus those without ($Med=3.0$ seconds) such that events labelled with imperfective morphology were longer. Summing the data across all ages, for Eleanor, the same comparison showed the same result. There was a significant difference ($U(113)$, $Z=5.06$, $p<.001$) between

GRAMMATICAL ASPECT USE IN ENGLISH

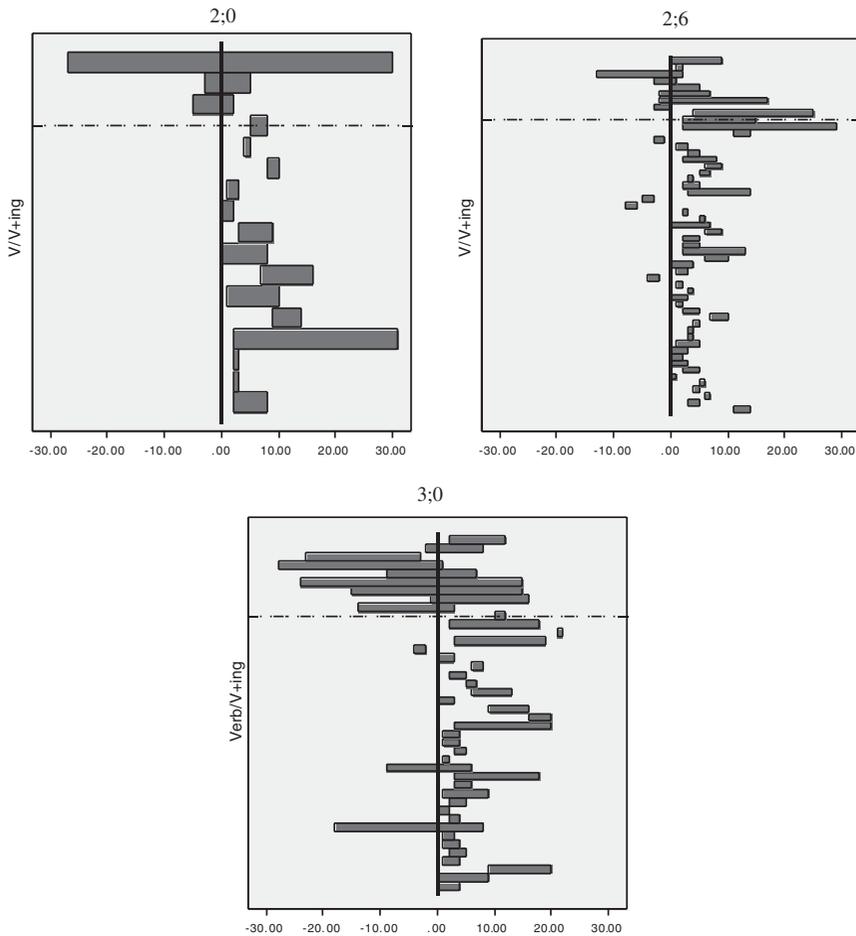


Fig. 5. Eleanor: duration of event and grammatical aspect.

events labelled with *-ing* ($Med=12.5$ seconds) versus those without ($Med=3.0$ seconds), such that events labelled with imperfective morphology were longer. On average, across all time periods and both children, events described with *-ing* were over four times longer than those without.

Verb timing relative to event

We examined the timing of the verb since construing the event from within is a crucial part of the imperfective function. We calculated the proportion of verbs that were uttered as the child was performing the action (concurrent)

TABLE 1. *A summary of frequency counts for those events that crossed the verbal utterance (X) versus those that did not (no X), displayed by aspectual marking*

	Thomas		Eleanor	
	X	no X	X	no X
-ing	38	26	15	6
non-ing	8	146	2	90

against those which were not (non-concurrent). This was done by counting the frequency of those event times with a negative AND a positive value, showing that they had crossed the 0 second timeline when the verb was uttered. We then divided those verbs into ones which had imperfective *-ing* and those which did not (Table 1).

Chi-squared (with Yates' correction) on Thomas' data showed that there was a significant difference between the likelihood of imperfective and perfective verbs overlapping with the event ($\chi^2(1, N=218) 76.49, p < .001$). For Thomas a word with *-ing* morphology is 1.5 times more likely to occur within the boundary of the event (38/26). A word lacking *-ing* morphology is 18.2 times more likely to occur outside the boundaries of the event (146/8).

Chi-squared (with Yates' correction) on Eleanor's data showed that there was a significant difference between the likelihood of imperfective and perfective verbs overlapping with the verb ($\chi^2(1, N=113) 58.85, p < .001$). For Eleanor a word with *-ing* morphology is 2.5 times more likely to occur within the boundary of the event (15/6). A word lacking *-ing* morphology is 45 times more likely to occur outside the boundaries of the event (90/2).

The results from the reliability analysis show that imperfective form *V-ing* was associated with longer actions on average and much more likely to occur while the child is performing the action to which the verb refers. It is also worth noting that most of the perfective events are displaced to the right of the verb line, according with Tomasello and Kruger (1992) who found that majority of verbs were used by mothers to refer to actions that mothers wished children to perform or that they were anticipating their performing.

Finally, we wanted to know whether verb type is confounded with aspectual morphology. Where an *-ing* verb occurred in a grounded context we examined whether that verb occurred at least once without *-ing* (Table 2). For example, for the verbs *doing*, *taking*, and *giving* we looked for at least one instance of *do*, *take*, and *give*.

There is clearly variability between the two children's CDS, but on average, when a *-ing* verb occurs it is available 64% of the time in a non-*-ing* verb. This shows that verb type disassociates with respect to aspectual morphology in CDS. Interestingly, for both children this proportion

TABLE 2. *Percentage of -ing verbs that also occur in the unmarked form (raw figures in parentheses)*

	2;0	2;6	3;0	Mean
Thomas age				
Raw figures	(5/23)	(13/25)	(11/16)	
%	22	52	68	47
Eleanor age				
Raw figures	(2/3)	(7/9)	(9/9)	
%	66	77	100	81

increases with age, providing suggestive further evidence to the claim that, early on, children's verb-specific use falls into one of two prototypical tense-aspect clusters (cf. Li & Shirai, 2000; Wagner, 2009).

In summary, Figures 4 and 5 demonstrate that for both Thomas and Eleanor the perfective-imperfective language used in CDS maps consistently onto the ongoingness of the event and how likely the verb is uttered at a time when the child will be in the midst of performing the action.

DISCUSSION

We set out to characterize the pragmatic context in which children hear grammatical aspect in English. We looked at the duration and timing of children's actions relative to the verb that described these actions in CDS. We interpret these findings following the structure of the data analysis; first we discuss availability and then the reliability results.

We established to what extent caregivers are grounding their descriptions of their child's activities in the immediate context. Averaging over both children and all developmental time-points, we found 30% of verbs referred to an event that was recoverable within a 60-second window of the verb being uttered. We also noted that the unmarked perfective verbal forms outnumber imperfective forms 4 to 1. This gave us a broad estimation of how available and recoverable the referent of the verb was from the immediate context.

At a general level, it is possible to think of the child's task in working out how language is used as a signal detection task in a noisy environment (e.g., MacWhinney *et al.*, 1984). The 30% figure of the imperfective/perfective distinction in grounded contexts represents the availability of the signal. All following claims regarding the RELIABILITY of this signal or cue must be made with respect to this baseline denominator. In theory, the hardest pattern to detect would be an infrequent and unreliable cue. We don't know for sure whether the 30% rate reaches some critical threshold of tokens whereby we could say this cue is now 'available'. But what we can say is that the less available the cue is—and in this case it is more unavailable than available—the more it will need to be reliable (when it is available) if it is to be

detectable or learned. Of course, displaced referents can still be construed as either more ongoing – *do you remember running yesterday* – or less ongoing – *remember when you broke your leg*. Recalling these events (or projecting into the future) still construes in the mind of the speaker/addressee an event as seen from within and thus ongoing or from a ‘bird’s-eye’ position and thus completed. One might speculate that a proportion of these displaced actions also received descriptions contemporaneously. It would indeed be interesting to look at these cases, but the present study did not analyze these examples for two practical reasons. First, the coding of the speakers’ referential intention became unreliable much beyond a 60-second window from the verbal utterance, and second, in the case of events that happened off camera, they were completely uncodable in terms of measuring the duration and timing of the event.

One final point to be made regarding the availability analysis is that, for Thomas, the first development window showed a much higher level (80%) than the average grounding figure (30%). For Thomas at least, CDS was more grounded with respect to verbal referents early on in development. We know Eleanor’s language ability is advanced in comparison to other children in this corpus (e.g., Lieven, Salomo & Tomasello, 2009). It could be that we would have also seen higher levels of grounded CDS for Eleanor had we been able to measure at an earlier stage than our corpus allowed. The pattern seen in Thomas raises the possibility that adults gradually talk about a larger proportion of displaced events as language develops. More children and a more detailed analysis of the data would confirm the precise nature of this interaction. For example, do adults expand the proportion of displaced events as a result of children talking more about displaced events, or do adults, realizing that children have reached a certain proficiency in their language, start to introduce more displaced events into the conversation? There are claims that CDS can be characterized as relatively more grounded in the here-and-now relative to adult-to-adult speech (Brown & Bellugi, 1964; Cross, 1977; Dil, 1971; Snow & Ferguson, 1977). However, as our study testifies, there are large proportions of development in which speech is significantly displaced (see also Gleitman, 1990).

We now turn to the reliability of the aspectual distinction. That is, regardless of how frequent the cue is, how consistently does the ongoingness of the action as performed by the child map onto the perfective/imperfective aspect used by the caregiver to describe that event? We found that the length of the child’s action predicts both the caregivers’ choice of grammatical aspect and the co-occurrence of action and verb. These features correspond with two functions of imperfective aspect, namely, that events are construed as ongoing and from within.

First, the association between child event duration and CDS imperfective morphology is fairly clear. Longer actions are more likely to be associated

with verbs that have been inflected with *-ing*. This provides a strong cue to the child that one of the functions of imperfective morphology is to indicate how ongoing the activity is. The reliability of this cue is useful to the child in terms of signal detection, given the low availability rate of grounded events. The second claim regarding the timing of the verb and the child's action perhaps requires more unpacking. There is mounting evidence from the embodied cognition literature that language develops in a tight association with perception and action (Childers & Tomasello, 2003; Ejiri & Masataka, 2001; Glenberg, 2007; Goodwyn, Acredolo & Brown, 2000; Hahn & Gershkoff-Stowe, 2010; Iverson & Fagan, 2004; Iverson & Goldin-Meadow, 2005). When one hears a linguistic symbol, for example *running*, this re-activates stored sensorimotor information associated with running events (Barsalou, 1999, 2008). The embodied cognition view emphasizes that cognition is a situated activity and as such we should include sensorimotor activities as part of the pragmatic context. This is very similar to a claim made by early developmental theorists such as Piaget (1952), Werner and Kaplan (1963), and Nelson (1974). With respect to grammatical aspect, the hypothesis is that the embodied event of performing the action facilitates construing the scene from within and blurring the boundaries of the event. This is because, if the verb has imperfective *-ing*, the child is much more likely to be in the midst of action than if the verb is in its perfective form. By building a record of these embodied events in association with the aspectual morphology they do or do not hear, children can begin to form a verb-general notion of what the function of *-ing* is. The Chi-squared analysis provides support for this hypothesis, namely, more events described with *-ing* morphology occur concurrently with the verbal utterance than those without *-ing*. Again, the reliability of this cue is useful to the child in terms of signal detection given the low availability rate of grounded events. However, it is noticeable that there is an even stronger pattern for verbs that lack *-ing* to not occur concurrently with the event. In short, the majority of *-ing* forms occur concurrently with the event but sometimes fall outside of it; the non-*ing* forms very rarely occur concurrently with the utterance. One could reason from this data (if one were a language learner) that whatever the function of the unmarked form, it is not associated with contexts where one has an internal perspective on an event. As mentioned above, there are a minority of grounded *-ing* forms that do not overlap with the child performing the event, for example sentences of the form *you were hitting him, weren't you?* Further analysis of the pragmatic context and associated linguistic material may be able to distinguish between grounded overlapping *-ing* use and non-overlapping *-ing* use.

As we have emphasized, the claims regarding reliability need to be understood in the context of the base rate 30% average verbal grounding figure (with the qualification about Thomas' developmental differences).

More generally, the methodology we have used here advocates exploring linguistic cues in combination with the social–pragmatic context. By using video evidence we have been able to consider a broader range of cues than a traditional corpus-based approach. By doing so, we have been able to get closer to reconstructing the rich social–pragmatic–linguistic world in which the child grows up. The challenge is to explore ways in which social–pragmatic skills interact with prodigious pattern-finding abilities in a way that explains the emergence of linguistic knowledge. For instance, French has numerous homophones, especially in the form of infinitives and past participles for class 1 verbs. Similar to the approach taken in this study, Clark and de Marneffe (2012) examined the pragmatic contexts in which children hear the crucial alternations. They found that adult speech to young children consistently distinguishes infinitival uses from participial ones: they are produced in different contexts, depending on the timing of the event being talked about, and they are presented in different constructions. Thus adults consistently mark the potentially ambiguous homophones by distinguishing the meaning carried by the infinitive from that carried by the past participle in the communicative context.

Grammatical aspect is potentially an abstract notion. However, the social, pragmatic, and referential context in which a child hears aspectual language is a rich source of information that may limit the degrees of freedom on what these constructions mean. Detecting the way adults reliably use language across different situations is a powerful learning mechanism (Clark & de Marneffe, 2012; Scott & Fisher, 2012; Smith & Yu, 2008). The data from these observations suggest the main perfective/imperfective distinction is available in grounded communicative contexts for approximately 30% of verbal utterances. Of these utterances there are two features of the communicative context that reliably map onto the functions of the imperfective–perfective distinction, namely the duration and the timing of the event. This may go some way to explaining why it is one of the first morphological markers to come under the child’s control (Brown, 1973). The data also provide one explanation of why imperfective aspect coerces people into construing the event from within. When the function of *-ing* was being constructed in development, the majority of actions associated with *-ing* are performed from within the action.

REFERENCES

- Ambalu D., Chiat S. & Pring T. (1997). When is it best to hear a verb? The effects of the timing and focus of verb models on children’s learning of verbs. *Journal of Child Language* **24**, 25–34.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences* **22**, 577–609.
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology* **59**, 617–45.

- Brown, R. W. (1973). *A first language*. Cambridge, MA: Harvard University Press.
- Brown, R. W. & Bellugi, U. (1964). Three processes in the child's acquisition of syntax. *Harvard Educational Review* 34, 133–51.
- Childers, J. B. & Tomasello, M. (2003). Children extend both words and non-verbal actions to novel exemplars. *Developmental Science* 6, 185–90.
- Clark, E. V. (1996). Early verbs, event types and inflections. In C. E. Johnson & J. H. V. Gilbert (eds.), *Children's Language* 9, 61–73. Mahwah, NJ: Erlbaum.
- Clark, E. V. and de Marneffe, M.-C. (2012). Constructing verb paradigms in French: adult construals and emerging grammatical contrasts. *Morphology* 22, 89–120.
- Comrie, B. (1976). *Aspect*. Cambridge: Cambridge University Press.
- Croft, W. (2012). *Verbs: aspect and clausal structure*. Oxford: Oxford University Press.
- Cross, T. (1977). 'Mothers' speech adjustments: the contribution of selected child listener variables'. in C. Snow and C. Ferguson (eds.), *Talking to children: language input and acquisition*, 151–88. Cambridge: Cambridge University Press.
- Dil, A. (1971) 'Bengali baby talk'. *Word* 27, 11–27.
- Ejiri, K. & Masataka, N. (2001). Co-occurrence of preverbal vocal behavior and motor action in early infancy. *Developmental Science* 4, 40–48.
- Gleitman, L. R. (1990). The structural sources of verb meanings. *Language Acquisition* 1, 1–55.
- Glenberg, A. M. (2007). Language and action: creating sensible combinations of ideas. In G. Gaskell (ed.), *The Oxford handbook of psycholinguistics*, 361–70. Oxford: Oxford University Press.
- Goodwyn, S. W., Acredolo, L. P. & Brown, C. A. (2000). Impact of symbolic gesturing on early language development. *Journal of Nonverbal Behavior* 24, 81–103.
- Hahn, E. & Gershkoff-Stowe, L. (2010). Children and adults learn actions for objects more readily than labels. *Language Learning and Development* 6, 283–308.
- Iverson, J. M. & Fagan, M. K. (2004). Infant vocal-motor coordination: precursor to the gesture-speech system? *Child Development* 75, 1053–66.
- Iverson, J. M. & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science* 16, 367–71.
- Johnson, B. W. & Fey, M. E. (2006). Interaction of lexical and grammatical aspect in toddlers' language production. *Journal of Child Language* 33, 419–35.
- Kuczaj, S., Carter, A., Sherman, J. & Borys, R. (1989). On children's use of experience to determine the meanings of words. Paper presented to the Biennial Meeting of the Society for Research in Child Development, Kansas City.
- Li, P. & Shirai, Y. (2000). *The acquisition of lexical and grammatical aspect* (Studies on Language Acquisition, 16). Berlin: Mouton de Gruyter.
- Lieven, E., Tomasello, M., Behrens, H. & Spears, J. (2003). Early syntactic creativity: a usage-based approach. *Journal of Child Language* 30, 333–70.
- Lieven, E., Salomo, D. & Tomasello, M. (2009). Two-year-old children's production of multiword utterances: a usage-based analysis. *Cognitive Linguistics* 20(3), 481–508.
- MacWhinney, B., Bates, E. & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior* 23, 127–50.
- Nelson, K. (1974). Concept, word, and sentence: interrelations in acquisition and development. *Psychological Review* 81, 267–85.
- Piaget, J. (1952). *Origins of intelligence in children*. New York: International Universities Press.
- Pinker, S. (2007). *The stuff of thought: language as a window into human nature*. New York: Viking.
- Scott, R. & Fisher, C. (2012). 2;5-year-olds use cross-situational consistency to learn verbs under referential uncertainty. *Cognition* 122(2), 163–80.
- Shirai, Y. & Andersen, R. (1995). The acquisition of tense/aspect morphology: a prototype account. *Language* 71, 743–62.
- Smith, L. & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition* 106, 1558–68.

- Snow, C. E. & Ferguson, C. A. (1977). *Talking to children*. Cambridge: Cambridge University Press.
- Tomasello, M. (1992). *First verbs*. Cambridge: Cambridge University Press.
- Tomasello, M. & Kruger, A. C. (1992). Joint attention on actions: acquiring verbs in ostensive and non-ostensive contexts. *Journal of Child Language* **19**(2), 311–33.
- Tomasello, M. & Todd, J. (1983). Joint attention and lexical acquisition style. *First Language* **4**, 197–212.
- Wagner, L. (2001). Aspectual influences on early tense comprehension. *Journal of Child Language* **28**, 661–81.
- Wagner, L. (2009). I'll never grow up: continuity in aspectual representations. *Linguistics* **47**(5), 1051–74.
- Werner, H. & Kaplan, B. (1963). *Symbol formation*. New York: Wiley.
- Whitehurst, G., Kedesdy, J. & White, T. (1982). A functional analysis of meaning. In S. Kuczaj (ed.), *Language development, Vol. 1: syntax and semantics*, 397–427. Hillsdale, NJ: Erlbaum.

Appendix: all verb forms analyzed in the study

		Eleanor	
2;0	grounded	-ing	<i>fixing, drawing, drawing</i>
		bare verb forms	<i>put, tell, saw, put, count, turn, rub, draw, look, draw, change, find, stick, sing</i>
	displaced	-ing	<i>wanting, drawing, playing</i>
		bare verb forms	<i>bring, sing, bought, mix, fix, use, push, show, blow, make, play, find, play, give, hold, draw, tell, drawn, show, tell, lift, put, squeeze, clap, drink</i>
2;6	grounded	-ing	<i>eating, flying, having, going, looking, looking, shopping, swimming, taking</i>
		bare verb forms	<i>bring, close, come, dance, do, do, do, do, do, do, draw, draw, drawn, fallen, find, find, get, give, give, go, hold, hold, hold, join, kiss, kiss, like, look, look, look, put, see, see, show, show, show, swim, swim, switch, tell, tell, tell, tell, tell, write</i>
	displaced	-ing	<i>sitting, running, playing, running, going, making, going, swimming, swimming, singing, crying, having, cooking, picking, eating, doing, going, doing, trying</i>
		bare verb forms	<i>go, fall, show, show, look, says, do, bought, runs, do, see, bring, fallen, hurt, told, climbed, fall, hurt, do, take, do, hit, put, walk, eat, take, go, have, went, sing, get, post, delivers, come, get, delivers, get, bring, get, learn, swim, clap, do, look, do, make, look, finished, draw, do, do, draw, hold, draw, look, do, do, see, write, think, drew, help, draw, take, draw, help, do, have, wear, take, go, tell, wear, put, live, call, live, eat, go, took, play, look, call, sing, play, say, sit, pretend, talks, say, sleep, pretend, lie, have, count, lie, wake, do, have, write, take, drink, get, ride, see, changed, put, bought, give, come, put, lose, take, get, have, bought, get, put, get, bought, told, watch, do, like, gone, go, see, pack, put, swim, wear, swim, wave, tell, do, ask, eat, go, eat, get, go, drop, said, tell, go, watch, knock, come, give, get, find, sit, see, look, find, see, put, go, look, breathe, look, see, open, seen, wear, put, breathe, look, did</i>
2;11	grounded	-ing	<i>doing, writing, licking, talking, putting, telling, writing, having, colouring</i>
		bare verb forms	<i>come, count, tell, hold, dropped, look, do, do, write, tell, wait, show, give, write, lick, lick, play, push, tell, tell, put, put, put, press, do, push, press, guess, put, make, roll, look, press</i>
	displaced	-ing	<i>fishing, doing, telling, going, getting, drawing, going, having, getting, playing, going, going, going, writing, having</i>
		bare verb forms	<i>play, make, brought, like, put, get, play, do, write, do, stay, come, make, come, show, give, went, go, watch, wait, show, drive, do, look, put, read, talk, look, write, colour, do, come, give, look, press, see, tell, say, copy, watch, do, play, seen, play, make, think, play, give, watch, play, go, book, see, put, lost, do, play, do, make, do, tell, play, get, make, look, sit, seen, play, have, make, have, race, have, make, share, make, whisk, dip, fry, say, get, sleep, make, go, come, push, open, come, go, come, walk, see, walk</i>

			Thomas
2;0	grounded	-ing	<i>catching, drinking, flying, giving, going, hiding, kissing, looking, looking, playing, putting, seeing, spending, stacking, stacking, throwing, throwing, trying, using, walking, watching, watching, wearing</i>
		bare verb forms	<i>catch, catch, eat, eat, fallen, fallen, kick, put, ride, say, say, say, sit, smell, smell, throw, throw, throw, throw, throw, thrown</i>
	displaced	-ing	<i>bending, talking, walking, running</i>
		bare verb forms	<i>put, fell, open, got, fallen, meet, flew</i>
2;6	grounded	-ing	<i>bending, cleaning, clipping, coming, coming, coming, coming, coming, disposing, flashing, flying, flying, flying, flying, getting, jumping, jumping, walking, kicking, laughing, pressing, resting, running, running, saving</i>
		bare verb forms	<i>bend, clap, come, come, come, come, come, come, come, do, down, fall, fall, fallen, fight, find, flip, fly, fly, fly, fly, get, go, got, got, help, hold, hurt, hurt, hurt, kick, kiss know, land, lie, lie, lie, like, listen, listen look, look, look, look, look, look, look, open, place, point, press, press, press, push, put, save, save, say, say, says, says, send, send, send, send, show, show, sit, sit, sit, sit, sneek, stand, stay, take, talk, tickle, try, turn</i>
	displaced	-ing	<i>flashing, flashing, eating, flashing, flashing, going, talking, coming, saying, coming, flashing, looking, telling, chatting, calling, taking, flying, coming, feeling, feeling, feeling, saying, checking, making, giving, coming, going, working, dying, going, dying, going, dashing, singing, thinking, dashing, telling, telling, singing, going, cleaning, coming</i>
		bare verb forms	<i>flashes, stop, bright, stop, come, play, think, catch, do, do, mends, go, funny, need, go, work, fall, come, do, wash, shut, try, look, found, hide, see, think, forgot, wait, allowed, gave, tell, thought, look, loves, shown, gone, look, see, sleep, call, like, watch, come, caught, send, come, save, like, move, got, put, look, know, show, hear, turn, look, got seen, get, wakes, sleep, forgot, show, seen, play, know, came, pushed, press, get, hear, find, walk, see, stand, show, come, see, show, talk, sleep, say, know, heard, want, look, give, look, likes, have, hope, eat, check, help, give, sent, lie, help, come, fly, make, wash, help, look, worked, help, can, hear, loves, looks, come, play, crashed, fallen, look, got, think, hurt, speak, kick, help, fallen, cheer, bring, put, sing, say, save, go, save, know, go, get, look, see, help, take, check, says, hurts, come, do, help, open, fix, lift, open, fix, get, bring, put, take, collect, send, collect, go, feel, go, look, pick, see, like, look, bring, stop, help, collect, look, gone, think, help, look, need, know, tell, said, do, listen, tell collect sing remember stuck shout pull remember know put wind pull eaten think know sang know think look point care like eaten know told loved kill said play play drive put come get pretend think get love go speak come come play eat</i>

3;o	grounded	-ing	<i>doing, taking, giving, ringing, joining, going, putting, opening, giving, sticking, coming, going, taking, collecting, driving, coming</i>
		bare verb forms	<i>open, say, kiss, open, say, say, look, goes, swipe, put, see, put, put, put, open, found, find, give, give, swipe, do, give, keep, put, weigh, look, look, look, try, see, look, put, look, look, think, get, open, stick, stick, stick, stick, look, look, open, look, see, put, open, pour, eat</i>
	displaced	-ing	<i>running, making, crying, trying, delivering, stealing, snowing, posting, playing, holding, hanging, getting</i>
		bare verb forms	<i>move, show, look, see, found, play, see, go, give, buy, show, made, hang, show, look, post, put, see, buy, put, do, make, hold, put, ring, hang, look, goes, take, rob, come, buy, buy, bought, take, sign, go, buy, found, use, change, use, ring, go, buy, gives, make, go, give, sign, put, look, buy, go, done, went, sang, covered, look, went, make, sell, get, put, collect, go, put, pay, look, get, pull, come, shut, buy, get, buy, bought, have, stopped, post, buy, ask, buy, notice, have, think, buy, stamp, send, walks, says, buy, use, fall, get, open, stick, seen, do, fell, hold, try, put, think, get, waste, gone, put, left, gone, think, bites, collect, put, seen, take, look, sell, see, looked, say, sells, look, came, pushed, talked, pushed, said, gave, see, gives, put, fell, cry, come, thrown, look, see, gone, pull, bought, come, carries, taken, do, think, laugh, said, seen, come, put, blow, eat, blow, popped</i>
