Proceedings of the LSA

Telicity in context: Evidence from Mandarin-speaking children and adults

Jingying Xu & Cristina Schmitt*

Abstract. This study examines how Mandarin-speaking children and adults evaluate telic descriptions in incomplete situations. It focuses on incremental-theme predicates with verbal *-le* and tests whether truth-value judgments depend simply on event (in)completeness or on the interaction between visual event structure and determiner type of the direct object. Mandarin-speaking adults and children aged 4–6 judged sentences describing events in which either all three objects were partially affected (Each Object Incomplete, EOI) or two objects were fully affected and one was only partially affected (Third Object Incomplete, 3rdOI). The direct object appeared with either a numeral or a demonstrative determiner. Participants across age groups rejected incomplete events more often in 3rdOI than in EOI, showing that incomplete situations were not treated uniformly. Adults also showed a determiner contrast in 3rdOI, accepting demonstrative descriptions more readily than numeral descriptions. Children showed a developmental progression toward this pattern: all child groups were sensitive to Visual Completion Type, but only 6-year-olds showed a clear determiner contrast in 3rdOI. The results suggest that Mandarin telicity judgments are shaped by context-sensitive verification rather than by (in)completeness alone, and that developmental differences reflect how children map sentence meaning onto visual event representations.

Keywords. telicity; non-culmination; Mandarin; first language acquisition; truth-value judgment; event representation

1. Introduction.

1.1. TELICITY AND EVENT EVALUATION. We use language to describe events in the world around us. The same event can often be described in more than one way. For example, a scene in which a girl ate three cookies may be described as *She ate cookies*, *She ate three cookies*, or *She ate those cookies*. To evaluate whether a sentence truthfully describes a scene, a listener must map the event representation encoded by the sentence onto an event representation constructed from visual information and judge whether they align. In other words, we are not evaluating a linguistic representation directly against a visual scene, but to a representation of an event that is in our minds.

If the task is complicated for adults, for the learner, whose linguistic knowledge may be incomplete, the task is more complex as it adds a potentially incomplete knowledge of the linguistic pieces and/or how they should be combined. Children must not only use the linguistic ingredients to build the semantics of the sentence, but also determine how this information should be evaluated in context. For example, to build telic predicates (which arguably entail a logical end point), two main linguistic ingredients play a role: the verb and the object. The so-called *incremental-theme verbs*, such as *eat* and *build*, must combine with their objects in such a way

* We thank the members of the MSU Language Acquisition Lab for helpful discussion and feedback on this project. We also thank audiences at the 2026 LSA Annual Meeting for valuable comments on the earlier presentation of this work. We are also sincerely grateful to the child and adult participants for their time and participation, and the experiment helpers in China for their assistance with recruitment and data collection. Authors: Jingying Xu, Michigan State University(xujing21@msu.edu) & Cristina Schmitt, Michigan State University(schmit12@msu.edu).

that the object can be seen as determining whether the description includes a logical endpoint or not (Verkuyl 1972, 1993; Krifka 1989, 1992, 1998; Dowty 1991; Rothstein 2004; Borer 2005; Pin 2008; Filip 2008). When the object is quantized, as in *eat three cookies* or *eat those cookies*, the predicate is telic: it describes an event with a logical endpoint, and in the non-progressive use, it is typically understood to entail that the event reached its logical enpoint , i.e., the event has culminated. By contrast, when the object is non-quantized, as in *eat cookies*, the predicate is atelic, because no upper bound is provided for the event.

A large crosslinguistic literature following the seminal experimental work by van Hout (1998) has shown that children do not always behave in an adult-like way when judging whether a sentence entails event completion or not (English: Wagner 2001, 2009; Ogiela 2007; Jeschull 2007; Anderson 2017; van Hout et al. 2017; Patt et al. 2020; Mandarin: Chen 2006, 2017; Liu 2018; Li 2019; Russian: Kazanina & Phillips 2007; Spanish: Garcia del Real Marco 2015; van Hout et al. 2017; German: Schulz & Penner 2002; Wittek 2002; Schulz & Wittek 2003; Polish: van Hout 2008; Dutch: van Hout 2008; van Hout et al. 2017; Italian: van Hout 2008; Basque: van Hout et al. 2017; Hebrew: Hacohen 2010; see also van Hout 2018; Martin et al. 2020 for reviews). Explanations for this mismatch differ. Some accounts attribute it to lexical semantics, arguing that children initially assign different meanings to the relevant verbs or verb classes (Genter 1978; Wittek 2002; van Hout 2018; Li 2019). Others appeal to pragmatic development, proposing that children know the relevant semantic ingredients but have difficulty selecting the contextually appropriate interpretation or integrating linguistic and non-linguistic information under task demands (Jeschull 2007; Ogiela 2007; Anderson 2017; Liu 2018; Martin et al. 2020; Mathis & Papafragou 2021). Still others point to broader cognitive factors such as executive control (Blank et al. 2019). The present study adopts the view that children may already possess the core semantic ingredients needed to represent telic event descriptions, but differ from adults in how flexibly they verify those descriptions against visually presented scenes. We investigate this question through the case of Mandarin.

1.2. COMPETING ACCOUNTS OF MANDARIN NON-CULMINATION. The claim that telic predicates created with incremental-theme verbs carry a completion entailment has been put to test in Mandarin, where many have claimed that sentences with an incremental-theme predicate and a quantized direct object in perfective form entails event completion remains debated. Many researchers have pointed out that Mandarin allows an apparently telic predicate combined with the perfective marker *-le* (verbal *-le)* to describe an incomplete event that stops before reaching its inherent endpoint (Tai 1984; Sybesma 1997; Lin 2004; Soh & Kuo 2005; Koenig & Lian-Cheng 2008; Martin & Gyarmathy 2019; Martin 2019; Martin et al. 2023; Zhang 2018, 2020). As shown in (1a), denial of completion is often judged felicitous in Mandarin, unlike in the corresponding English example (1b) under the intended telic reading:

(1) a. ta chi-le na-kuai binggan, keshi mei chi-wan.
3SG eat-PFV that-CL cookie but NEG eat-finish
'She ate that cookie, but she didn't finish eating it.'
b. She ate that cookie, #but she didn't finish eating it.

Several accounts have been proposed to explain why apparently telic predicates in Mandarin can be accepted in incomplete situations. These accounts differ in the source of the apparent

lack of completion entailment. The first approach locates the source of non-culminating interpretations in the lexical semantics of Mandarin verbs (Tai 1984; Lin 2004; Sybesma 1997). On this view, monomorphemic verbs such as *chi* 'eat' do not themselves encode a strong culmination condition in the way often assumed for their English counterparts. In acquisition terms, this approach predicts that child–adult differences should reflect differences in the lexical-semantic representations assigned to these verbs (Li 2019). If children initially treat such verbs as result-entailing predicates, they should be more likely to reject incomplete events, and this difference should be relatively robust across contexts.

A second approach shifts the source of the effect from the verb to aspectual morphology, especially verbal *-le*. On this view, the crucial factor is not the lexical semantics of the verb itself, but the contribution of *-le*, which has been argued to mark event termination without necessarily requiring full completion (Martin & Gyarmathy 2019; Martin 2019; Martin et al. 2023). From an acquisition perspective, this view predicts that developmental differences should reflect how children interpret verbal *-le*. If children assign *-le* a stronger completion requirement than adults do, they should reject incomplete events more often, again in a way that is relatively stable across different visual contexts.

In simple clauses such as the example above, it is impossible to isolate the role of *-le* from that of the verb because because removing *-le* yields a generic rather than an episodic interpretation. For the present purposes, however, the two proposals make the same predictions. We therefore treat them together as *no-completion-entailment approaches*.

A third approach retains the idea that Mandarin incremental-theme predicates with quantized objects compositionally encode telic structures and entail event completion with verbal *-le*, and therefore English and Mandarin are not radically different. On this view, differences in the acceptance of telic predicates in incomplete situations reflect differences in how the different linguistic ingredients afford particular evaluations depending on visual and/or pragmatic constraints. In acquisition terms, this approach predicts that children and adults may share the same core semantic ingredients, while differing in how flexibly they verify those meanings against visual scenes. Developmental differences should therefore vary with contextual factors, rather than reflecting a fixed difference in verb meaning or in the semantics of verbal *-le*.

These approaches therefore make distinct predictions for development. Under the *no-completion-entailment approaches*, child–adult differences should primarily reflect differences in underlying semantic representations and should therefore be relatively stable across contexts. Under the *completion-entailment approach*, by contrast, developmental differences should emerge in how children evaluate a telic event description against a scene, and should therefore be sensitive to contextual factors that affect verification.

1.3. THE PRESENT STUDY. The present study provides a way to test these competing accounts. If child–adult differences primarily reflect differences in verb meaning or in the interpretation of verbal *-le*, then judgments should be relatively stable across different kinds of incomplete situations. If, by contrast, children and adults share the same core semantic ingredients but differ in how they verify telic descriptions against visual scenes, then judgments should vary depending on how completion is visually structured and how the direct object guides verification. The current experiment therefore examines whether truth-value judgments are shaped simply by event incompleteness, or by the interaction between visual event structure and the nominal form of the direct object.

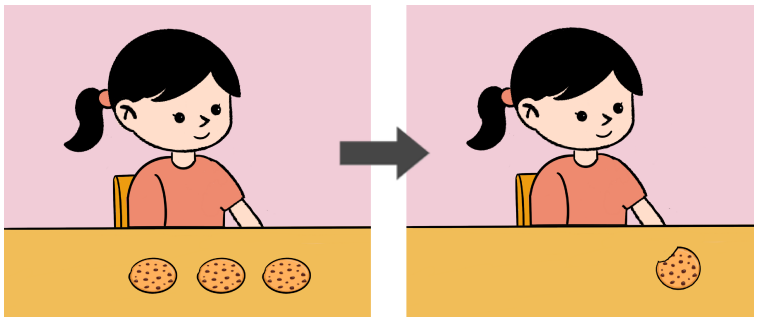

Figure 1. First and last scenes of the video for the 3rdOI cookie-eating event.

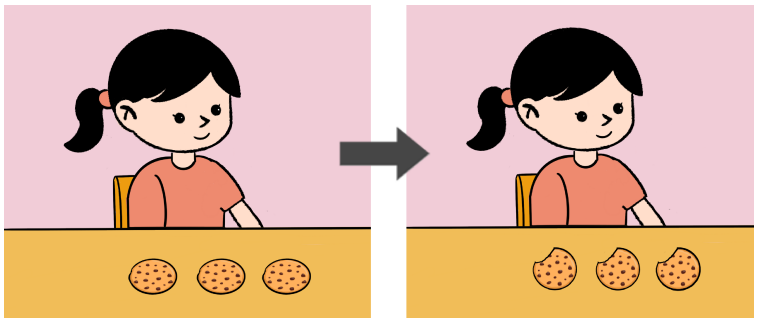

Figure 2. First and last scenes of the video for the EOI cookie-eating event.

To address this question, the experiment contrasts two types of incomplete events. The visual manipulation distinguishes between two kinds of incomplete events. In Third Object Incomplete (3rdOI) events (Figure 1), two of the three objects are fully affected, and the third is only partially affected. In Each Object Incomplete (EOI) events (Figure 2), all three objects are only partially affected. This contrast is theoretically important because the two event types differ not just in completion, but in their internal organization. In 3rdOI, the scene contains an asymmetry: two objects form a fully affected subset, while the third object remains incomplete. This creates an internal contrast within the event between completed and incomplete objects, and also provides a potential standard of comparison for what counts as completion in the scene. In EOI, by contrast, no such subset is available. All three objects are affected to the same partial degree, so there is no completed subset and no internal contrast among the objects. If participants treat all incomplete events alike, then 3rdOI and EOI should pattern similarly. If, however, judgments are sensitive to the how completion is visually distributed, then the two event types may yield different response patterns.

The determiner manipulation provides a further test of how linguistic form interacts with visual information about the event-type and distribution. Numeral and demonstrative DPs place different demands on verification. A numeral DP such as *three cookies* directly encodes the full cardinality of the affected set, and strongly supports an evaluation that includes all three objects. A demonstrative DP such as *those cookies*, by contrast, does not specify cardinality in the same way and may allow more flexible accommodation to the visually salient portion of the scene. This difference becomes especially important in 3rdOI. Because 3rdOI contains a fully affected subset of two objects, a demonstrative description may be more readily verified against that completed subset (excluding the incompletely affected object), whereas a numeral description still requires

verification against the full set of three. In EOI, however, all objects are uniformly incomplete, so no fully affected subset is available to support such accommodation. As a result, a determiner contrast is expected to be strongest in 3rdOI and weaker or absent in EOI.

Taken together, these manipulations make it possible to test whether truth-value judgments are driven simply by the absence of full completion, or whether they depend on the interaction between visual event structure and the semantic properties of the direct object. This contrast is particularly important for distinguishing among competing accounts of Mandarin telicity. Under the verb-semantic account and the *-le*-based account, the observed judgments are not expected to depend strongly on the difference between 3rdOI and EOI or on the interaction between Visual Completion Type and Determiner Type. Under the completion-entailment account, by contrast, such effects are expected, because the crucial issue is how a bounded event description is verified against the scene.

Under the verb-semantic account and the *-le*-based account, the systematic interaction of Visual Completion Type and Determiner Type is not strongly expected, since the source of the relevant interpretation lies in the verb or in *-le* itself. Under the completion-entailment account, however, such an interaction is expected. In particular, adults should distinguish 3rdOI from EOI, and demonstrative descriptions should be more acceptable than numeral descriptions in 3rdOI, where the scene contains a fully affected subset that can support more flexible verification. Developmentally, children are expected to show sensitivity to Visual Completion Type earlier than adult-like sensitivity to Determiner Type, with older children showing a clearer determiner contrast in 3rdOI.

2. Experiment.

2.1. MATERIALS, DESIGN, AND PROCEDURE. The target verbs were incremental-theme verbs including *chi* 'eat', *he* 'drink', *zao* 'build', and *hua* 'draw'. Determiner Type of the direct object (Numeral vs. Demonstrative) was manipulated within subjects. All target sentences in simple clauses with verbal *-le*. as in (2).

(2) a. wo chi le san-kuai binggan.
1SG eat LE three-CL cookie
'I ate three cookies.'

b. wo chi le na-ji-ge xiaodaogao.
1SG eat LE that-several-CL cupcake
'I ate those cupcakes.'

Each test sentence was paired with a complete scene and an incomplete scene. Video Type in the incomplete videos (EOI vs. 3rdOI) was manipulated between subjects. Each participant completed 16 target trials. In each trial, participants first watched a short video clip and then heard a test sentence, together with one filler sentence for children or three filler sentences for adults. Their task was to judge whether each sentence truthfully described the video.

2.2. PARTICIPANTS. Participants were first-language speakers of Mandarin and included children aged 4–6 and a control group of adults. Three 4-year-olds were excluded for failing to meet the 75% accuracy criterion on the filler items. The final sample therefore included 97 children (4-year-olds: $n = 30$, mean age = 4;6, $SD = 0.27$; 5-year-olds: $n = 32$, mean age = 5;6,

$SD = 0.28$; 6-year-olds: $n = 35$, mean age = 6;5, $SD = 0.29$), along with 26 adults (mean age = 29;11, $SD = 11.56$). Because Visual Completion Type (EOI vs. 3rdOI) was manipulated between subjects, participants were distributed across two experimental conditions. None of the participants included in the final analyses had a diagnosed language or hearing impairment at the time of testing. Adults were tested online using JATOS, whereas children were tested in person by an experimenter using a computer.

2.3. RESULTS. As a baseline check, both adults and children showed floor-level rejection in the complete conditions (adults: 2%; children: 2.5%), indicating that participants were attentive and readily accepted telic statements when the events were fully completed. We therefore focus below on responses in the incomplete conditions. Figure 3 presents the group-level rejection rates for adults and children by age.

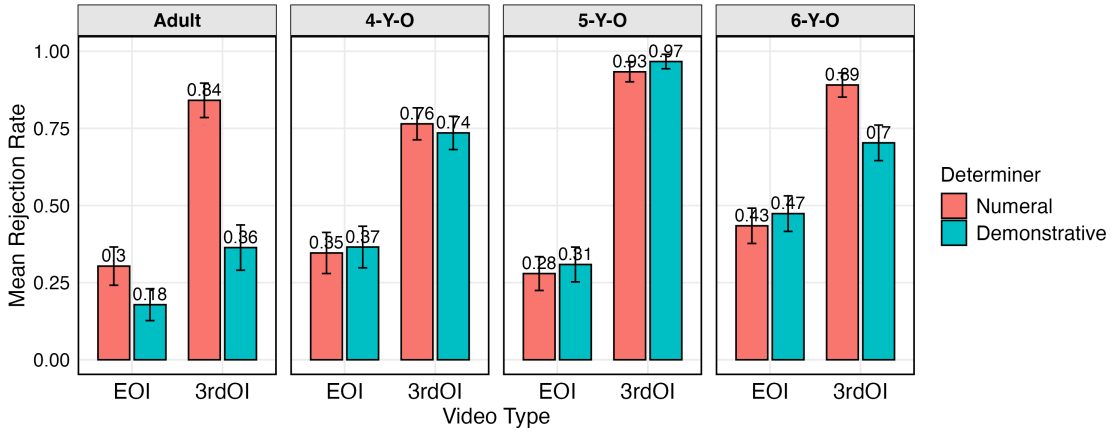

Figure 3. Group-level rejection rates for incomplete situations in for adults and each child age group

2.3.1. ADULTS. In the Each Object Incomplete (EOI) condition, adults generally showed low rejection rates for both determiner types. In contrast, rejection increased in the 3rd Object Incomplete (3rdOI) condition, particularly for numeral descriptions.

To assess these patterns statistically, we fitted a generalized linear mixed-effects model (GLMM) with a binomial link predicting rejection from Video Type, Determiner Type, and their interaction, with a random intercept for subjects. The model revealed a significant effect of Video Type ($\beta = 4.39$, $SE = 1.26$, $z = 3.48$, $p = .0005$), indicating higher rejection in 3rdOI than in EOI. The interaction between Video Type and Determiner Type was also significant ($\beta = -2.47$, $SE = 0.97$, $z = -2.54$, $p = .011$), while the main effect of Determiner Type was marginal ($\beta = -1.11$, $SE = 0.59$, $z = -1.89$, $p = .0588$) (Table 1). Pairwise comparisons further showed that in the 3rdOI condition, demonstrative descriptions were rejected significantly less often than numeral descriptions, whereas no reliable determiner difference emerged in EOI (Table 7). Thus, adults were sensitive both to visual completion and to the determiner contrast, but the determiner effect was restricted to the condition in which only two of the three subevents were complete.

2.3.2. CHILDREN. As in adults, children showed relatively low rejection in EOI and substantially higher rejection in 3rdOI. At the same time, children showed a developmental pattern in

their sensitivity to determiner type, with adult-like determiner effects emerging only in the oldest age group.

A GLMM fitted to the child data, with centered age included as a continuous predictor, revealed a significant effect of Video Type ($\beta = 12.73$, $SE = 3.47$, $z = 3.67$, $p = .0002$), indicating greater rejection in 3rdOI than in EOI (Table 2). The main effect of Determiner Type was not significant ($\beta = 0.49$, $SE = 0.47$, $z = 1.06$, $p = .2914$), but the interaction between Video Type and Determiner Type was significant ($\beta = -1.48$, $SE = 0.69$, $z = -2.13$, $p = .0332$). Neither the main effect of age nor its interactions reached significance, although the three-way interaction between Video Type, Determiner Type, and Age was marginal ($\beta = -0.11$, $SE = 0.06$, $z = -1.89$, $p = .0586$).

To better understand developmental differences, we fitted separate GLMMs for each child age group.

For 4-year-olds, the model showed a significant effect of Video Type ($\beta = 16.52$, $SE = 4.53$, $z = 3.64$, $p = .0003$), but no significant effect of Determiner Type ($\beta = 0.33$, $SE = 0.81$, $z = 0.40$, $p = .6866$) and no interaction ($\beta = -1.07$, $SE = 1.20$, $z = -0.89$, $p = .3723$) (Table 3). Thus, 4-year-olds strongly distinguished EOI from 3rdOI, but did not reliably differentiate numeral and demonstrative descriptions.

For 5-year-olds, the same general pattern held. There was a significant effect of Video Type ($\beta = 16.29$, $SE = 4.17$, $z = 3.91$, $p = .0001$), but neither the main effect of Determiner Type ($\beta = 0.61$, $SE = 0.79$, $z = 0.77$, $p = .4437$) nor the interaction ($\beta = 0.45$, $SE = 1.33$, $z = 0.34$, $p = .7318$) was significant (Table 4). Like 4-year-olds, 5-year-olds showed clear sensitivity to visual completion type, but no determiner-based modulation of rejection.

For 6-year-olds, however, the pattern changed. The model again revealed a significant effect of Video Type ($\beta = 7.16$, $SE = 2.55$, $z = 2.81$, $p = .0049$), but crucially, there was also a significant interaction between Video Type and Determiner Type ($\beta = -3.38$, $SE = 1.07$, $z = -3.15$, $p = .0016$) (Table 5). Pairwise comparisons showed that in the 3rdOI condition, demonstrative descriptions were rejected significantly less often than numeral descriptions ($estimate = 2.824$, $SE = 0.867$, $p = .0011$), whereas no reliable contrast emerged in EOI (Table 7). Thus, by age 6, children began to show the same determiner-sensitive flexibility observed in adults.

Taken together, the age-group analyses reveal a developmental pattern: all child groups were sensitive to the type of event completion, but only 6-year-olds showed clear evidence that determiner type modulated rejection in incomplete situations.

2.3.3. ADULTS VS. CHILDREN. To directly compare adults and children, we fitted a combined GLMM including Age Group (Adult vs. Child), Video Type, Determiner Type, and their interactions. This model revealed a significant effect of Video Type ($\beta = 6.90$, $SE = 2.31$, $z = 2.98$, $p = .0029$), indicating greater rejection in 3rdOI than in EOI, and a significant effect of Determiner Type ($\beta = -1.27$, $SE = 0.63$, $z = -2.02$, $p = .0434$), indicating an overall lower rejection for demonstratives in comparison to the rejection to sentences with numerals (Table 6). The interaction between Age Group and Determiner Type was also significant ($\beta = 1.73$, $SE = 0.75$, $z = 2.33$, $p = .0200$), as was the interaction between Video Type and Determiner Type ($\beta = -3.02$, $SE = 1.14$, $z = -2.64$, $p = .0082$). The three-way interaction was not significant ($\beta = 1.44$, $SE = 1.29$, $z = 1.12$, $p = .2637$).

These results indicate that adults and children patterned similarly in showing greater rejection in 3rdOI than in EOI, but differed in how strongly determiner type affected their judgments.

In particular, adults showed a clearer determiner-based reduction in rejection when compared to children.

Term	Estimate	SE	z	p
Intercept	-1.663	0.797	-2.086	0.0370*
Video Type: 3rdOI	4.390	1.263	3.477	0.0005***
Determiner: Demonstrative	-1.107	0.586	-1.890	0.0588
Video Type × Determiner	-2.472	0.972	-2.544	0.0110*

Table 1. GLMM results for adults.

Term	Estimate	SE	z	p
Intercept	-4.949	2.220	-2.230	0.0258*
Video Type: 3rdOI	12.732	3.474	3.665	0.0002***
Determiner: Demonstrative	0.492	0.467	1.055	0.2914
Centered Age	0.135	0.106	1.279	0.2007
Video Type × Determiner	-1.475	0.692	-2.130	0.0332*
Video Type × Age	-0.047	0.140	-0.333	0.7389
Determiner × Age	0.005	0.039	0.122	0.9026
Video Type × Determiner × Age	-0.114	0.061	-1.891	0.0586

Table 2. GLMM results for children overall.

Term	Estimate	SE	z	p
Intercept	-7.514	2.937	-2.559	0.0105*
Video Type: 3rdOI	16.519	4.533	3.644	0.0003***
Determiner: Demonstrative	0.328	0.812	0.403	0.6866
Video Type × Determiner	-1.068	1.197	-0.892	0.3723

Table 3. GLMM results for 4-year-olds.

2.3.4. INDIVIDUAL RESPONSE PATTERNS. Figure 4 presents individual rejection patterns for adults and each child age group. Each point represents a participants number of rejected trials for numeral sentences (x-axis) and demonstrative sentences (y-axis) within each video type. The diagonal line indicates equal rejection of the two determiner types.

For adults (Panel A), participants in the EOI condition cluster near the lower-left region of the plot, indicating low rejection for both determiner types. In the 3rdOI condition, many participants fall below the diagonal, indicating greater rejection of numerals than demonstratives. This mirrors the group-level pattern and shows that the adult determiner effect was not driven by only a small number of participants.

Notably, although the group-level analysis showed no significant difference between adults and children in the EOI condition, the individual-level plots reveal a bimodal distribution among children. For 4-year-olds (Panel B), participants in the EOI condition cluster around the lower-left and upper-right corners, indicating that many children either rejected both determiner types

Term	Estimate	SE	z	p
Intercept	-7.571	2.493	-3.037	0.0024**
Video Type: 3rdOI	16.290	4.167	3.910	0.0001***
Determiner: Demonstrative	0.605	0.790	0.766	0.4437
Video Type × Determiner	0.454	1.326	0.343	0.7318

Table 4. GLMM results for 5-year-olds.

Term	Estimate	SE	z	p
Intercept	-1.238	1.360	-0.910	0.3628
Video Type: 3rdOI	7.155	2.545	2.812	0.0049**
Determiner: Demonstrative	0.560	0.619	0.904	0.3661
Video Type × Determiner	-3.384	1.073	-3.154	0.0016**

Table 5. GLMM results for 6-year-olds.

or accepted both. In the 3rdOI condition, many points are concentrated near the upper-right corner, showing that many 4-year-olds rejected both numerals and demonstratives. 5-year-olds (Panel C) show a similar pattern: many participants in EOI again cluster near the lower-left and upper-right corners, while responses in 3rdOI are tightly concentrated near the upper-right corner. 6-year-olds (Panel D) show a more mixed pattern. As in the younger groups, many participants in the EOI condition cluster near the diagonal. In the 3rdOI condition, several points remain near the upper-right corner, but more participants also appear below the diagonal, indicating greater rejections of numerals than demonstratives. This pattern suggests that 6-year-olds begin to show the determiner-sensitive flexibility seen more clearly in adults, consistent with the group-level results.

Overall, the individual-level plots confirm that the group-level differences are not driven by a small number of outliers. At the same time, they reveal that childrens responses in EOI were more categorical than the group means alone suggest: many children either accepted all trials or rejected all trials. Thus, while adults showed relatively stable high acceptance in EOI and a demonstrative advantage in 3rdOI, children showed a more bimodal pattern in EOI and stronger overall rejection in 3rdOI, with adult-like determiner-sensitive flexibility beginning to emerge by age 6.

3. Discussion. The current study investigated how Visual Completion Type and Determiner Type of the direct object affect how children and adults evaluate telic descriptions in incomplete situations. Across all age groups, participants rejected incomplete events more often in 3rdOI than in EOI, showing that incomplete scenes were not evaluated uniformly. This finding is important because it indicates that judgments were sensitive not simply to whether an event was incomplete, but to how incompleteness was distributed across the affected objects. Under the *no-entailment-completion account*, all sentences should be acceptable independent of the type of visual completion. It is therefore not straightforward for this account to capture the systematic sensitivity to Visual Completion Type and Determiner Type observed here.

Furthermore, children and adults showed the same distinction between completion types, suggesting that both groups were sensitive to the contextual structure of the scene in similar

Term	Estimate	SE	z	p
Intercept	-3.348	1.617	-2.071	0.0384*
Age Group: Child	0.917	1.707	0.537	0.5912
Video Type: 3rdOI	6.896	2.313	2.981	0.0029**
Determiner: Demonstrative	-1.274	0.631	-2.020	0.0434*
Child × Video Type	1.370	2.443	0.561	0.5750
Child × Determiner	1.733	0.745	2.326	0.0200*
Video Type × Determiner	-3.015	1.141	-2.642	0.0082**
Child × Video Type × Determiner	1.437	1.285	1.118	0.2637

Table 6. GLMM results for adults and children combined.

Group	Video Type	Contrast	Estimate	SE	p
Adults	EOI	Numeral - Demonstrative	1.107	0.586	0.0588
Adults	3rdOI	Numeral - Demonstrative	3.580	0.803	0.0000***
4-Y-O	EOI	Numeral - Demonstrative	-0.328	0.812	0.6866
4-Y-O	3rdOI	Numeral - Demonstrative	0.740	0.878	0.3992
5-Y-O	EOI	Numeral - Demonstrative	-0.605	0.790	0.4437
5-Y-O	3rdOI	Numeral - Demonstrative	-1.060	1.067	0.3204
6-Y-O	EOI	Numeral - Demonstrative	-0.560	0.619	0.3661
6-Y-O	3rdOI	Numeral - Demonstrative	2.824	0.867	0.0011**

Table 7. Pairwise contrasts of Determiner Type within each Video Type for adults and each child age group.

ways. For adults, 3rdOI-demonstrative descriptions were accepted more readily than 3rdOI-numeral descriptions, while the two EOI conditions patterned more similarly. This pattern is consistent with the predictions of a *completion-entailment account*. In 3rdOI, two objects are fully affected while a third remains incomplete, creating a completed subset that a demonstrative DP can more readily pick out. A numeral DP, by contrast, continues to require verification against the full set of three objects, and therefore remains less acceptable when one member of that set is not fully affected. In EOI, all objects are equally partially affected, so no completed subset is available and no comparable determiner contrast is expected. Because EOI lacks an internal contrast between completed and incompleted objects, participants may also be more likely to treat the event as globally finished, despite the absence of full completion for each individual object.

Children showed a developmental progression toward this adult pattern. Like adults, all child groups were sensitive to Visual Completion Type, rejecting 3rdOI more often than EOI. However, only the 6-year-olds showed a clear determiner contrast in 3rdOI. This suggests that sensitivity to the visual organization of incompleteness emerges earlier than adult-like flexibility in how demonstrative noun phrases can be verified against the same scene. At the same time, the individual-level results suggest that children were not simply noisier versions of adults. In particular, childrens responses in EOI were often bimodal, with many children either accepting or rejecting both determiner types. This pattern is consistent with the idea that children may share the relevant semantic ingredients with adults, but differ in how they implement them during truth-value judgment. In other words, the difference between children and adults may be explained by

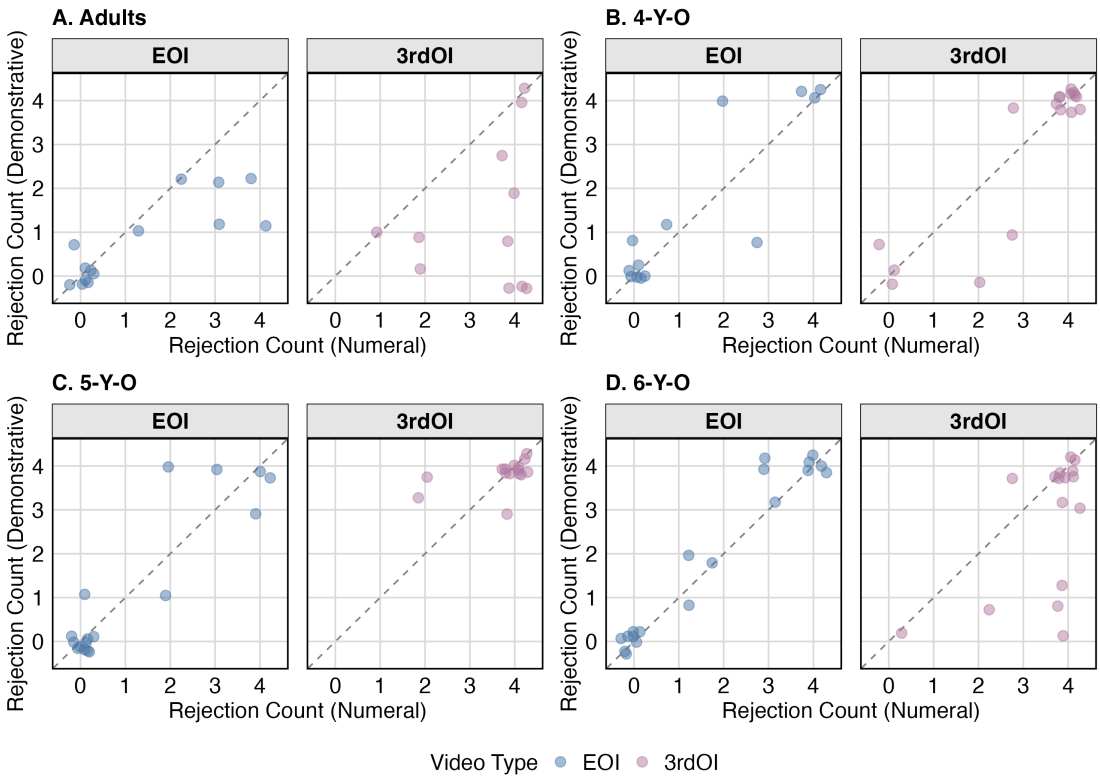

Figure 4. Individual rejection patterns for incomplete situations for adults and each child age group.

the fact that, until around age 6, children do not treat the demonstrative as allowing more accommodation than the numeral in this context. This does not necessarily mean that children do not understand the meaning of the demonstrative. Rather, it suggests that in this situation they still prefer not to construct a subset from the set made available by the scene.

Taken together, this experiment established a baseline pattern in which judgments are shaped by the interaction of visual structure and determiner type, rather than by event incompleteness alone. More broadly, the results support the view that the interpretation of telic descriptions in Mandarin cannot be reduced solely to verb meaning or to aspectual morphology. Instead, they suggest that truth-value judgments reflect how linguistically encoded completion conditions are evaluated against visually represented events. On this view, developmental differences do not necessarily indicate that children assign a different semantics to incremental-theme predicates. Rather, they point to a difference in how children map a bounded event description onto a scene, and in how flexibly they accommodate partial completion when the visual context allows multiple ways of organizing the event.

The present findings therefore favor an analysis on which Mandarin incremental-theme predicates retain a compositional completion entailment, while acceptance in incomplete situations depends on context-sensitive verification. Adults appear able to use both the visual structure of the scene and the nominal properties of the object DP in deciding whether a telic description is satisfied. Children, by contrast, appear to rely more heavily on a stricter completion standard, es-

pecially at younger ages, and only gradually develop the adult-like flexibility that allows demonstrative descriptions to be accepted more readily in contexts such as 3rdOI.

As for the evidence that Mandarin adults may accept telic descriptions in incomplete situations, based on examples such as (1a) (repeated in (3a)), we would like to point out two facts about these cases. First, although the English translation of the example in (3a) seems odd, the Mandarin and English sentences are not fully identical. In Mandarin, there is no overt object for *chi-wan*, and therefore one can interpret the second clause as 'not finish eating' rather than 'not finish eating that cookie', which would be closer to the English tranlation. Second, when tested with standard diagnostics for telicity, the Mandarin sentences do pass the equivalent of the *in-X-time* test for telic predicates, as shown in (4). Finally, it is important to note that English-speaking adults have also been shown to accept telic predicates in incomplete situations, as reported in previous experimental work (van Hout 1998; Ogiela 2007; Ogiela et al. 2014; Anderson 2017; Patt et al. 2020).

(3) a. ta chi-le na-kuai binggan, keshi mei chi-wan.
3SG eat-PFV that-CL cookie but NEG eat-finish
'She ate that cookie, but she didn't finish eating it.'
b. She ate that cookie, #but she didn't finish eating it.

(4) ta shi fenzhong nei chi-le na-kuai binggan.
3SG ten minute in eat-PFV that-CL cookie
'She ate that cookie in ten minutes.'

At the same time, the current study has some limitations. As in all the previous studies, all target sentences contained verbal *-le*, so the present results do not by themselves determine whether *-le* contributes independently to participants judgments. In the current paper, however, the main point is that even with *-le* held constant, participants judgments varied systematically with Visual Completion Type and Determiner Type. This already poses a challenge for accounts that attribute Mandarin non-culmination primarily to verb meaning or to *-le* alone. Future work can further investigate how these factors interact with aspectual marking in a broader range of constructions.

4. Conclusion. To conclude, the present study provides evidence to support a developmental view in which children and adults share the core semantic ingredients relevant to telicity, but differ in how they verify telic descriptions against complex visual scenes. The path toward adult-like interpretation lies not in replacing one lexical semantics with another, but in learning how to apply the same semantic instructions with greater contextual flexibility.

References.

Anderson, Curt. 2017. Contextual factors in childrens calculation of telicity. In *Proceedings of the 41st Annual Boston University Conference on Language Development (BUCLD 41)*, 18–31. Boston University, MA.

Blank, Andrew, Rachael Frush Holt & Laura Wagner. 2019. Inhibitory control and receptive vocabulary influence aspect comprehension in children. *Applied Psycholinguistics* 41(1). 133–151. https://doi.org/10.1017/s0142716419000432.

Borer, Hagit. 2005. *Structuring sense: The normal course of events*, vol. 2. Oxford: Oxford University Press.

Chen, Jidong. 2006. The acquisition of verb compounding in Mandarin. In Eve V. Clark & Barbara F. Kelly (eds.), *Constructions in acquisition*, 111–136. Stanford, CA: CSLI.

Chen, Jidong. 2017. When transparency doesn't mean ease: Learning the meaning of resultative verb compounds in Mandarin Chinese. *Journal of Child Language* 44(3). 695–718.

Dowty, David R. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547–619. https://doi.org/10.2307/415037.

Filip, Hana. 2008. Events and maximalization: The case of telicity and perfectivity. In Susan Rothstein (ed.), *Theoretical and crosslinguistic approaches to the semantics of aspect*, 217–256. Amsterdam: John Benjamins.

Genter, Dedre. 1978. On relational meaning: the acquisition of verb meaning. *Child Development* 49. 988–998.

Hacohen, Aviya. 2010. *On the acquisition of (Hebrew) compositional telicity*: Ben-Gurion University of the Negev dissertation.

van Hout, Angeliek. 1998. On learning the role of direct objects for telicity in Dutch and English. *University of Massachusetts Occasional Papers in Linguistics* 24(1). Article 9.

van Hout, Angeliek. 2008. Acquiring perfectivity and telicity in Dutch, Italian and Polish. *Lingua* 118(11). 1740–1765. https://doi.org/10.1016/j.lingua.2007.08.011.

van Hout, Angeliek. 2018. *On the acquisition of event culmination* 95–121. John Benjamins Publishing Company. https://doi.org/10.1075/tilar.24.05hou.

van Hout, Angeliek, Mara J Arche, Hamida Demirdache, Isabel Garca del Real, Ainara Garca Sanz, Anna Gavarr, Luca Gmez Marzo, Saar Hommes, Nina Kazanina, Jinhong Liu et al. 2017. Agent control and the acquisition of event culmination in Basque, Dutch, English, Spanish and Mandarin. In *Proceedings of the 41st Annual Boston University Conference on Language Development (BUCLD 41)*, 323–332. Boston University, MA.

Jeschull, Liane. 2007. The pragmatics of telicity and what children make of it. In *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America*, 180–187. Somerville, MA: Cascadilla Press.

Kazanina, Nina & Colin Phillips. 2007. A developmental perspective on the imperfective paradox. *Cognition* 105(1). 65–102. https://doi.org/10.1016/j.cognition.2006.09.006.

Koenig, Jean-Pierre & Chief Lian-Cheng. 2008. Scalarity and state-changes in mandarin (and other languages). In Olivier Bonami & Patricia Cabredo Hofherr (eds.), *Empirical issues in syntax and semantics 7*, 241–262. Paris: CSSP.

Krifka, Manfred. 1989. Nominal reference, temporal constitution and quantification in event semantics. In Renate Bartsch, Johan van Benthem & Peter van Emde Boas (eds.), *Semantics and contextual expression*, 75–115. Dordrecht: Foris.

Krifka, Manfred. 1992. Thematic relations as links between nominal reference and temporal constitution. *Lexical Matters* 29–53.

Krifka, Manfred. 1998. The origins of telicity. In Susan Rothstein (ed.), *Events and grammar*, 197–235. Dordrecht: Kluwer Academic Publishers.

Li, Tianshu. 2019. *Incomplete perfectivity and its acquisition in Mandarin Chinese*: The Chinese University of Hong Kong dissertation.

Lin, Jingxia. 2004. *Event structure and the encoding of arguments: The syntax of the mandarin and english verb phrase*: Massachusetts Institute of Technology Ph.d. dissertation.

Liu, Jinhong. 2018. *(Non)-culminating causative accomplishments in mandarin: Experimental evidence*. Nantes: The University of Nantes Ph.D. dissertation.

Martin, Fabienne. 2019. Non-culminating accomplishments. *Language and Linguistics Compass* 13(8). e12346. https://doi.org/10.1111/lnc3.12346.

Martin, Fabienne, Hamida Demirdache, Isabel Garca del Real, Angeliek van Hout & Nina Kazanina. 2020. Childrens non-adultlike interpretations of telic predicates across languages. *Linguistics* 58(5). 1447–1500.

Martin, Fabienne & Zsfia Gyarmathy. 2019. A finer-grained typology of perfective operators. In Christopher Pin (ed.), *Empirical issues in syntax and semantics 12*, 263–290. CSSP.

Martin, Fabienne, Chih-Chen Sun, Fang Liu, Hamida Demirdache & Yicheng Su. 2023. Why one can kill Rasputin twice in Mandarin. In Jin Tanaka & Michael Yoshitaka Erlewine (eds.), *Proceedings of the 25th Meeting of the Austronesian Formal Linguistics Association (AFLA)*, vol. 96, 223–238. Cambridge, MA: MIT Working Papers in Linguistics.

Mathis, Ariel & Anna Papafragou. 2021. Do children use pragmatic (goal) information to compute event culmination? Poster presented at the 46th Boston University Conference on Language Development. Poster presentation.

Ogiela, Diane. 2007. *Development of telicity interpretation: Sensitivity to verb-type and determiner-type*: Michigan State University dissertation.

Ogiela, Diane, Cristina Schmitt & Michael Casby. 2014. Interpretation of verb phrase telicity: Sensitivity to verb-type and determiner-type. *Journal of Speech, Language, and Hearing Research* 57(3). 1–11.

Patt, Raymond, Laura Wagner & Sudha Arunachalam. 2020. The development of a sense of an ending. In *Proceedings of the 44th Annual Boston University Conference on Language Development (BUCLD 44)*, 492–502. Boston University, MA.

Pin, Christopher. 2008. Aspectual composition with degrees. In Louise McNally & Christopher Kennedy (eds.), *Adjectives and adverbs: Syntax, semantics, and discourse*, 183–219. Oxford: Oxford University Press.

Garcia del Real Marco, Isabel. 2015. *The acquisition of tense and aspect in Spanish*: University of the Basque Country dissertation.

Rothstein, Susan. 2004. *Structuring events: A study in the semantics of lexical aspect*. Blackwell.

Schulz, Petra & Zvi Penner. 2002. How you can eat the apple and have it too: Evidence from the acquisition of telicity in German. In J. Costa & M. J. Freitas (eds.), *Proceedings of the generative approaches to language acquisition conference*, .

Schulz, Petra & Angelika Wittek. 2003. Opening doors and sweeping floors: What children with specific language impairment know about telic and atelic verbs. In B. Beachley, A. Brown & F. Colin (eds.), *Proceedings of the 27th Annual Boston University Conference on Language Development*, Cascadilla Press.

Soh, Hooi Ling & Jenny Yi-Chun Kuo. 2005. Perfective aspect and accomplishment situations in Mandarin Chinese. In H. J. Verkuyl, H. de Swart & A. van Hout (eds.), *Perspectives on aspect*, 199–216. Springer.

Sybesma, Rint. 1997. Why chinese verb-le is a resultative predicate. *Journal of East Asian Linguistics* 6(3). 215–261.

Tai, James H.-Y. 1984. Verbs and times in chinese: Vendlers four categories. In *Papers from the 20th Annual Meeting of the Chicago Linguistic Society (CLS 20)*, 286–296. University of Chicago, IL.

Verkuyl, Henk J. 1972. *On the compositional nature of the aspects*. Dordrecht: D. Reidel.

Verkuyl, Henk J. 1993. *A theory of aspectuality: The interaction between temporal and atemporal structure*. Cambridge: Cambridge University Press.

Wagner, Laura. 2001. Aspectual influences on early tense comprehension. *Journal of Child Language* 28(3). 661–681. https://doi.org/10.1017/s0305000901004792.

Wagner, Laura. 2009. Ill never grow up: continuity in aspect representations. *Linguistics* 47(5). https://doi.org/10.1515/ling.2009.037.

Wittek, Angelika. 2002. *Learning the meaning of change-of-state verbs: A case study of German child language*. Berlin: De Gruyter. https://doi.org/10.1515/9783110907988.

Zhang, Anqi. 2018. *On non-culminating accomplishment in Mandarin*: The University of Chicago dissertation.

Zhang, Anqi. 2020. Referentiality, individuation and incompletive readings. *Journal of East Asian Linguistics* 29. 435–468. https://doi.org/10.1007/s10831-020-09201-0.