Cognition 249 (2024) 105811

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Non-signing children's assessment of telicity in sign language

Laura Wagner [a,*], Carlo Geraci [b], Jeremy Kuhn [b], Kathryn Davidson [c], Brent Strickland [b,d]

[a] Ohio State University, USA
[b] Institut Jean Nicod, CNRS, EHESS, ENS-PSL, France
[c] Harvard University, USA
[d] UM6P Africa Business School and School of Collective Intelligence, Morocco

ARTICLE INFO

Keywords:
Sign language
Iconicity
Telicity
Universal grammar

ABSTRACT

Adults with no knowledge of sign languages can perceive distinctive markers that signal event boundedness (telicity), suggesting that telicity is a cognitively natural semantic feature that can be marked iconically (Strickland et al., 2015). This study asks if non-signing children (5-year-olds) can also link telicity to iconic markers in sign. Experiment 1 attempted three close replications of Strickland et al. (2015) and found only limited success. However, Experiment 2 showed that children can both perceive the relevant visual feature and can succeed at linking the visual property to telicity semantics when allowed to filter their answer through their own linguistic choices. Children's performance demonstrates the cognitive naturalness and early availability of the semantics of telicity, supporting the idea that telicity helps guide the language acquisition process.

A central claim of nativist theories of language is that there is a privileged set of conceptual categories around which our linguistic systems are organized. Depending on the specific flavor of nativist theory, these conceptual categories may reflect the fundamental organization of human thought (e.g. Culbertson & Newport, 2015; Hespos & Spelke, 2004; Lakusta & Wagner, 2016) or may have been pre-packaged for the human linguistic system over evolutionary time into a language specific construct, such as Chomsky's Universal Grammar (Chomsky, 1988). Regardless of their origin, the nativist position predicts that these privileged conceptual categories will be cross-linguistically common (if not universal) and will be well integrated into the language acquisition process, typically leading to early learning (Strickland, 2016). Indeed, these predictions are causally linked: by hypothesis, innate conceptual categories are the kernels that guide language development and lead to cross-linguistic commonalities.

One concept that has strong claim for being part of this privileged set is telicity. Telicity refers to the event property of boundedness. Telic predicates describe an event as having an inherent end boundary (e.g. "enter a house" ends with someone inside of a house) while atelic predicates describe an event with no specific endpoint in mind (e.g. "swim" can continue indefinitely). Recent evidence suggests that precisely this kind of event boundedness is called upon in nonlinguistic categorization tasks (Ji & Papafragou, 2022). The formal linguistic properties of telicity have been extensively studied in spoken languages (e.g. Comrie, 1976; Dowty, 1979; Vendler, 1967), and its grammatical reflexes have been identified for a wide variety of languages (e.g. Smith, 1991). Telicity is not encoded in any transparent way across spoken languages but emerges out of a combination of linguistic elements including argument structure, lexical semantics, case marking, determiners, auxiliary selection, and verbal morphology. The specific cluster of elements that contributes to telicity varies across languages but all languages appear to have a means of encoding it, satisfying the requirement for universality.

Sensitivity to telicity also shows the hallmark of being acquired early and showing a distinctive acquisition trajectory. Researchers have documented that children acquiring many different languages are sensitive to the telicity value of a predicate when choosing their tense and grammatical aspect morphology in naturalistic speech (see Li & Shirai, 2000 for a review). Children preferentially use past tense and perfective marking ("Mary entered the house") with telic predicates and present tense and imperfective marking ("Mary is swimming") with atelic predicates. In controlled comprehension studies which more tightly constrain how telicity is signaled than free production can, children show strong knowledge of its grammatical reflexes and interactions by age 4 years. Children can correctly understand that an event is being described as telic (or atelic) and can use different event descriptions to shift their construal of a scene (Wagner, 2006, 2010). Thus acquisition evidence suggests that telicity is an early learned and early used

* Corresponding author at: Department of Psychology, OSU, 1835 Neil Ave., Columbus, OH 43210, USA.
E-mail address: Wagner.602@osu.edu (L. Wagner).

https://doi.org/10.1016/j.cognition.2024.105811
Received 12 September 2023; Received in revised form 15 March 2024; Accepted 3 May 2024
Available online 21 May 2024

conceptual category in language acquisition.

All the work noted so far has been on spoken languages, but researchers (Malaia & Wilbur, 2012; Wilbur, 2008) have documented a potential marker of telicity in American Sign Language (ASL), which also appears in several other sign languages. Specifically, verbs contributing to telic meaning often use ENDSTATE, which involves a distinctive change in the hands (in terms of shape, orientation, or position) from the beginning to the end of a sign. The sign itself instantiates an inherent visual ending which iconically corresponds to the fact that telicity describes events as having an inherent end-point. By contrast, atelic predicates are more often signed using hand motions that are repeated. The open-ended nature of the event is iconically signaled through a visual representation of ongoingness. These iconic markers are probabilistic in nature and admit exceptions (cf. Davidson, Kocab, Sims, & Wagner, 2019).

Strickland et al. (2015) tested the idea that the iconic representation of telicity in sign languages would be visible even to those who did not know a sign language (and see also Kuhn, Geraci, Schlenker, & Strickland, 2021). Across several studies, non-signing participants were shown individual signs from several sign languages and asked to match them to one of two English verbs. The English choices differed in their telicity value (e.g. "die" vs "swim") and thus only one of them matched the sign's (iconic) telicity. Adults successfully matched the telic signs to the English telic verbs at above chance levels and matched the atelic signs to the English atelic verbs (even when neither verbs were actual translations of the sign). These results suggest that the iconic marking of telicity in sign is more than just a fact about sign languages but raises the possibility that there is something cognitively natural about the marking of event boundaries such that we can easily analogize a visual boundary into a linguistic system – even without knowing that system.

Additional evidence for this idea comes from non-linguistic event understanding: even without linguistic descriptions, Ji and Papafragou (2020) have shown that young children can categorize events as being bounded or non-bounded and analyze their structures differently; moreover, adults spontaneously use telicity in categorizing events (Ji & Papafragou, 2022). Cognitive naturalness is consistent with a nativist approach to language, but it further suggests that domain-general cognition (as opposed to something language-specific) might be the true origin of a learning bias in this domain.

Strickland et al. (2015) show that adults can associate gestural stops with the endpoints of events, but the findings do not address the extent to which this ability arises from cognitively natural biases or is learned in development. The current work addresses this question by extending the Strickland et al. (2015) finding to pre-school aged children. The materials and methodological approach are very closely inspired by that work. As noted above, the nativist position argues that privileged semantic elements gain cross-linguistic pervasiveness by biasing the way children learn their native language. By hypothesis, therefore, the iconic instantiation of telicity in sign languages arises from learning pressures. Adults could have succeeded in previous work because they possess sophisticated problem-solving skills that allow them to leverage categories (such as telicity) with which they have long experience. It is therefore important to show that these iconic markers are visible to a population that clearly does understand how telicity is marked in their native language (i.e. children over the age of 4 years old) but is not as likely to be able to draw on explicit strategies in the task. Finding that hearing non-signing children, like their adult counterparts, can use iconic marking to infer telicity meanings would strengthen the case for telicity as part of the privileged set of cognitively natural concepts.

1. Experiment 1: Replicating Strickland et al. (2015) with preschool aged children

Our goal with this set of experiments was to replicate two experiments from Strickland et al. (2015). We presented non-signing children in the United States with videos of 12 signs produced in the sign language of the Netherlands, Nederlande Gebarentaal (NGT) and asked them to match the sign to one of two English words. One of the English words always matched the sign in terms of telicity, but only in Experiment 1A was there a correct translation of the sign itself. All of the videos, general methods, as well as the English word choices in Experiments 1A and 1B were drawn from among those used in the original Strickland et al. work.

1.1. Participants

Participants were recruited from a local science museum and tested in an on-site research lab. Specific demographics were not collected but the overall composition of the pool is 88% white and has a high SES. There were 24 participants in each experiment, just as there were in Strickland et al.'s original study. The specific ages and gender breakdowns are described below. No child participated in more than one experiment.

1.2. Stimuli

Table 1 shows the 12 NGT signs used across all experiments and the English verbs they were paired with in Experiments 1A, 1B, and 1C (full videos of the signs can be seen in our OSF site: https://osf.io/t92r6/). Each sign was produced by the same native speaker of NGT. The classification of the signs as telic vs. atelic was done the same way as in Strickland et al. (2015). All telic signs meet the form criteria laid out for ENDSTATE (Malaia & Wilbur, 2012; Wilbur, 2008): The hands are in different positions from the beginning to the end of the sign. All atelic signs involve repeating the same motion multiple times. The signs were drawn equally from three semantic domains – physical actions, social interactions, and mental actions.

Each trial consisted of a power-point slide that contained a movie showing the target sign and the two English choices printed in large capital letters beneath the movie, on opposite edges of the screen. Although most children in the study could not read well enough to rely on the printed words, their presence assisted the experimenter and generally reinforced the structure of the task for the child by showing two clear choices. The slide for one trial is shown in Fig. 1. The specific English words presented varied by Experiment (see details below). However, in all cases, one word matched the sign in terms of telicity and one did not. Four counterbalancing lists were created and varied the presentation order of the signs, the side on which each English word (left or right) and the specific words used (details below). Full counterbalancing lists are available on our OSF site. Children were sequentially assigned to a counterbalancing list. None of these variations was expected to impact the results but the variability limited the possibility that unanticipated item differences would drive the effects. The full set of signs and English choices is shown in Table 1 and the lists are available on our OSF site.

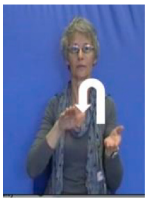
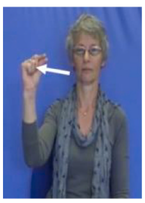
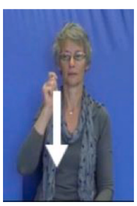
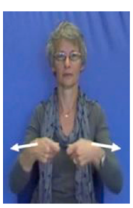
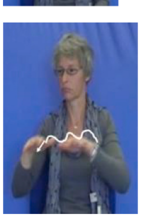
1.3. Procedures

First, all children were asked if they knew what a sign language was; if they did not (most of them did not), they were told that people who could not hear used a language where they talked with their hands which was called a sign language.

Before beginning the experiment itself, children were introduced to the task procedures using four practice items that used videos of signs from American Sign Language (ASL). All of these signs were nouns (*can, candle, hanger,* and *shoe)* and one of the two English word choices was always the correct translation. These items used the identical procedures to the experiment (described next). We did not score these items and the only feedback provided was about paying attention to the movie and listening to the choices.

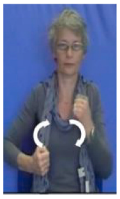
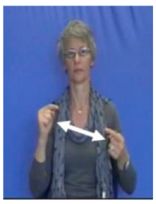
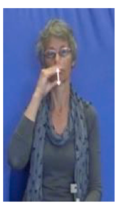
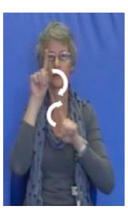
Children then proceeded to the 12 experimental trials. The experimenter recorded children's responses on paper during the experiment

Table 1
Experimental Materials. In Experiments 1a, 1b and 1c, the matching English verb matches the NGT sign in terms of telicity (and counts as the correct answer); the English foil verbs had the opposing telicity value to the NGT sign. In Experiment 1a, the matching verb is a direct translation of the sign; in Experiments 1b and 1c, the matching verb only matches the telicity value of the sign. The English verbs in Experiments 1a and 1b were the same as those used in Strickland et al. (2015); the English verbs in Experiment 1c were chosen as likely to be known well by children. Experiment 2 only used the video materials. The videos used are available on OSF https://osf.io/t92r6/.

NGT Sign	NGT Gloss	Telicity Value	Domain	Experiment 1		Experiment 2		Experiment 3	
				Match	Foils	Matches	Foil	Matches	Foils
	DIE	Telic	Physical	Die	Float Run	Buy Sell	Think	Drop Make Open	Play Swim
	LEAVE	Telic	Physical	Leave	Float Run	Buy Sell	Imagine	Make	Dance Swim
	BUY	Telic	Social	Buy	Discuss Talk	Decide Forget	Float	Build Drop Make	Dance Hug Play
	SELL	Telic	Social	Sell	Discuss Talk	Decide Forget	Run	Find	Love Ride
	DECIDE	Telic	Mental	Decide	Imagine Think	Die Leave	Discuss	Break Build Drop	Dance Hug Ride
	FORGET	Telic	Mental	Forget	Imagine Think	Die Leave	Talk	Break Build Drop Open	Dance Hug Love
	FLOAT	Atelic	Physical	Float	Die Leave	Discuss Talk	Forget	Dance Hug Ride Swim	Break Drop Make Open

(continued on next page)

Table 1 (*continued*)

NGT Sign	NGT Gloss	Telicity Value	Domain	Experiment 1		Experiment 2		Experiment 3	
				Match	Foils	Matches	Foil	Matches	Foils
	RUN	Atelic	Physical	Run	Die Leave	Imagine Think	Sell	Play Ride	Break Build Find
	DISCUSS	Atelic	Social	Discuss	Buy Sell	Imagine Think	Leave	Dance Love Ride	Break Drop Find
	TALK	Atelic	Social	Talk	Buy Sell	Float Run	Decide	Hug Play	Open
	IMAGINE	Atelic	Mental	Imagine	Decide Forget	Float Run	Buy	Love Swim	Break Drop Find
	THINK	Atelic	Mental	Think	Decide Forget	Discuss Talk	Die	Play Swim	Build Make Open

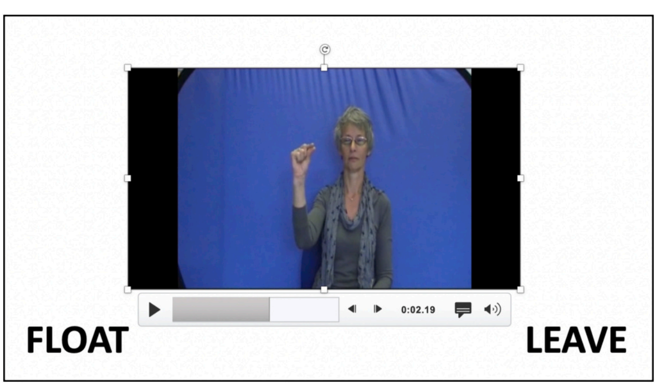


Fig. 1. What was on the laptop screen for a sample trial.

and these responses were checked via video-tape for accuracy. For each item, the experimenter put the trial slide up on the screen of a laptop where the child could see it. The experimenter pointed to each of the printed English words and read each of them to the child, for example: “This one says ‘float’.” In Experiments 1A and 1B, children were asked to provide definitions of each one, for example: “Do you know what ‘float’ means? Can you tell me?” A child was considered to know the words if they provided a sensible definition or, more commonly, could use the word appropriately in a sentence. When needed, a child-friendly definition was provided, for example: “Float is when you rest on top of the water.” Most (over 90%) of children required at least one word to be defined, and the average number of definitions provided was 2.2 (the modal number of definitions provided was 1). In Experiment 1C, children were not asked to provide definitions because the words were all ones they were expected to know well.

The experimenter then played the movie of the NGT sign on the slide and asked the children which English word they thought it meant, for example: “Do you think that word was Leave or Float?” If children were non-responsive or inattentive, the experimenter would repeat the trial (show the movie and ask for a choice) up to three times before moving to the next item.

1.4. Analysis plan

The primary dependent variable was the telicity of the English word children chose on each trial (telic choices were scored as 1 and atelic choices were scored as 0). These data are graphed in Fig. 2. We compared children’s rate of telic choices as a function of the telicity of the sign using the non-parametric Mann-Whitney test. A secondary DV,

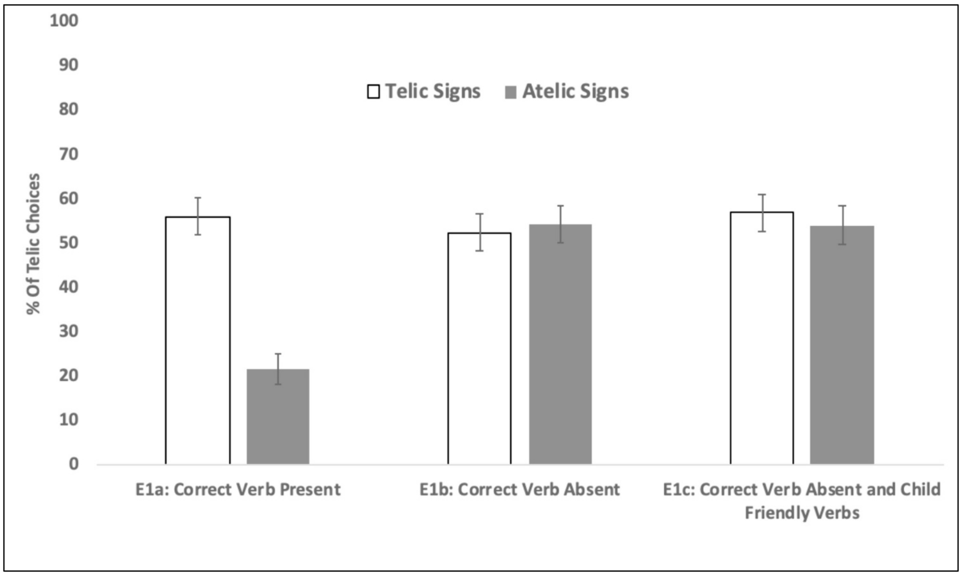


Fig. 2. Percentage of Telic Choices for Experiments 1a, 1b, and 1c.

derived from this one, was accuracy: did children choose the English word that matched the sign for telicity? The primary independent variable was the telicity of the sign. This value was tested using the non-parametric Wilcoxon Signed Rank test. However, to be sure that no spurious factors were responsible for our results, we did a preliminary set of binary logistic regressions using gender and counterbalancing list as additional independent variables. In no analyses was gender a significant predictor of our results. In Experiment 1A, counterbalancing list was a significant factor (children in the most accurate list did significantly better than children in the least accurate list) but there was no interaction between list and our critical variable, the telicity of the sign. Thus, we did not include either gender or list in the main analyses.

1.5. *Experiment 1A*

In this study, one of the two English words presented on each slide was a correct translation for the sign. The incorrect foils for each sign were all correct answers for other signs and matched the same sematic domain as the sign (physical, social, mental). We varied the foils (among two options) across the counterbalancing lists. This study is modeled on Strickland et al.'s Experiment 2.

1.5.1. *Participants*

The mean age of the 24 participants was 5.54 years (range: 4.3–6.5 years; 17 females). One additional child's data was not included because over half the trials were not completed.

1.5.2. *Results*

A Mann-Whitney *U* test showed that children were significantly more likely to choose the telic English word with the telic sign than they were with the atelic sign ($N = 284$, $U = 5.93$, $p < .001$). This pattern also led children to be above chance in choosing the correct translation for the signs by a Wilcoxon Signed rank test ($Z = 5.7$, $P < .05$). Table 2 ranks all of the items according to the percentage of times they were matched to the telic English word. As can be seen there, most individual items were consistent with the overall pattern.

1.6. *Experiment 1B*

In this study, neither of the two English words was a correct translation for the verb. However, one of the words matched the sign in terms of telicity and the other did not. The matching English was chosen from a

Table 2
The NGT signs ordered for each experiment from those generating the least telic responses to those yielding the most. In Experiments 1a, 1b, 1c, the relevant measure is the percentage of telic English word choices; In Experiment 2's feature identification task, the relevant measure is the percentage of "stops suddenly" choices; In Experiment 2's free translation task, the relevant measure is the mean telicity rating of the children's translations for that sign (the scale ran from 1 to 7, with higher numbers indicating more likely to be telic). Atelic signs are in gray.

Exp.1a: Signs (% of Telic Choices)	Exp.1b: Signs (% of Telic Choices)	Exp.1c: Signs (% of Telic Choices)	Exp.2, Features: Signs (% of "Stops Suddenly" Choices)	Exp. 2, Free Translations: Signs (Telicity ratings)
Run (4%)	Talk (38%)	Run (28%)	Think (8%)	Run (2.0)
Talk (13%)	Forget (42%)	Float (33%)	Run (8%)	Imagine (2.2)
Float (21%)	Imagine (46%)	Think (46%)	Discuss (13%)	Think (2.6)
Discuss (26%)	Think (48%)	Buy (46%)	Imagine (25%)	Discuss (2.7)
Forget (33%)	Decide (45%)	Die (50%)	Float (25%)	Float (2.8)
Imagine (33%)	Die (50%)	Sell (54%)	Talk (46%)	Talk (3.8)
Think (33%)	Buy (54%)	Leave (58%)	Die (50%)	Leave (4.8)
Die (48%)	Sell (54%)	Decide (63%)	Sell (50%)	Buy (4.9)
Decide (57%)	Discuss (56%)	Talk (67%)	Leave (71%)	Sell (4.9)
Buy (60%)	Leave (58%)	Forget (71%)	Buy (75%)	Forget (5.2)
Sell (63%)	Run (58%)	Discuss (75%)	Forget (75%)	Decide (5.5)
Leave (75%)	Float (79%)	Imagine (75%)	Decide (79%)	Die (6.4)

different semantic domain (physical, social, mental) than the sign to minimize accidental iconic overlap with the meaning. In addition, we varied the matching words (among two options) across the counterbalancing lists. But to limit the number of lists needed, we used a single incorrect foil word for each sign. This study is modeled on Strickland et al.'s Experiment 4.

1.6.1. *Participants*

The mean age of the 24 participants was 5.22 years (range: 4.8–5.9

years; 13 females).

1.6.2. Results

A Mann-Whitney U test showed that children did not choose the telic English word more with the telic sign than they did with the atelic sign ($N = 286$, U $= 0.48$, $p = .63$). Children were also not above chance in correctly matching the telicity of the English word with the telicity of the sign by a Wilcoxon Signed rank test ($Z = 0.47$, $P = .64$). Moreover, as can be seen in Table 2, there was no consistent pattern across the items.

1.7. Experiment 1C

Although all the children in Experiments 1A and 1B demonstrated reasonable knowledge of the English word choices (as noted above), we worried that their knowledge was not robust enough to support this task. For example, when asked to define "talk" one child said "when you talk you have to use your mouth." While that response makes it clear that the child knows something about the meaning of the word, it is possible that their representation was not sufficiently specific to allow them to abstract telicity out as a feature. The original English words were chosen to be used with adults, and children might benefit from a more child-friendly set of words. We consulted the Macarthur Bates Communicative Development Inventory (Fenson et al., 1991) and chose a set of verbs that were listed as being understood by over 90% of 36-month-olds (see Table 1). Telicity classifications of these verbs were made by the first author. All but one of the child-friendly words came from the physical semantic domain (the exception was "love") and the counter-balancing lists mixed-and-matched the items across trials, preserving only the crucial element that one word was telic and the other atelic. Because these words are typically quite well known by children in the age range we tested, we did not ask children to demonstrate their knowledge of them nor did we provide definitions.

1.7.1. Participants

The mean age of the 24 participants was 5.4 years (range: 4.1–6.9 years; 12 females).

1.7.2. Results

A Mann-Whitney *U* test showed that children again did not choose the telic English word more with the telic sign than they were with the atelic sign ($N = 288$, U $= 0.47$, $p = .64$). Children were also not above chance in correctly matching the telicity of the English word with the telicity of the sign by a Wilcoxon Signed rank test ($Z = 0.471$, $P = .64$). Moreover, as can be seen in Table 2, there was again no consistent pattern.

1.8. Discussion

The results of these three studies show that children had difficulty abstracting telicity per se from the signs. They did succeed in guessing the correct translation of the signs in Experiment 1A at above chance levels – and therefore also succeeded in guessing the correct telicity value of the sign – but they were at chance once the actual translation was missing (1B and 1C). The fact that children could use the real meanings of the signs more effectively suggests that there was some residual iconicity within the signs beyond the telicity information that children were able to perceive and leverage in this forced choice task.

But when forced to rely solely on telicity, as in Experiments 1B and 1C, children failed. Even when the English choices were words that the children had likely been producing for well over a year, they still could not reliably link the telicity in the English words to the properties in the signs. We note that providing children with the easier English items did marginally improve performance, raising the possibility that our sample size was just too small to detect such a weak effect. However, as our concern is about children's appreciation of telicity and not their ability to pass this particular task, in Experiment 2 we opted to take a different tack and make the task itself more open-ended.

2. Experiment 2: Free Translations and Feature Identification

The results from the forced-choice matching tasks cast doubt on children's ability to map the abstract feature of telicity between the signs and the English verbs. In Experiment 2, children were freed from the constraints of the previous studies and asked to offer a translation of each sign of their own choosing. Children had the opportunity to demonstrate their ability to link the visual features of the signs to something in their lexicon, even if that something is not couched precisely in the terms expected based on work with adults. In addition, we also included a forced-choice feature identification task to verify if children had a necessary pre-cursor skill for all of these tasks: Could children perceive the relevant visual feature in the signs at all? The feature identification task is modeled on Experiment 8 from Strickland et al. (2015).

2.1. Participants

The participants were recruited from the same pool as those in the previous experiments. The 24 children (12 female) who participated in this study had a mean age of 5.47 years (range: 4.4–6.2 years).

2.2. Stimuli

The stimuli consisted of the same set of 12 target signs, but they were now presented with no accompanying English choices. The signs were presented in one of two fixed orders. No practice trials were presented.

2.3. Procedures

As in the previous experiment, children were asked if they knew what a sign language was, and it was explained to them if needed. Children completed the free translation task for all signs first. For the free translation task, children were told that the woman in the videos was using sign language to say one word in each video and to guess what the word was. They were reassured that guessing was OK and that there were no right or wrong answers. Children could watch each video up to three times before moving on. Several children who were reluctant to provide translations for the first few videos but warmed as the task proceeded were allowed to go back and provide glosses for the skipped videos. For ambiguous responses, children were asked to use their response in a sentence. All answers were recorded on paper by the experimenter and checked via video.

Once children had provided a translation for each sign, they were asked if they had noticed that some of the signs looked like they "stopped suddenly" while others looked like they were "repeated over and over". These phrases expressed the essential visual properties in terms that were accessible to the child. Children then re-watched the videos (in the same order) and were asked after each one, "Did it stop suddenly or did it repeat over and over?". If children were non-responsive or inattentive, the experimenter would repeat the trial up to three times before moving on. Responses were recorded on a coding sheet.

2.4. Coding

To determine the telicity of the children's translation, each response was coded by two linguists (authors 2 and 3) on a seven point scale, where a 1 indicated that the response was definitely Atelic and a 7 indicated that the response was definitely Telic. The coders were aware of the set of videos used, but were blind to which video the individual translations were in response to. We opted to use this graded scale because the translations themselves were often ambiguous. Children's responses included nouns ("alligator") and adjectives ("silly") and routines (“Good Morning”), as well as verbs (“run”), prepositions, (“up”), partial phrases that included morphology (“riding a motorcycle”), as well as relatively complete sentences (“Alligator eats somebody”). Thus, while telicity classifications are generally made in the literature through the application of classic tests (cf. Smith, 1991; Vendler, 1967) much of what children said did not readily permit the use of those tests as they are traditionally done. Both of our coders are familiar with the range of traditional telicity tests, and they largely agreed on the appropriate rating (their scores differed on average by 0.95, which was less than half of a standard deviation of the overall average rating). Translations which either coder deemed to be “uncodable” for telicity were omitted (a total of 33.9% of all responses) The statistical analyses were conducted over the average of the coders' ratings.

2.5. Results

The results for both tasks are shown in Fig. 3. Note that the feature identification graph (panel A) uses percentage of “stops suddenly” choices as the y-axis, which is parallel to the way the Experiment 1 data were graphed.

2.5.1. Feature Identification Task

Given the similar structure of this task to those in Experiment 1, we analyzed it the same way. A Mann-Whitney *U* test showed that children were significantly more likely to say that the telic sign “stopped suddenly” than they were for the atelic sign ($N = 288$, $U = 7.83$, $p < .001$). This pattern also led children to be above chance in identifying the correct features for the signs by a Wilcoxon Signed rank test ($Z = 16.7$, $P < .001$). Table 2 (column 4) shows the ranking of the items according to the percentage of times children indicated that they “stopped suddenly”. As can be seen there, 4 (of 6) telic signs were matched to that feature over 50% of the time and all 6 atelic signs were matched to the feature of “repeated over and over” over 50% of the time.

2.5.2. Free Translation Task

A selection of representative glosses for the signs is provided in Table 3. The overall mean telicity score for the telic signs was 5.13 (SD: 1.43) and for atelic signs was 2.78 (SD: 1.15). These scores are significantly different from each other ($t\ (23) = 6.48$, $p < .001$). Moreover, the score for the telic signs is significantly higher (that is, more telic) than the mid-point of the scale ($t\ (23) = 5.6$, $p < .001$) and the score for the atelic signs is significantly lower (that is, less telic) than the mid-point of the scale ($t\ (23) = -3.08$, $p < .005$).

Inspection of the glosses themselves revealed that many children were able to correctly guess the true meaning of two atelic signs. Specifically, 14 children (58%) correctly translated the sign for RUN as some variant of “run” or “jog” and 7 children (29%) correctly translated the sign for TALK as some kind of talking (“talk”, “chat”, “blah-blah-blah”). When a child correctly translates a sign into English, she will necessarily correctly match the telicity value of that sign, but it is not clear that telicity is a determining factor in the response. Thus, we removed the two items with high numbers of correct glosses and re-ran the analyses. Since both items were atelic, the overall value of the telic signs remains the same. But even without these items, the score for the atelic signs was still significantly lower (less telic) than the telic signs ($t\ (22) = 6.37$, $p < .001$) and still significantly below the mid-point of the scale ($M = 2.58$; $t\ (22) = -3.45$, $p < .002$).[1]

In addition, as can be seen in Table 2, most signs (11 of 12) generated translations that were coded as matching in telicity, assuming signs above the rating mid-point are classified as telic and those below it are classified as atelic. Moreover, the table also reveals that the atelic sign that generated translations with the highest telic ratings still had a rating that was lower than the telic sign that generated the lowest telic ratings.

2.6. Discussion

Children succeeded with both the feature identification and the free translation task. The fact that children can correctly identify which signs contained the physical features relevant for telic and atelic interpretations suggests that the failures in Experiments 1B and 1C were most likely about difficulties in linking the perceptual feature to the telicity in the English translations, rather than in identifying the feature in the first place.

The strong results in the free translation task suggests further that children are indeed capable of linking those perceptual features to the semantics of telicity in the predicted way. Children's ideas about how to encode telicity in English may be somewhat different than the set-up in Experiment 1 expected, but children could nonetheless connected the sign's visual characteristics to event boundary points.

3. General Discussion

This study investigated if non-signing pre-school aged children, like non-signing adults, could use signs with iconic markers of telicity to infer telicity semantics in their own language. Children's success on the task would provide additional support for including telicity in the native endowment supporting language development. Experiment 1 provided only minimal support for the idea that children could succeed as adults did. Although children did successfully link the signs to their correct translations in Experiment 1A, they failed to make the connection when forced to rely solely on information related to telicity (Experiment 1B), even when that information was encoded in highly familiar vocabulary (Experiment 1C). However, the results from Experiment 2 showed that not only can children identify the relevant markers in the signs, but when allowed to freely translate the signs for themselves, children did preferentially link ENDSTATE forms to telic interpretations and repetitive hand motions to atelic interpretations. These telicity markers are visible and interpretable to these pre-school aged children.[2]

Why did children succeed so robustly in the free translation task and not in the more constrained forced choice tasks? One possibility stems from the specific task demands in Experiment 1. Children may have found it difficult to juggle all the cognitive components of the task. Although they could do the task in a general sense (they did perform above chance in Experiment 1A), the added burden of abstracting the telicity feature out of the English choices and matching to a perceptual feature that is instantiated somewhat differently across each sign may have been more than the children could handle. Using more child-friendly English verbs did help children perform incrementally better, though, so it is possible that children might have succeeded with a sufficiently easy version of the task. Alternatively, it's possible that a much larger sample size would have revealed an ability that produced

[1] A reviewer pointed out that the transparency of RUN and TALK could also be core drivers of the results in Experiment 1A. We therefore also re-ran the analyses in that study taking those two items out. The revised results were qualitatively the same: Children performed above chance overall in the task (Wilcoxon Signed Rank = 3.65, $p < .001$) and they chose the telic English word significantly more with the telic signs than the atelic signs (Mann-Whitney U = 4.17, $p < .001$).

[2] We note that we have replicated the results from Study 2 using the sign language of Italy (LIS). Those data are reported in our OSF data sheet.

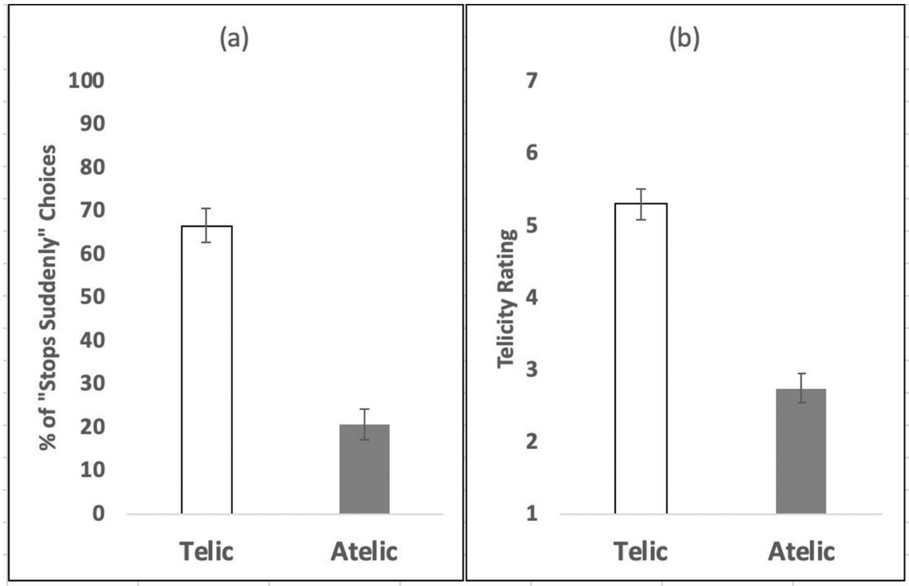


Fig. 3. Results for Experiment 2.
Panel (a) shows the percentage of "Stops Suddenly" choices for the feature identification task. Panel (b) shows the telicity rating given to the child's free translation. Telicity ratings were done on a 7 point scale, with higher ratings indicating that the response was more likely to be telic.

Table 3
Representative translations for each NGT sign given by the children in Experiment 2.

NGT Sign	Telicity Value	Child Translations
DIE	Telic	Flip it over, Flipping a pancake over, Spilled
LEAVE		Shut your mouth, snap
BUY		Pull down, Sit down, Stop it
SELL		Break, Pull apart
DECIDE		Ride, Setting an egg on a pan, Sit down
FORGET		Blow my mind, A chicken popping out
FLOAT	Atelic	Walking like a rabbit, Playing a piano, Hop
RUN		Jogging, Run, Dancing
DISCUSS		Braid, Knitting, Scribble
TALK		Chatting, Eat, Quacking like a duck, Talk
IMAGINE		Think, Wiggle
THINK		Climbing, Drawing, Wiggle

only a small effect.

But the strong success found in Experiment 2 shows that, task difficulties notwithstanding, children are indeed capable of making the mapping between the telicity in the sign and in the English words. We note that the translations provided freely by the children were often quite different from the choices they were given in Experiment 1. Their own translations of the signs rarely involved a single word in citation form (e.g. "float") but instead included particle verbs ("sit down"), lone prepositions ("up"), morphology ("dancing"), and even full phrases and clauses ("a chicken popping out"). These responses suggest that the sparely elegant parsing asked of children in the first three experiments (one verb to one sign) may not correspond well to their natural units of meaning.

And indeed, the children were perhaps not wrong. After all, English cues to telicity include prepositions, direct objects, and even verb morphology. To the extent that children's success requires them to filter their analysis through their own native language, then their spontaneous responses are in fact a better representation of how their language actually does encode the concept of telicity. Once children's language skills have been accommodated, they can build upon their knowledge of telicity in their native language to map entirely novel forms (here, sign language verbs) to these categories (telic/atelic) via the iconicity present in these forms, the same iconicity that seems to drive aspects of their meaning in the larger linguistic context of NGT.

Children's difficulties in Experiment 1 demonstrate that they are by no means expert problem-solvers. Their success in Experiment 2 demonstrates the potency of telicity as a guide for interpreting linguistic input. That strength is precisely what supports the inclusion of telicity among the semantic elements to which our innate language learning capacity is sensitive.

CRediT authorship contribution statement

Laura Wagner: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing, Funding acquisition. **Carlo Geraci:** Conceptualization, Data curation, Methodology, Resources, Validation, Writing – review & editing. **Jeremy Kuhn:** Conceptualization, Data curation, Formal analysis, Methodology, Resources, Validation, Writing – review & editing. **Kathryn Davidson:** Validation, Writing – review & editing, Conceptualization, Methodology, Resources. **Brent Strickland:** Conceptualization, Data curation, Funding acquisition, Methodology, Resources, Validation, Writing – review & editing.

Data availability

We have included our video stimuli and our data spreadsheet in an OSF site.

Acknowledgements

We would like to thank the students who assisted with collecting these data, Aria Dockham, Michelle Indarjit, and Alice Cusick; the NSF REU Site grant (SMA # 1461039) which funded those students; and the Center of Science and Industry (COSI), where the data were collected. We thank Andrea Sims and Annemarie Kocab for their insights into telicity in ASL. We also thank the venues where portions of these data have been presented and the people there who provided feedback: the Language Acquisition Reading Group (Lacqueys) at OSU, the International Association for the Study of Child Language Conference, and the BU Conference on Language Development. This research was also supported by grants ANR-17-EURE-0017 FrontCog and ERC H2020 788077

Orisem.

References

Chomsky, N. (1988). *Language and the problems of knowledge: The Managua lectures*. Cambridge, MA: MIT Press.

Comrie, B. (1976). *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.

Culbertson, J., & Newport, E. L. (2015). Harmonic biases in child learners: In support of language universals. *Cognition, 139*, 71–82.

Davidson, K., Kocab, A., Sims, A. & Wagner, L. (2019). The relationship between verbal form and event structure in sign languages. *Glossa: A journal of general* Linguistics*, 4 (1),* 123. 1–36. DOI: http://doi.org/https://doi.org/10.5334/gigl.924.

Dowty, D. (1979). *Word meaning and Montague grammar*. Dordrecht: Kluwer.

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., … Reilly, J. S. (1991). *Technical manual for the MacArthur communicative development inventories*. San Diego, CA: San Diego State University.

Hespos, S., & Spelke, E. (2004). Conceptual precursors to language. *Nature, 430*, 453–456.

Ji, Y., & Papafragou, A. (2020). Midpoints, endpoints and the cognitive structure of events. *Language, Cognition and Neuroscience, 35*(10), 1465–1479.

Ji, Y., & Papafragou, A. (2022). Boundedness in event cognition: Viewers spontaneously represent the temporal texture of events. *Journal of Memory and Language, 127*, 104353.

Kuhn, J., Geraci, C., Schlenker, P., & Strickland, B. (2021). Boundaries in space and time: Iconic biases across modalities. *Cognition, 210*, Article 104596.

Lakusta, L., & Wagner, L. (2016). Conceptualizing the event: The relationship between infants' representations and linguistic organization. In A. Baron, & D. Barner (Eds.), *Core Knowledge and Conceptual Change: Essays in Honor of Susan Carey* (pp. 245–260). Oxford University Press.

Li, P., & Shirai, Y. (2000). *The acquisition of lexical and grammatical aspect*. Walter de Gruyter.

Malaia, E., & Wilbur, R. B. (2012). Kinematic signatures of telic and atelic events in ASL predicates. *Language and Speech, 55*(3), 407–421.

Smith, C. (1991). *The Parameter of Aspect*. Dordrecht: Kluwer.

Strickland, B. (2016). Language reflects "core" cognition: A new hypothesis about the origins of cross linguistic regularities. *Cognitive Science*. https://doi.org/10.1111/cogs.12332

Strickland, B., Geraci, C., Chemla, E., Schlenker, P., Kelepir, M., & Pfau, R. (2015). Event representations constrain the structure of language: Sign language as a window into universally accessible linguistic biases. *Proceedings of the National Academy of Sciences, 112*(19), 5968–5973.

Vendler, Z. (1967). *Linguistics in philosophy*. Ithaca, NY: Cornell University Press.

Wagner, L. (2006). Aspectual bootstrapping in language acquisition: Telicity and transitivity. *Language Learning and Development, 2*, 51–77.

Wagner, L. (2010). Inferring meaning from syntactic structures in acquisition: The case of transitivity and telicity. *Language & Cognitive Processes, 25*, 1354–1379.

Wilbur, R. B. (2008). *Complex predicates involving events, time and aspect: Is this why sign languages look so similar. Theoretical issues in sign language research* (pp. 217–250).