Object-Event Correspondences Across Languages

Sarah Hye-yeon Lee (sarahhl@sas.upenn.edu)
Department of Linguistics, University of Pennsylvania
Philadelphia, Pennsylvania, USA

Anna Papafragou (anna4@sas.upenn.edu)
Department of Linguistics, University of Pennsylvania
Philadelphia, Pennsylvania, USA

Abstract

Entities in the spatial domain (objects) and the temporal domain (events) are characterized by parallel distinctions that are supported by a shared notion of individuation that runs across domains. This work investigates whether conceptual considerations of individuation are language-independent. We test speakers of English, which uses count-mass syntax and telicity to mark linguistic individuals in the nominal and verbal domain respectively, and Mandarin, which lacks these linguistic features. Our results throw light onto the nature of entity categories in the human mind: both English-speaking and Mandarin-speaking viewers process individuated and non-individuated entities differently, with only the former having a well-defined (temporal/spatial) structure with integrally-ordered, distinct parts. Crucially, these features of non-linguistic individuation are conceptualized in similar ways cross-linguistically and are potentially universal.

Keywords: object; event; individuation; cross-linguistic studies

Introduction

The physical world provides humans with continuous streams of experience in both space and time. The human mind, however, can parse and organize this continuous input into different types of mental units. In the spatial domain, *objects* are understood as fundamental units for many human cognitive processes from infancy to adulthood (Piaget, 1955; Quine, 1960; Scholl, 2001; Spelke, 1988, 1994). In the temporal domain, *events* are considered the foundational entities for human perception and cognition across the human lifespan (Baldwin, Baird, Saylor, & Clark, 2001; Zacks & Tversky, 2001).

Furthermore, closer inspection reveals deep commonalities between these two domains. Within the spatial domain, *objects* (e.g., table, ball) are characterized by well-defined boundaries and, as such, are distinct from *substances* (e.g., sand, water). Within the temporal domain, *bounded* events (sometimes simply referred to as *events*) are internally structured in terms of distinct temporal stages and have a well-defined endpoint (e.g., piling up a deck of cards), unlike *unbounded* events that lack well-defined structure and hence an inherent endpoint (e.g., shuffling a deck of cards) (Ji & Papafragou, 2020, 2022). Importantly for present purposes, these distinctions are linked because they are supported by a shared notion of individuation that runs across domains (see Papafragou & Ji, 2023; Lee, Ji, & Papafragou, in press; for logico-philosophic discussion, see Bach, 1986; Jackendoff, 1991; Taylor, 1977; cf. Champollion, 2015, 2017; Filip, 2012; Truswell, 2019; Wellwood, Hespos, & Rips, 2018).

Specifically, objects and bounded events both qualify as individuals because they possess a well-defined internal structure (including boundaries); by contrast, substances and unbounded events are non-individuals because they lack such an inherently well-defined structure (or boundaries) (see Prasada, Ferenz, & Haskell, 2002 for related discussions on objecthood). Support for this proposal comes from a study by Papafragou and Ji (2023) showing that there is a strong homology between cognitive representations of events and objects. In that study, after brief training, viewers were able to extend categories of bounded or unbounded events to objects or substances, respectively. For example, they were able to extend a category of bounded events (e.g., dress a teddy bear) to also include novel objects (e.g., a solid, ring-like entity), and a category of unbounded events (e.g., pat a teddy bear) to also include novel substances (e.g., a white non-solid mass). Importantly, viewers were able to draw such connections between events and objects even in the absence of prior training. Thus, the cognitive representations of bounded/unbounded events and objects/substances seem to be strongly aligned.

Further experiments revealed that objects and events show specific underlying signatures of individuation (Lee et al., in press). For instance, across domains, individuated entities resist restructuring ("**No Restructuring**"), presumably because they are organized within a specific (spatial or temporal) structure that cannot be rearranged. When an image of an object (e.g., a vase) was edited to depict changes in structure, viewers were more likely to notice it than when an image of a substance (e.g., some clay) was edited in the same way. Similarly, when a video depicting a bounded event (e.g., folding a handkerchief) was edited, viewers were more likely to notice it than when a video depicting an unbounded event (e.g., waving a handkerchief) was edited. Similarly, across domains, individuated entities alone have distinct parts ("**Distinct Parts**"): in Lee et al. (in press), viewers perceived subparts of individuated entities (e.g., two different parts of a vase or two different parts of a folding event) as more distinct from one another than subparts of non-individuated entities (e.g., two different parts of some clay or two different parts of a waving event).

Current study

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

An issue left open by previous studies is whether mental individuation of spatial or temporal entities might arise through (or be affected by) one's native language. The studies summarized above - including the object-event correspondence studies in Papafragou and Ji (2023) and Lee et al. (in press) (see also Wellwood et al., 2018) - have only tested English speakers. In count-mass languages like English, count syntax (e.g., *a table*) (as opposed to mass syntax, e.g., *wood*) provides a cue to individuation (Bloom, 1999; Gordon, 1985; Link, 1983)[1]: Speakers of English know that *a table* refers to a discrete individuated entity, as opposed to some arbitrary portion of a table. Furthermore, in English, telic predicates with verbs denoting an action leading to natural endpoint (e.g., *fold a handkerchief*) can describe bounded events, while atelic predicates with verbs denoting actions that lack a natural endpoint (e.g., *wave a handkerchief*) describe unbounded events (Jackendoff, 1991; Mourelatos, 1978; Parsons, 1990; van Hout, 2016; Vendler, 1957).

Other languages, however, have different means for encoding both distinctions. For instance, Mandarin lacks count-mass syntax, thus all nouns can appear in their bare form (Chierchia, 1998), as in (1), where the bare noun *shū* 'book' in (1) can either denote a single book or plural books.

(1) Zuótiān wǒ mǎi le **shū**
Yesterday I buy ASP books
'Yesterday, I bought one or more books.'
(Rullman & You, 2006, p. 175)

When quantifying nouns, Mandarin uses classifiers both with nouns that refer to individuals (e.g., sān běn shū 'three CL book') and those that do not (e.g., sān píng jiǔ 'three CL wine'). In the verbal domain, while English speakers can use different verbs (e.g., *fix vs. drive a car*) to denote boundedness contrasts, in Mandarin, mono-morphemic verbs (e.g., *kai* 'drive') are generally inherently unbounded (Lin, 2004; Sybesma, 1997). In Mandarin, resultative verb compounds (e.g., *da-po* 'hit-break'), which comprise of a second verb (*po* 'break') that describes the end state associated with an event described by the first verb (*da* 'hit'), are pervasively used to form telic verb phrases. Furthermore, because a bare noun phrase in Mandarin does not carry any information about the quantity of its referent, it cannot delimit events. A bare noun (e.g., *sha-zi* 'sand') and a mono-morphemic verb (e.g., *wan* 'play') form atelic verb phrases denoting an activity (e.g., *wan sha-zi* 'play with sand'). These cross-linguistic differences raise the possibility that Mandarin speakers might have different individuation patterns in non-linguistic tasks compared to English speakers.

Whether speakers of classifier languages like Mandarin may think differently about entities relative to speakers of mass-count languages such as English has been debated (Barner, Inagaki, & Li, 2009; Imai & Mazuka, 2003; Li, Dunham, & Carey, 2009). Available evidence seems to suggest that linguistic individuation does not limit object construal (e.g., Li et al., 2009; for background, see Imai & Gentner, 1997). The extent to which event cognition is susceptible to linguistic effects is also a highly debated topic (for different perspectives, see Choi & Bowerman, 1991; McDonough, Choi & Mandler, 2003; Landau, 2022; Konishi, Brezack, Golinkoff, & Hirsh-Pasek, 2019; Papafragou, Hulbert, & Trueswell, 2008; Sakarias & Flecken, 2019, among many others).

Here we investigate whether (un)individuated entities across both object and event domains are conceptualized in similar ways cross-linguistically and contribute to this discussion. Specifically, we compare conceptions of objecthood and eventhood (and their correspondence) in speakers of English and Mandarin Chinese using the paradigm of Lee et al. (in press). If conceptual individuation is a product of the human mind prior to (and independently from) language, speakers of English and Mandarin should both abide by the No Restructuring and Distinct Parts principles (just as in the original findings from Lee and colleagues.). If, however, conceptual individuation arises from or is shaped by the way languages encode individuation, only English speakers, but not Mandarin speakers, should be sensitive to the No Restructuring and Distinct Parts principles. We test these predictions in Experiments 1 (No Restructuring; 1a: Objects, 1b: Events) and 2 (Distinct Parts; 2a: Objects, 2b: Events).

Experiment 1: No Restructuring

Recall that the No Restructuring principle predicts that observers should find restructurings to individuated entities (objects, bounded events) more noticeable than those to unindividuated entities (substances, unbounded events). If conceptual individuation arises independent of language, both English and Mandarin-speaking adults should show sensitivity to the No Restructuring principle. If, however, conceptual individuation is a product of linguistic encoding, only English speakers should show sensitivity to the No Restructuring principle.

Experiment 1a: Objects

Participants 40 adult native speakers of English and 40 adult native speakers of Mandarin Chinese participated. (In all experiments reported in this study, Mandarin-speaking participants were recruited from Prolific and from the Beijing Institute of Technology. We included language questions on Mandarin to ensure that they were indeed native speakers of Mandarin. All instructions in the Mandarin version of the experiment were in Mandarin.)

Stimuli Following Lee et al. (in press), we used sixteen pairs of images, each depicting a familiar object (e.g., crystal swan) and substance (e.g., crystal). In ten pairs, the object was the artifact made from the substance counterpart (e.g., vase-clay)

[1] It is important to note, however, that there is not a one-to-one correspondence the object vs. substance distinction and count vs. mass syntax (e.g., Barner & Snedeker, 2005; Barner, Wagner, & Snedeker, 2008).

and in two pairs, the object was a natural kind and the substance was an artifact made from the object counterpart (e.g., onion-chopped onion). In the remaining four pairs, both the object and the substance were artifacts (e.g., roll of toilet paper-toilet paper). We created spatially restructured versions of each entity by switching the positions of the second and third vertical strips of the image (see Table 1). Images were edited using the Adobe Photoshop 2022 software. All images were in 400 x 400 pixel dimensions.

Table 1: Sample images in Experiment 1a.

	Original	Restructured
Object		
Substance		

The original stimuli came from a pool of images that were normed in a manner similar to Li et al.'s (2009) Experiment 3, where participants were asked to rate the entities in their original (not restructured) form on a scale of 1-7, with 1 being a good object and 7 being a good substance. The stimuli were rated by 15 naïve native English speakers who did not participate in any of the other experiments reported in this study. Items categorized as objects had a mean rating of 2.62 (SD=2.25), and items categorized as substances had a mean rating of 4.81 (SD=2.31), with people reasonably rating substances higher than objects on our response scale ($t(14) = -7.1$, $p<.001$).

Procedure All experiments reported in this study were hosted online on PennController IBEX (Zehr & Schwarz, 2018; https://www.pcibex.net/), and participants completed them remotely via the internet. In the Mandarin experiments, all instructions were given in Mandarin Chinese. At the beginning of each trial, a fixation cross was displayed for 1000ms. After the fixation cross, the original image was displayed for 100ms. Then, the screen was masked for 3000ms. Afterwards, the restructured image was displayed for 100ms. There was no post-mask after the restructured image. At the end of each trial, participants were asked to identify whether the two items they saw were identical.

Results Results from Experiment 1a are shown in Figure 1. In all experiments reported in this study, the accuracy of the participants' responses was analyzed using Generalized Linear Mixed Effects models (*glmer*). We coded Condition (Object vs. Substance) using centered contrasts (Object=-0.5, Substance=0.5) and included it as the fixed effect. As random

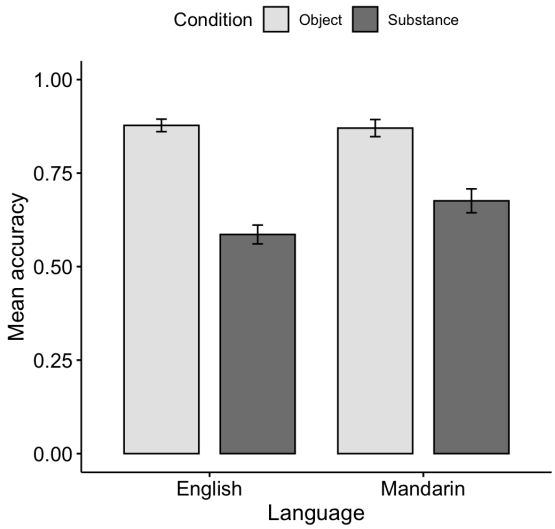


Figure 1: Mean accuracy by language group and condition in Experiment 1a (Error bars represent ±SE).

effects, we entered intercepts for subjects and items, as well as by-subject and by-item random slopes for the effects of Condition. In order to test whether both English- and Mandarin-speaking participants are sensitive to the No Restructuring principle, we conducted these analyses separately for English-speaking participants' and Mandarin-speaking participants' responses.

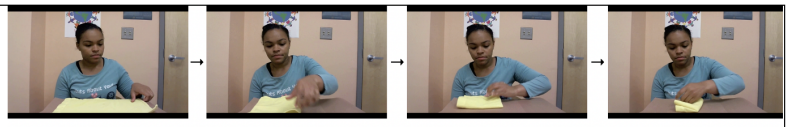
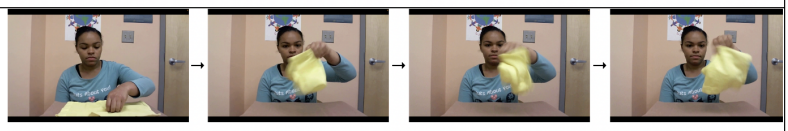
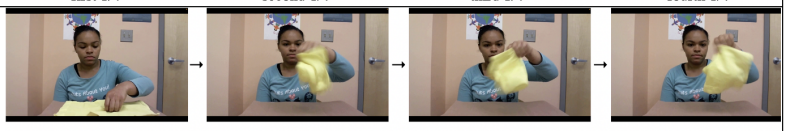
We found that English-speaking participants were more likely to accurately judge that the original and the restructured images were different when presented with Objects (M=87.8%, SD=0.33) than with Substances (M=58.6%, SD=0.49) (β=-1.26, SE=0.53, z=-2.38, p=.018). Mandarin-speaking participants were also more likely to accurately judge that the original and the restructured images were different when presented with Objects (M=87.0%, SD=0.34) than with Substances (M=67.6%, SD=0.47) (β=-1.27, SE=0.26, z=4.87, p<.0001). As expected, both groups of participants were better at detecting restructurings to Objects than to Substances.

Experiment 1b: Events

Participants 40 new adult native speakers of English and 40 new adult native speakers of Mandarin Chinese participated.

Stimuli For target items, we used sixteen pairs of videos from Ji and Papafragou (2020). All videos involved the same girl doing a familiar everyday action in a lab room. Paired videos had the same duration and showed a bounded and an unbounded event. For the sixteen pairs used as target stimuli in Experiment 1b, we created temporally restructured versions of each event by dividing each video into four temporal segments of equal duration and switching the second and third segments (see Table 2). This mirrors the stimuli design in Experiment 1a, where each entity was divided into four segments of equal widths and the second

Table 2: Sample videos in Experiment 1b.

bounded event (fold a handkerchief)	original	 first 1/4 second 1/4 third 1/4 fourth 1/4
	restructured	 first 1/4 third 1/4 second 1/4 fourth 1/4
unbounded event (wave a handkerchief)	original	 first 1/4 second 1/4 third 1/4 fourth 1/4
	restructured	 first 1/4 third 1/4 second 1/4 fourth 1/4

and third segments were switched. Videos were edited using the Adobe Premiere Pro 2022 software.

The original versions of these videos were drawn from a pool of 20 pairs of bounded-unbounded videos in the earlier study (duration range: 4.5s–13s; M=7.98s). That set had been normed to ensure that all video stimuli would illustrate the contrast in boundedness (Ji & Papafragou, 2022): participants (n = 40) judged videos of bounded events as "something with a beginning, midpoint and specific endpoint" 87% of the time but said the same for videos of unbounded events only 21.5% of the time (a significant difference, $t(39) = 20.33$, $p<.001$).

Procedure At the beginning of each trial, a fixation cross was displayed for 1000ms. Once the fixation cross disappeared, participants watched the original video. Then, the screen was masked for 1500ms. Afterwards, participants watched the restructured video. At the end of each trial, participants were asked to identify whether the two videos they watched were identical.

Results Results from Experiment 1b are shown in Figure 2. English-speaking participants were more likely to accurately judge that the original and the restructured videos were different when presented with Bounded events (M=78.5%, SD=0.41) than with Unbounded events (M=60.1%, SD=0.49) (β=-1.26, SE=0.53, z=-2.38, p=.018). Mandarin-speaking participants were also more likely to accurately judge that the original and the restructured images were different when presented with Objects (M=79.4%, SD=0.41) than with Substances (M=63.0%, SD=0.48) (β=-1.58, SE=0.49, z=-3.25, p=.001). As expected, both groups of participants were better at detecting restructurings to Bounded events than to Unbounded events.

Discussion Experiments 1a and 1b showed that both English- and Mandarin-speaking adults were better at detecting

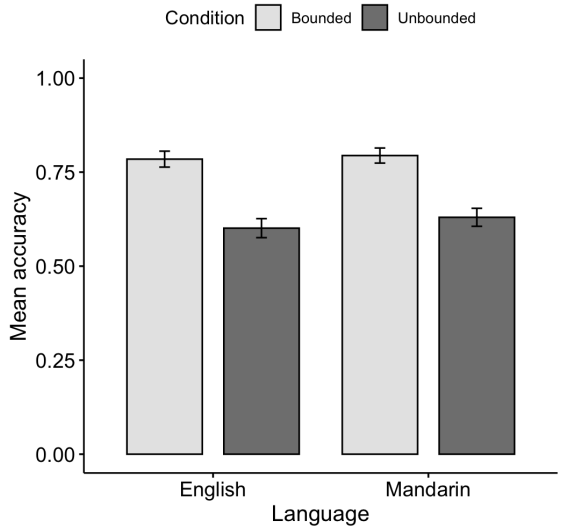


Figure 2: Mean accuracy by language group and condition in Experiment 1b (Error bars represent ±SE).

structural changes to individuated entities (objects and bounded events) than to non-individuated entities (substances and unbounded events), as predicted by the No Restructuring principle. Taken together, these results suggest that regardless of one's native language, the cognitive system is sensitive to conceptual considerations (No Restructuring) of individuation. Therefore, these results support the hypothesis that conceptual individuation is a product of the human mind prior to (and independently from) language.

Experiment 2: Distinct Parts

As alluded to already, the Distinct Parts principle predicts that observers should be better at detecting the difference between two random subparts of individuated entities (objects, bounded events) than the difference between two

random subparts of unindividuated entities (substances, unbounded events). If conceptual individuation arises independent of language, both English and Mandarin-speaking adults should show sensitivity to the Distinct Parts principle. If, however, conceptual individuation is a product of linguistic encoding, only English speakers should show sensitivity to the No Restructuring principle.

Experiment 2a: Objects

Participants 40 new adult native speakers of English and 40 new adult native speakers of Mandarin Chinese participated.

Stimuli Following Lee, Ji, and Papafragou (2023, Under Review), we used the 16 original images used in Experiment 1a to extract two different 80 x 80 pixel parts from each image. One part was extracted from the center of each entity (middle part), and another part from the top right corner of each entity (edge part). See Table 3 for examples. We selected the center and edge parts of each entity to make the parts maximally distinct from each other, for both objects and substances. Moreover, our selection of center and edge parts of spatial entities mirrors our selection of middle and end segments of temporal entities in Experiment 2b.

Table 3: Sample images in Experiment 2a.

	Middle part	Edge part
Object		
Substance		

Procedure The trial structure of Experiment 2a was similar to that of Experiment 1a. At the beginning of each trial, a fixation cross was displayed for 1000ms. After the fixation cross, one part of an entity was displayed for 100ms. Then, the screen was masked for 3000ms. Afterwards, the other part of the entity was displayed for 100ms. There was no post-mask after the other part of the entity. At the end of each trial, participants were asked whether the two images were identical or not. The ordering of the segments was counterbalanced so that in one half of the trials, participants saw the middle segment first, and in the other half, they saw the edge segment first.

Results Results from Experiment 2a are shown in Figure 3. English-speaking participants were more likely to accurately identify the two segements as distinct for Objects (M=74.5%, SD=0.44) than for Substances (M=54.2%, SD=0.50) (β=1.29, SE=0.40, z=3.27, p=.001). Mandarin-speaking participants were also more likely to accurately judge the two segements as distinct when presented with segments of Objects (M=80.3%, SD=0.40) than with segments of Substances (M=59.5%, SD=0.49) (β=1.27, SE=0.41, z=3.09, p=.002).

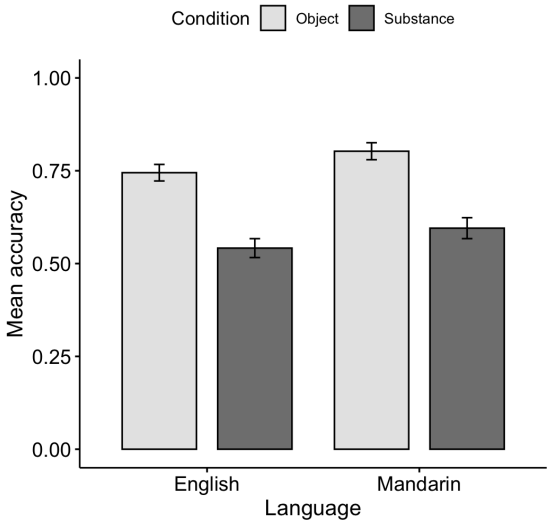


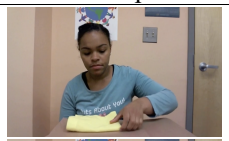
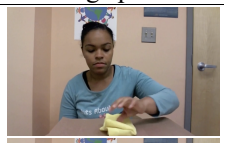
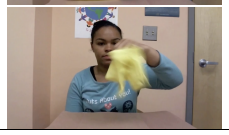
Figure 3: Mean accuracy by language group and condition in Experiment 2a (Error bars represent ±SE).

Experiment 2b: Events

Participants 40 new adult native speakers of English and 40 new adult native speakers of Mandarin Chinese participated.

Stimuli We segmented each original video from Experiment 1b into nine temporal segments, and used the fifth (middle) and the eighth (end) segments (see Table 4.) This mirrors the stimuli design in Experiment 2a, where the middle and edge parts of spatial entities were selected.

Table 4: Sample images in Experiment 2b.

	Middle part	Edge part
Bounded event		
Unbounded event		

Procedure The trial structure of Experiment 2b was similar to that of Experiment 1b. At the beginning of each trial, a fixation cross was displayed for 1000ms. Once the fixation cross disappeared, participants watched a video segment. Then, the screen was masked for 1500ms. Afterwards, participants watched the other video segment. At the end of each trial, participants were asked whether they were identical or not. The ordering of the segments was counterbalanced so that in half of the trials, participants saw the middle segment first, and in the other half, they saw the end segment first.

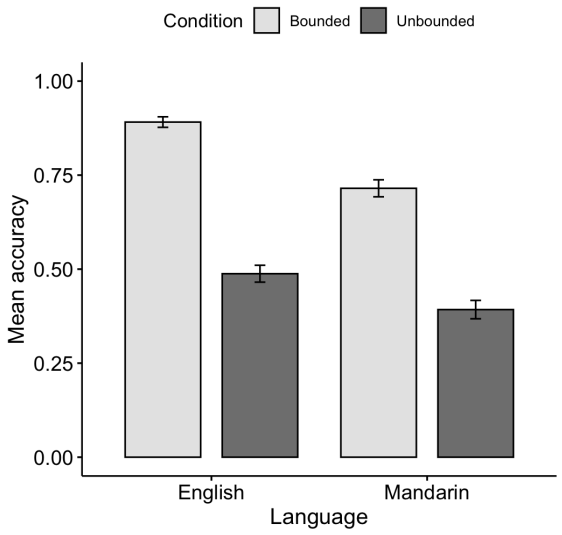


Figure 4: Mean accuracy by language group and condition in Experiment 2b (Error bars represent ±SE).

Results Results from Experiment 2a are shown in Figure 4. English-speaking participants were more likely to accurately identify the two segments as distinct for Bounded events (M=89.1%, SD=0.31) than for Unbounded events (M=48.8%, SD=0.50) (β=3.14, SE=0.76, z=4.12, p<.0001). Mandarin-speaking participants were also more likely to accurately judge the two segments as distinct when presented with segments of Bounded events (M=71.5%, SD=0.45) than with segments of Substances (M=39.2%, SD=0.49) (β=2.20, SE=0.58, z=3.79, p=.0001).

Discussion Experiments 2a and 2b showed that both English and Mandarin speakers were more likely to perceive two subparts of individuated entities (objects and bounded events) as distinct from one another than subparts of non-individuated entities (substances and unbounded events), as predicted by the Distinct Parts principle for individuated entities. These results again suggest that, regardless of one's native language, the cognitive system is sensitive to conceptual considerations (Distinct Parts) of individuation.

General Discussion

In this work, we investigated whether conceptual considerations of individuation are language-independent. We tested speakers of English, which uses count-mass syntax and telicity to mark linguistic individuals in the nominal and verbal domain respectively, and Mandarin, which lacks these linguistic features. In Experiment 1, we showed that both English and Mandarin speakers were sensitive to the No Restructuring principle, according to which individuated entities resist structural restructuring. In Experiment 2, we showed that both English and Mandarin speakers were sensitive to the Distinct Parts principle, according to which individuated entities possess distinct parts. Taken together, these results throw light onto the nature of entity categories in the human mind: both English-speaking and Mandarin-speaking viewers process individuated and non-individuated entities differently, with only the former having a well-defined (temporal/spatial) structure with integrally-ordered, distinct parts. Crucially, these features of non-linguistic individuation are conceptualized in similar ways cross-linguistically and are potentially universal.

These results, combined with earlier findings on object-event correspondences in English speakers (e.g., Lee et al., in press; Papafragou & Ji, 2023; Wellwood et al. 2018), highlight the robustness of the object-event correspondence across languages. Specifically, the common non-linguistic signatures of conceptual individuation across the domains of objects and events are shared across speakers of different languages.

Our findings support the idea that conceptual individuation is a product of the human mind prior to (and independently from) language. These conceptual representations may underpin and structure the linguistic encoding of individuation. Furthermore, early sensitivity to these representations may be used to map entity concepts onto foundational semantics in natural language during language development. This in turn supports the conclusion that an analysis of natural language can reveal meaning distinctions that characterize conceptual systems beyond language; furthermore, it suggests that the explanatory scope of linguistic theory should adjust accordingly so that it is not called upon to explain phenomena that could be in part explained by broader cognitive architecture.

Finally, our study contributes to the debate on the extent to which object and event cognition is susceptible to linguistic effects. We conclude that foundational aspects of object and event cognition, namely the conceptual features of basic units in both domains, are potentially universal in human cognition.

Acknowledgments

This work is supported by the NSF Grant BCS-2041171 awarded to Anna Papafragou.

References

Bach, E. (1986). The algebra of events. *Linguistics and Philosophy*, *9*(1), 5–16.

Barner, D., & Snedeker, J. (2005). Quantity judgments and individuation: Evidence that mass nouns count. *Cognition*, *97*(1), 41–66.

Barner, D., & Snedeker, J. (2006). Children's early understanding of mass-count syntax: Individuations, lexical content, and the number asymmetry hypothesis. *Language Learning and Development*, *2*, 163–194.

Barner, D., Inagaki, S., & Li, P. (2009). Language, thought, and real nouns. *Cognition*, *111*, 329–344.

Barner, D., Wagner, L., & Snedeker, J. (2008). Events and the ontology of individuals: Verbs as a source of individuating mass and count nouns. *Cognition*, *106*(2), 805–832.

Baldwin, D. A., Baird, J. A., Saylor, M. M., & Clark, M. A. (2002). Infants detect structure in human action: A first step toward understanding others' intentions? *Child Development*, *72*, 708–718.

Bloom, P. (1999). The role of semantics in solving the bootstrapping problem. In R. Jackendoff, P. Bloom, & K. Wynn (Eds.), *Language, logic, and concepts: Essays in memory of John Macnamara* (pp. 285–310). Cambridge, MA: MIT Press.

Chierchia, G. (1998). Reference to kinds across languages. *Natural Language Semantics, 6*, 339–405.

Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition*, *41*, 83–121.

Filip, H. (2012). Lexical aspect. In R. I. Binnich (Ed.), *The Oxford Handbook of Tense and Aspect*. Oxford University Press.

Imai, M., & Gentner, D. (1997). A crosslinguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition*, *62*, 169–200.

Imai, M., & Mazuka, R. (2007). Language-relative construal of individuation constrained by universal ontology: Revisiting language universals and linguistic relativity. *Cognitive Science*, *31*(3), 385–413.

Jackendoff, R. (1991). Parts and boundaries. *Cognition*, *41*, 9–45.

Ji, Y., & Papafragou, A. (2020). Is there an end in sight? Viewers' sensitivity to abstract event structure. *Cognition, 197,* 104197.

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

Konishi, H., Brezack, N., Golinkoff, R. M., & Hirsh-Pasek, K. (2019). Crossing to the other side: Language influences children's perception of event components. *Cognition*, *192*, 104020.

Krifka, M. (1989). Nominal reference, temporal constitution and quantification in event semantics. In R. Bartsch, J. van Benthem, & P. van Emde Boas (Eds.), *Semantics and contextual expression*. Foris Publications.

Lee, S. H., Ji, Y., & Papafragou, A. (in press). Signatures of individuation across objects and events. *Journal of Experimental Psychology: General.*

Li, P., Dunham, Y., & Carey, S. (2009). Of substance: the nature of language effects on entity construal. *Cognitive Psychology*, *58*(4), 487–524.

Lin, J. J. (2004). Event structure and the encoding of arguments: The syntax of the Mandarin and English verb phrase (Doctoral dissertation, Massachusetts Institute of Technology).

McDonough, L., Choi, S., & Mandler, J. M. (2003). Understanding spatial relations: Flexible infants, lexical adults. *Cognitive Psychology*, *46*(3), 229–259.

Mourelatos, A. P. (1978). Events, processes, and states. *Linguistics and Philosophy*, *2*(3), 415–434.

Papafragou, A., & Ji, Y. (2023). Events and objects are similar cognitive entities. *Cognitive Psychology*, 104353.

Parsons, T. (1990). *Events in the semantics of English: A study in subatomic semantics*. Cambridge, MA: MIT Press.

Piaget, J. (1955). *The child's construction of reality*. Oxford: Oxford University Press.

Prasada, S., Ferenz, K., & Haskell, T. (2002). Conceiving of entities as objects and as stuff. *Cognition*, *83*, 141–165.

Quine, W.V.O. (1960). *Word and object*. Cambridge, MA: MIT Press.

Rullmann, H., & You, A. (2006). General number and the semantics and pragmatics of indefinite bare nouns in Mandarin Chinese. In *Where semantics meets pragmatics* (pp. 175-196). Brill.

Sakarias, M., & Flecken, M. (2019). Keeping the result in sight and mind: General cognitive principles and language-specific influences in the perception and memory of resultative events. *Cognitive Science*, *43*(1), e12708.

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, *80*(1-2), 1–46.

Spelke, E. (1988). The origins of physical knowledge. In L. Weiskrantz (Ed.), *Thought without language* (pp. 168–184). Clarendon Press/Oxford University Press.

Spelke, E. (1994). Initial knowledge: Six suggestions. *Cognition*, *50*(1-3), 431–445.

Sybesma, R. (1997). Why Chinese verb LE is a resultative predicate. *Journal of East Asian Linguistics,* 6, 215–261.

Truswell, R. (Ed.). (2019). *The Oxford Handbook of Event Structure*. Oxford University Press.

van Hout, A. (2016). Lexical and grammatical aspect. In J. Lidz, W. Synder, & J. Pater (Eds.), *The Oxford Handbook of Developmental Linguistics* (pp. 587-610). Oxford University Press.

Vendler, Z. (1957). Verbs and times. *The Philosophical Review*, *66*, 143–160.

Wellwood, A., Hespos, S. J., & Rips, L. (2018). The object : substance :: event : process analogy. In T. Lombrozo, J. Knobe, & S. Nichols (Eds.), *Oxford Studies in Experimental Philosophy*, Volume 2 (pp. 183-212). Oxford University Press.

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*(1), 3–21.

Zehr, J., & Schwarz, F. (2018). PennController for Internet Based Experiments (IBEX).