Beat gestures help preschoolers recall and comprehend discourse information

Judith Llanes-Coromina a, Ingrid Vilà-Giménez a, Olga Kushch a, Joan Borràs-Comes b, Pilar Prieto a,c,*

a Universitat Pompeu Fabra, 08002 Barcelona, Catalonia, Spain
b Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Catalonia, Spain
c Institut Català de Recerca i Estudis Avançats (ICREA), 08010 Barcelona, Catalonia, Spain

A R T I C L E  I N F O

Article history:
Received 30 January 2017
Revised 12 February 2018

Keywords:
Beat gestures
Prosodic prominence
Word recall
Narrative comprehension
Saliency effect
Child development

A B S T R A C T

Although the positive effects of iconic gestures on word recall and comprehension by children have been clearly established, less is known about the benefits of beat gestures (rhythmic hand/arm movements produced together with prominent prosody). This study investigated (a) whether beat gestures combined with prosodic information help children recall contrastively focused words as well as information related to those words in a child-directed discourse (Experiment 1) and (b) whether the presence of beat gestures helps children comprehend a narrative discourse (Experiment 2). In Experiment 1, 51 4-year-olds were exposed to a total of three short stories with contrastive words presented in three conditions, namely with prominence in both speech and gesture, prominence in speech only, and nonprominent speech. Results of a recall task showed that (a) children remembered more words when exposed to prominence in both speech and gesture than in either of the other two conditions and that (b) children were more likely to remember information related to those words when the words were associated with beat gestures. In Experiment 2, 55 5- and 6-year-olds were presented with six narratives with target items either produced with prosodic prominence but no beat gestures or produced with both prosodic prominence and beat gestures. Results of a comprehension task demonstrated that stories told with beat gestures were comprehended better by children. Together, these results constitute evidence that beat gestures help
preschoolers not only to recall discourse information but also to comprehend it.

© 2018 Elsevier Inc. All rights reserved.

Introduction

Speakers express their thoughts in two dimensions: speech and gesture. Research over the past decades has shown that these two dimensions constitute a single communicative system that is tightly integrated semantically, pragmatically, and temporally (e.g., Kendon, 1980; Levinson & Holler, 2014; McNeill, 1992). In general, across studies, there is strong evidence that iconic and metaphorical gestures (i.e., gestures that either express specific semantic information in their own right or express more abstract ideas) have a positive effect on the recall of information by adults (Riseborough, 1981; So, Chen-Hui, & Wei-Shan, 2012; Thompson, 1995) and children (Cook, Mitchell, & Goldin-Meadow, 2008; Goldin-Meadow, Kim, & Singer, 1999; So et al., 2012; Tellier, 2008) as well as on comprehension processes by adult listeners (Cocks, Morgan, & Kita, 2011; Hostetter, 2011; Kelly, Barr, Church, & Lynch, 1999). In the area of child development, many studies have shown that iconic, metaphoric, and deictic gestures guide children toward the semantic content of a message; that is, they help children comprehend the discourse to which they are listening (Clark, Hutcheson, & Van Buren, 1974; Goldin-Meadow & Wagner, 2005; McNeil, Alibali, & Evans, 2000). Moreover, McNeil et al. (2000) pointed out that reinforcing gestures are an effective scaffold for children’s comprehension of complex spoken messages because they guide comprehension toward the meaning of the spoken language.

Whereas previous work has focused on the facilitation effects of gestures that convey meaning through their form and movement trajectory (iconic and metaphorical gestures) or signal reference (pointing gestures), less is known about the potential effects of beat gestures. Of the different kinds of gestures (see McNeill, 1992), beat gestures are unique in that they are nonreferential, typically co-occur with prosodically prominent positions in speech, and can have a set of discourse pragmatic functions (see Prieto, Cravotta, Kushch, Rohrer, & Vilà-Giménez, 2018, for a review of the properties of beat gestures). Beat gestures are often produced simultaneously with prominent positions in speech and are typically associated with pitch-accented syllables (Yasinik, Renwick, & Shattuck-Hufnagel, 2004). In a quantitative meta-analysis of 63 studies on the overall communicative role of co-speech gestures involving adults and/or children, Hostetter (2011) found strong evidence that gestures foster comprehension in listeners. However, all the gestures analyzed in those studies were iconic, metaphoric, or deictic gestures; none of them included beat gestures.

The current study investigated whether gestures must necessarily carry referential meaning to improve recall and comprehension in preschool children or may fully perform this function even when such referential meaning is absent, as in beat gestures. To explore the effect of context, Austin and Sweller (2014) tested whether beat gestures would help 3- and 4-year-old children and adults to recall spatial directions. In a between-
participants design, both children and adults were given verbal route directions to follow a path through a small-scale spatial array, with the speaker using neither beat gestures nor representational gestures. Although no significant results were found for adults, children recalled the directions better when they were accompanied by either a beat gesture or a representational gesture than when there was no gesture. Nonetheless, further converging evidence is needed to demonstrate that preschool children can exploit the beneficial effects of beats in natural interactions involving not only relevant contextual information but also more complex discourse inputs (see also discussion of Macoun & Sweller, 2016, below). Moreover, a complementary question, which has not been posed by previous studies, is whether children will also remember the information related to the target words better (not only the target words themselves) when these words are associated with beat gestures.

The majority of studies have provided insights into the role that beat gestures play in information recall, yet beat gestures may also play an important role in discourse comprehension processes. Here again, however, the way in which beat gestures are presented may matter. For example, in a study involving preschoolers, Macoun and Sweller (2016) found that iconic and deictic gestures provided comprehension benefits, whereas beat gestures did not. However, the way in which the beat gestures were presented was limited to focused positions. As we will see, in the current study beat gestures were introduced in a way that emphasized not only focused words but also discourse markers such as after, that, and then. Following on other work showing that beat gestures can benefit children when presented in pragmatically relevant discourse contexts (Igualada et al., 2017; Vilà-Giménez, Igualada, & Prieto, 2017), we hypothesized that positive effects might be observed if we exploited the full potential of beat gestures in narratives as pragmatic markers of both focus and discourse structure information.

Importantly, there is unassailable evidence that prosodic prominence alone (e.g., assignment of auditory saliency, generally through pitch accentuation) is beneficial for information comprehension and memorization (Bock & Mazzella, 1983; Fraundorf, Watson, & Benjamin, 2010). For example, Fraundorf et al. (2010) investigated whether the presence of prosodic prominence favored the recall of target items in a contrastive discourse and found that the L + H^* pitch accentuation had a positive effect on the recall of contrastive focus. Yet it is still not completely clear whether the presence of visible beat gestures working together with prosodic prominence represents a further benefit for recall and comprehension processes in comparison with the presence of prosodic prominence alone given that none of the developmental studies above—in particular those by So et al. (2012), Austin and Sweller (2014), and Igualada et al. (2017)—directly compared the effects of prosodic prominence alone with the effects of beat gestures and prosodic prominence used together. One study that pursued this line of research is Kushch and Prieto’s (2016) investigation of how beat gestures and prosodic prominence influence recognition memory (see also Kushch, Igualada, & Prieto, 2018). In their study, they showed that participants performed better in a recall task when the target item was associated with both prosodic and gestural prominence than when it was associated with prosodic prominence alone. In a separate study, Wang and Chu (2013) explored the effects of prosodic prominence alone versus prosodic prominence plus beat gestures, but this time with a focus on information comprehension as measured by brain response. The authors performed an event-related potential (ERP) experiment to assess how pitch accentuation and beat gestures influenced semantic processing. The results showed that the target words accompanied by a beat gesture reduced negativities in the N400 time window more than those accompanied by a control hand movement or no gesture at all, suggesting that beat gestures enhance semantic processing during the comprehension of speech. Interestingly, there was no interaction between beat gesture and pitch accentuation, indicating that prosodic prominence and gestural prominence influence semantic processes independently. We explored this issue further in the first of two experiments described here by controlling for the effects of prosodic prominence on information recall by preschool children in the absence and presence of gestural prominence.

The aim of the current study was to investigate whether nonreferential beat gestures help preschool children in two complementary speech decoding tasks, namely recall and comprehension, by exploiting the discourse properties of beat gestures as highlighters of both information structure and discourse structure (Dimitrova, Chu, Wang, Özyürek, & Hagoort, 2016; McNeill, 1992; Shattuck-Hufnagel, Ren, Mathew, Yuen, & Demuth, 2016). According to Shattuck-Hufnagel et al. (2016), even though beat gestures are devoid of referential content, they can serve important pragmatic and discurs-
sive functions such as focus marking, rhythmic marking, and discourse structure marking. We hypothesized that when focused information and discourse structure information are highlighted by beat gestures, this would help children not only to recall target words better but also to comprehend the content of a story better. The study consists of two experiments. Experiment 1 examined the potential effects of three types of prominence (gestural and prosodic prominence combined, prosodic prominence alone, and no prosodic prominence) on the recall by preschool children of a set of contrastively focused items in short discourses as well as information related to those items. Following recent results by Igualada et al. (2017) and Austin and Sweller (2014), we expected that children would remember the items conveyed in the prominence in speech and gesture condition better than in the nonprominent speech and prosodic prominence conditions. Moreover, we also expected that children would remember the information related to the target item in an optimal way when it was conveyed to them with prominence in both speech and gesture. Experiment 2 examined the comprehension of short narratives presented under two experimental conditions, namely no beat (i.e., with target items produced with prosodic prominence but no beat gestures) and beat (i.e., with target items produced with both prosodic prominence and beat gestures). Following Wang and Chu's (2013) electrophysiological results, we also hypothesized that information featuring focused items in the beat condition would be comprehended more easily. The discourse materials in both experiments were presented in child-directed speech and used a type of beat gesture that is commonly associated with child-directed narratives, namely the movement of both hands, palms open, outward away from the body (see “Experimental materials” section below).

Whereas 4-year-old children were recruited for Experiment 1, 5- and 6-year-old children were recruited for Experiment 2. Our rationale for using 4-year-olds for Experiment 1 but 5- and 6-year-olds for Experiment 2 is linked to the relative complexity of the tasks involved. Experiment 1 consisted of a recall task framed in short contextual discourses involving contrastive focus structures, and Experiment 2 consisted of a comprehension task involving narratives with a more complex narrative structure. In developmental terms, 4-year-olds have been shown to have already acquired the notion of contrastive focus (Chen, in press; Höhle, Fritzsche, & Müller, 2016). In particular, Höhle et al. (2016) showed that 4-year-olds are able to comprehend sentences containing focus particles such as the German particle nur ‘only’. By contrast, it has been shown that narrative abilities appear during the late preschool ages and become apparent only at around 5 or 6 years of age. For example, Demir and Küntay (2014) claimed that 5-year-olds are able to link their utterances only with sequential conjunctions or discourse markers such as then and after, whereas children older than 5 years can use more complex structures. Crucially, 4-year-olds’ narratives are often out of sequence and shorter, hence the greater age of the children who participated in our Experiment 2.

Experiment 1

In a within-participants experimental design, children were presented with three discourse contexts containing a set of contrastively focused items in one of the following three conditions: (1) prominence in speech accompanied by beat gestures, (2) prominence only in speech with no gestures, and (3) nonprominent speech and likewise without gestures. After the presentation of each discourse, children were asked a recall question about the contrastively focused items, followed by another recall question about information related to these contrastively focused items. The experimental design was adapted from Fraundorf et al. (2010) and Kushch and Prieto (2016) to be appropriate for children and to capture both recall and comprehension of information.

Method

Participants

A total of 51 preschool children ($M_{age} = 4.57$ years, $SD = 0.26$) from five schools (Escola Estalella i Graells, Escola Casa Nostra, Escola Pública Joan Bruguera, Escola Bora Gran, and Escola Can Puig) in the provinces of Barcelona and Girona, Spain, took part in our study. All of them were bilingual in Catalan and Spanish but were overwhelmingly Catalan dominant (mean dominance of daily use of Catalan...
relative to Spanish = 89.71%, SD = 9.96). Parents were informed about the experiment’s goal, completed a questionnaire (Bosch & Sebastián-Galles, 2001) about their children’s language use, and signed a consent form allowing us to run the experiment and process the data. All parents reported that their children had exhibited a typical pattern of development and did not have any language disorders.

Materials
Preliminary study. To find out what kinds of beat gestures and intonation are typically used in child-directed discourses containing contrastive information, three adult native Catalan speakers participated in a dramatization task in which they were asked to imagine that they were preschool teachers and needed to tell two very short stories to 3- and 4-year-old children. They were first asked to read and memorize the two stories (see Appendix A). Each story contained two contrastive items, one of which was named again in the last sentence of the story.

After they had read and understood the stories, these adult participants recounted them as if speaking to a group of preschool children. They were asked to place special emphasis on the target contrastive item in the last sentence to make sure that the children would remember it. They were then asked to repeat the procedure, but this time to “overemphasize” the target word.

The dramatization task resulted in a total of 12 video-recordings (3 participants x 2 discourse situations x 2 conditions [emphasized vs. overemphasized]). The recordings of all target contrastive words were later extracted and analyzed for gestural and prosodic information. In terms of their prosodic content, the words were annotated using Cat_TOBI, the Catalan version of ToBI (Prieto et al., 2015). Gestures were analyzed using the Massachusetts Institute of Technology (MIT) Gesture Studies Coding Manual (http://web.mit.edu/pelire/www/manual), specifically following the sections on Hand Shape (open or steepled), Palm Orientation (palms up or palms together), and Trajectory Movement (outward or inward). The results showed that the most common pitch accent in the overemphasized target words was L + H L%. Regarding gestures, all target words were seen to be accompanied by beat gestures in both the emphasized and overemphasized conditions. Although these gestures varied in form and size, the most common beat gesture was a movement away from the body of both hands, palms open, at waist height (see Appendix B).

On the basis of these results, it was decided that gestural and prosodic focal marking in the experimental materials would be carried out using the outward open-palm gesture and the L + H L% nuclear pitch configuration (the typical configuration used for contrastive focus), respectively.

Experimental materials. The materials used in this experiment were adapted from Fraundorf et al. (2010), Kushch and Prieto (2016), and Igualada et al. (2017). They consisted of a set of 36 videos (approximate length of 24–35 s) created on the basis of three short discourses in Catalan (see Example [1] below for the English translation of one of them, and see Appendix C for translations of the full set of three discourses). The two main characters in the three discourses were Elmer, an elephant taken from the series of children’s books about Elmer the Patchwork Elephant by David McKee, and Núria, his female human friend. Because Elmer has trouble remembering things, Núria needs to constantly remind him about what they have done in their trips together. Similar to the short stories in the preliminary study, each discourse context mentioned two alternative pairs of items (e.g., roses ‘roses’ and peixos ‘fishes’; peixos ‘fishes’ and ànecs ‘ducks’ in Example [1]), which were then followed by a continuation sentence that mentioned the contrastively focused item. These contrastively focused words were the target words intended to receive the focus marking (as indicated by the boldface in Example [1]). All target contrastive items were disyllabic words in Catalan with stress on the first syllable and were deemed to be familiar to 4-year-old Catalan-speaking children (58.3% of them can be found in the Catalan version of the MacArthur–Bates Communicative Development Inventories for children aged 16–30 months; Serrat et al., 2010).

[1] Example of a stimulus narrative:
Núria addressing Elmer the elephant: Elmer, do you remember our trip to the field? In the morning, we went for a walk in the field. We walked for many hours! We noticed that near the lake there were roses and leaves, and you picked the roses, which soon wilted! Later on, you saw that in
the lake there were fishes and ducks, and you picked up the fishes. They were very small! We had a great time!

Recall question (researcher asks child participant): Now help Elmer remember what he picked up when he went to the field. What did he pick up? What happened to the roses? What were the fishes like?

The contrastively focused items were counterbalanced according to whether they were the first or second words within each pair of items. Thus, a total of 36 stories (3 discourses × 2 contrastive pairs × 2 possible orders × 3 conditions) were prepared.

Video-recordings were then made of the 36 stories. Recordings took place in an experimental lab at the Universitat Pompeu Fabra with a Marantz PMD660 professional portable digital video-recorder and a Rode NTG2 condenser microphone, and later they were edited with the Adobe Premiere Pro CC video-editing program. Three recordings were made for each discourse, one for each of three conditions, namely with prominence in both speech and gesture, with prominence only in speech, and with nonprominent speech. For prominence in both speech and gesture, the speaker was asked to produce an L + H L% nuclear pitch configuration and an open-palm outward beat gesture on the target item. For prominence only in speech, the speaker was asked to produce an L + H L% nuclear configuration with no beat gestures. For nonprominent speech, the speaker was asked to produce the target items with a neutral L L% nuclear pitch configuration prosody and no hand gestures. The stories were performed by a Catalan native speaker who previously had been trained to produce the target discourses in child-directed speech and with the target words in the three experimental conditions. The speaker was also asked to use neutral facial gestures in a consistent way across conditions, which basically consisted of showing a smiling face. Special care was taken during the recordings to make sure that the target items were produced with the same prosody in prominence in both speech and gesture and prominence only in speech.

Fig. 1. Still images of the speaker saying the word zebres ‘zebras’ in the three conditions: prominence in both speech and gesture (Condition 1, left panel), prominence in speech only (Condition 2, middle panel), and nonprominent speech (Condition 3, right panel).
The three panels in Fig. 1 correspond to still images of the speaker saying the target word *zebras* in the three conditions together with the waveform and pitch contours of the target words. The left panel shows the production of the target word in prominence in both speech and gesture, the central panel shows the production of the word in prominence only in speech, and the right panel shows the production of the word in nonprominent speech.

To check that the speaker used identical prosody in the target items across prominence in speech and gesture and prominence only in speech (both containing prominent speech), and that this differed from nonprominent speech, the target items were acoustically analyzed using Praat software (Boersma & Weenink, 2016). The target words were prosodically coded using Cat_ToBI (Prieto et al., 2015). The H and L target F0 points in every pitch accent were manually coded, as were the starting and end points of the target accented syllables and target words. Pitch range was calculated as the distance in semitones between the minimum and maximum pitches of L + H⁄ and end points of the target accented syllables and target words. Three generalized linear mixed models (GLMMs; West, Welch, & Galecki, 2007) were performed for target word duration, target syllable duration, and pitch range. In all analyses, condition was set as a fixed factor and a random intercept was set for the difference of discourse within each condition. The effect of condition was found to be significant for word duration, F(2, 69) = 3.463, p = .037, and pitch range, F(2, 69) = 16.546, p < .001, but not for syllable duration, F(2, 69) = 2.115, p = .128. For word duration, pairwise contrasts identified a significant difference such that prominence in both speech and gesture > nonprominent speech (β = .103, SE = .039, p = .011), but not between the other two conditions of prominence in both speech and gesture vs. prominence only in speech in both speech and gesture > nonprominent speech, β = .059, SE = .039, p = .139; prominence only in speech vs. nonprominent speech, β = .044, SE = .039, p = .264). For pitch range, pairwise contrasts identified a significant difference between nonprominent speech and the other two conditions (prominence in both speech and gesture > nonprominent speech, β = 43.321, SE = 7.964, p < .001; prominence only in speech > nonprominent speech, β = 34.567, SE = 7.964, p < .001), but not between prominence in both speech and gesture and prominence only in speech (β = 8.754, SE = 7.964, p = .275). As expected, all three measures (pitch range, syllable duration, and word duration) indicated, crucially, that items in the prominence in both speech and gesture condition and in the prominence only in speech condition were prosodically similar. By contrast, although the pitch range measures corroborated the expected distinction between the prominent and nonprominent items, this was not the case for the two duration measures (i.e., syllable and word duration).

Our interpretation of these results is that in our data target stressed syllables without pitch accents (labeled nonprominent speech) have similar durations as stressed syllables with pitch accents (labeled prominent speech) and that the main difference between the two lies in the F0 pitch range patterns (i.e., the presence of a rising pitch accent in prominent speech but not in nonprominent speech). This result is not surprising because previous studies have reported that stressed syllables in unaccented contexts in Catalan are marked (and perceived) with duration cues (see Ortega-Llebaria, Vanrell, & Prieto, 2010).

**Procedure**

The 51 children were tested individually in a quiet classroom at their school, and the entire session was videotaped. During each session, two researchers were present in the classroom; Experimenters 1 interacted with the children and sat next to them in front of the laptop computer, and Experimenters 2 controlled the video-recorder and also observed and coded the children’s responses. The child participants were seated in front of the laptop computer to watch the video-recorded discourses. To control for memory span, all children were administered a memory span task to determine the maximum number of items they could recall. In this test, which followed the same procedure as that used in Igualada et al. (2017), children were asked to repeat a series of word lists produced by the experimenter whereby the first list consisted of one item, the second list consisted of two items, the third list consisted of three items, and so on. This procedure continued until the children could no longer succeed in recalling all the words in the list. Once children’s apparent limit had been reached, they were asked to repeat three lists of the same length to confirm that this word count was indeed the limit of their memory span.

The memory span task was followed by the experiment per se, which consisted of watching the recordings in which the human character Núria reminded Elmer what they had done together. Each
of the discourses was presented to the child participants under one of the three experimental conditions (prominence in both speech and gesture, prominence in speech only, or nonprominent speech) in a within-participants experimental design. Each participant saw a total of three videos (one video per condition), with the children being randomly assigned to different discourse and condition orders. After each of the three videos, two prompt slides with the two recall questions appeared in the PowerPoint presentation, and the experimenter asked the children whether they could help Elmer remember them (see Appendix C for the recall questions corresponding to the target word [Recall Question 1] and information related to the target word [Recall Question 2]). The chart in Fig. 2 summarizes the procedure followed in Experiment 1.

Fig. 2. Experimental procedure for Experiment 1.
Results

In this experiment, the responses to the recall questions were regarded as correct and coded as 1 if the children correctly recalled the target word or the information related to the target word, and the responses were regarded as incorrect and coded as 0 if they did not. For each condition, the total number of correct responses from all 51 participants was averaged, such that a final score of 1 would indicate that all 51 had recalled that word correctly.

A GLMM analysis was run with recall of target item (e.g., responses to the first recall question) as the dependent variable (binomial distribution, logit link). Condition, memory span, and condition trial—the order in which each condition was presented to each participant—were set as fixed factors. A random intercept was specified for both participant (significant: $\beta = .862, SE = .370, p = .020$) and test word (not significant: $\beta = .299, SE = .233, p = .201$). Memory span was not found to be significant ($\beta = .367, SE = .262, p = .168$), and neither was condition trial, $F(2, 288) = .509, p = .602$, suggesting that there had been no primacy or recency effects for the order of presentation of conditions. We found a significant main effect of condition on recall, $F(2, 288) = 5.276, p = .006$. Children were significantly more likely to recall information in the prominence in both speech and gesture condition compared with either of the speech-alone conditions (prominence in both speech and gesture vs. prominence only in speech: $\beta = .241, SE = .077, p = .006$; prominence in both speech and gesture vs. nonprominent speech: $\beta = .191, SE = .081, p = .038$). No difference was found in the proportion of items recalled between the prominence only in speech condition versus the nonprominent speech condition ($\beta = -.050, SE = .082, p = .537$). A graph of the results can be seen in Fig. 3.

![Graph of results](image-url)

**Fig. 3.** Mean recall rates for the target item as a function of the three within-participant conditions (i.e., prominence in both speech and gesture, prominence in speech only, and nonprominent speech). CI, confidence interval. * stands for $p \leq .05$, ** stands for $p \leq .01$.
For the second question, a GLMM was run with recall of information related to target item (i.e., responses to the second recall question) as the dependent variable (binomial distribution, logit link). Condition, memory span, and condition trial were set as fixed factors. A random intercept was specified for both participant (not significant: $\beta = .636, SE = .355, p = .073$) and test word (not significant: $\beta = .526, SE = .374, p = .159$). Memory span was not found to be significant ($\beta = .368, SE = .265, p = .172$), and neither was condition trial, $F(2, 288) = 2.541, p = .081$. However, we found a significant effect of condition on information recall, $F(2, 288) = 5.693, p = .004$. Children were significantly more likely to recall information related to the target item in the prominence in both speech and gesture condition compared with either of the speech-alone conditions (prominence in both speech and gesture vs. prominence only in speech: $\beta = .207, SE = .070, p = .011$; prominence in both speech and gesture vs. nonprominent speech: $\beta = .173, SE = .069, p = .025$). No difference was found in the proportion of items recalled between the prominence only in speech and nonprominent speech conditions ($\beta = -.034, SE = .055, p = .534$). A graph of the results can be seen in Fig. 4.

All in all, the results of Experiment 1 reveal the positive effects of beat gestures on the recall by preschoolers of contrastive information in discourse. Surprisingly, they do not reveal any effects for prosodic prominence alone. Presumably, the null effects of prosody are due to the fact that the target words were included in child-directed discourse, which is typically characterized by systematic use of pitch accentuation in nearly every word (see General Discussion and Conclusions), which neutralized the prosodic salience of the target words relative to the other words in the discourse. Be that as it may, to further explore the enhancing effects of beat gestures on young children’s comprehension of information, we decided to carry out a second experiment.

![Graph](image_url)

**Fig. 4.** Mean recall rates for the information related to the target item as a function of the three within-participant conditions (i.e., prominence in both speech and gesture, prominence in speech only, and nonprominent speech). CI, confidence interval. * stands for $p \leq .05$. 


Experiment 2

The main goal of the second experiment was to assess the effects on comprehension of beat gestures and compare them with the effects of the no-beat condition. Although beat gestures typically accompany prosodic prominence in natural discourse, given the results of the previous experiment, where no effects of prosodic prominence alone were obtained, we decided not to test this condition further.

In contrast to Experiment 1, which had a within-participants design, this experiment had a between-participants design, whereby each child was presented with experimental stimulus materials in only one condition. In other words, all the short story videos that a child watched were either accompanied by beat gestures or produced without beat gestures. Our concern was that if stimulus conditions were mixed, the presence of beat gestures in some narratives would have some kind of carryover effect on narratives where beat gestures were absent.

Method

Participants

A total of 59 5- and 6-year-old preschool and first-grade children from five schools (Escola Casa Nostra, Escola Pública Joan Bruguera, Escola Bora Gran, Escola Can Puig, and Col·legi Dr. Masmitjà) in the province of Girona, Spain, took part in our study. As in Experiment 1, all of the children were bilingual in Catalan and Spanish but were overwhelmingly Catalan dominant (mean dominance of daily use of Catalan relative to Spanish = 90.65%, SD = 8.86). Among the original data set, 4 participants were excluded from the final analysis (2 girls and 2 boys): 1 for not showing good narrative skills in the narrative production task, another 1 for not showing good fluency skills in the narrative production task, and the other 2 for technical and procedural problems that occurred during the experiment. Parents were informed about the experiment's goal, completed a questionnaire (Bosch & Sebastián-Galles, 2001) about their children's language use, and signed a consent form allowing us to run the experiment and process their children's data. All parents reported that their children had exhibited a typical pattern of development and did not have any language disorders. Thus, the results of this study are based on data from the remaining 55 participants (M = 5.84 years, SD = 0.56).

Materials

Preliminary study. Similar to Experiment 1, before preparing recordings of the experimental materials, we carried out a preliminary study to determine the types of beat gestures that naturally occur in stories directed to children and the types of words with which such gestures are typically associated. To do this, two female preschool teachers were video-recorded expressively reading five short stories, resulting in 10 recordings (2 participants × 5 narratives). As in the previous experiment, they were asked to imagine that they were telling each story to 5- and 6-year-old children. The five stories used for this preliminary study were very similar in structure to the narratives later used in the experimental materials. This structure is explained in detail in the “Experimental materials” section below (see also Demir, Fisher, Goldin-Meadow, & Levine, 2014; Demir, Levine, & Goldin-Meadow, 2015). Gestures were analyzed using the MIT Gesture Studies Coding Manual, specifically following the sections on Hand Shape (open or steepled), Palm Orientation (palms up or palms together), and Trajectory Movement (outward or inward). Interestingly, the results of this preliminary study showed that of the 73 beat gestures observed, 49 (67.12%) were associated with focal words. The remaining 24 beat gestures (32.88%) were associated with discourse markers. In terms of the type of gesture, the most common was the open-palm outward hand movement gesture (n = 31, 42.47%) seen in the preliminary study for Experiment 1 (Appendix D, left panel). A second gesture in which the hands were brought together palm inward at waist level (Appendix D, right panel) was also observed. With respect to the distribution of different types of beat gestures over discourse markers and/or content words, 42.47% (n = 31) were open-palm outward hand movements accompanying focal words as well as discourse markers (such as “suddenly” and “in the end”), whereas 6.85% (n = 5) were palm-inward hand movements, which were exclusively associated with focal words, not discourse markers. On the basis of these
observations, we decided to use the outward hand movement as the beat gesture accompanying discourse markers in our stimulus materials and the inward hand movement as the beat gesture accompanying focal words.

**Experimental materials.** A total of six short narratives in Catalan were prepared as the main experimental materials. Each story featured an animal that lived on a farm and followed a similar narrative structure (Demir et al., 2014, 2015). First, the animal was introduced in a particular context such as “Once upon a time, a duck was walking to school” or “Once upon a time, a pig was playing football.” Then, the animal was faced with a problem (initiating event) such as “It started to rain” or “He realized that it was late.” The animal came up with a solution (goal and attempt to achieve the goal) that led to the outcome of the story. This structure lent itself naturally to the formulation of two subsequent comprehension questions: one about the problem and the other about the solution. The English translation of one of these narratives is shown in Example [2] below (English translations of the six target narratives are reproduced in Appendix E). An effort was made to ensure that all discourse markers as well as focal words would be familiar to 5- and 6-year-olds Catalan children (65.7% of all these items were found to be listed in the Catalan version of the MacArthur–Bates Communicative Development Inventories for children aged 16–30 months; Serrat et al., 2010). In Example [2], both discourse markers and focal words appear in boldface.

[2] Example of a stimulus narrative

*Once upon a time, a duck* was walking to school. *Suddenly,* it started to *rain,* and the duck didn’t have an *umbrella.* In the *end,* he came up with a *solution:* He put his hood on his *head* to protect himself from the rain.

Video-recordings were then made of the six narratives being read off a teleprompter by the same two preschool teachers who had participated in the preliminary study, again performing as if they were addressing very young children. Each storyteller recounted two versions of each of the six narratives in two different conditions: with beat gestures and without beat gestures (2 storytellers × 6 narratives × 2 conditions × 2 versions). The recordings were carried out in a lab at the Universitat Pompeu Fabra and were later edited using the AVID video-editing program (Technology, 2016). In the no-beat condition, the storytellers were asked to use child-directed speech but to not produce any beat gestures (see Appendix F, left panel). In the beat condition, they were asked to use child-directed speech accompanied by beat gestures (see Appendix F, right panel) using the outward hand movement to signal discourse markers and inward hand movements to signal focal words (it was observed that both types of hand movement were also typically associated with a slight head nod, a slight widening of the eyes, and eyebrow raising). The two storytellers were trained to produce all target items with the same prosody in both conditions, and the recordings were carefully monitored by two of the authors in this study to make sure that both storytellers maintained consistency in their synchronization between the beat gestures and the corresponding target items. During the recordings, the target items appeared in boldface to remind the speaker to match these items with the appropriate beat gesture on the focused words (see Example [2]). Moreover, we also asked the two storytellers to use child-directed speech with nonexaggerated facial gestures in a consistent way across items in the two conditions, which basically consisted of using a smiling face. After the recording session, the second author of this study compared the two versions of each narrative and selected the best performance, yielding a total of 24 recordings.

To confirm that each storyteller had consistently used the same prosody across no-beat and beat conditions, the audio tracks were acoustically analyzed using Praat (Boersma & Weenink, 2016). In terms of their prosodic content, first the target words were prosodically coded using Cat_ToBI (Prieto et al., 2015). The H and L target F0 points in every pitch accent were manually coded, as were the starting and end points of the target accented syllables and target words. Similarly, a set of three automatic measures were obtained for the whole speech fragment, namely average pitch, average duration, and average intensity. A set of GLMMs were performed to assess the differences among the following prosodic measures across the two conditions: (a) the pitch range, meaning the distance between the minimum and maximum pitches of L + H* pitch accents; (b) the duration of the target
accented syllables; (c) the duration of the target words; (d) the mean pitch of the whole narrative; (e) the mean duration of the whole narrative; and (f) the mean intensity of the whole narrative. In all analyses, condition was set as the fixed factor and a random intercept was set for both speaker and narrative. No statistically significant effects of condition were found for any of the three prosodic variables, namely pitch range, \( F(1, 179) = 0.001, \beta = .374, SE = 13.228, p = .977 \) (results in Hertz) or \( F(1, 179) = 0.092, \beta = -.269, SE = .885, p = .761 \) (results in semitones), accented syllable duration, \( F(1, 202) = 0.865, \beta = -.008, SE = .008, p = .353 \), or target word duration, \( F(1, 126) = 0.006, \beta = .001, SE = .016, p = .938 \). As for general phonetic measures covering each full discourse, no significant effect was found for mean pitch (in Hertz), \( F(1, 20) = 0.001, \beta = .165, SE = .4735, p = .972 \), mean duration of the discourse, \( F(1, 20) = 0.003, \beta = .075, SE = .437, p = .865 \), or mean intensity of the discourse, \( F(1, 20) = 0.028, \beta = .090, SE = .534, p = .868 \). Thus, as expected, all measures yielded nonsignificant differences across the no-beat and beat conditions.

As noted, the recording sessions obtained a total of 24 videos (2 storytellers \( \times \) 6 narratives \( \times \) 2 conditions), each lasting between 18 and 32 s. These recordings were then embedded in PowerPoint presentations to make four different sequences of stimulus materials for each condition, with each set containing all six narratives but in different orders and with the two storytellers also alternating in different orders. The first four slides served to introduce the context of the farm and the animal that would be featured in the narrative. The fifth slide contained the embedded video of the storyteller recounting the story either with or without accompanying beat gestures. The sixth and final slide showed the animal and a question mark, serving as a prompt for the researcher to ask the two comprehension questions.

**Procedure**

The 55 child participants were tested individually in a quiet classroom at their school. Children sat in front of the laptop as they watched each presentation and then stood up in front of the researcher, who asked them two comprehension questions about the narrative. A video camera was situated just behind the researcher to record the children’s behavior both while they watched the presentation and then when they stood up and answered the comprehension questions.

Before the experiment proper commenced, participants were randomly assigned to one of two groups, such that 28 children were shown the stimulus materials in the no-beat condition and 27 children were shown the materials in the beat condition. Thus, each child was shown presentations of all six narratives in one of the two conditions but in varying orders of story and storytellers.

Prior to the experimental task, children were asked to participate in two tasks. First, to control for memory span, the same memory span task administered to children in Experiment 1 was administered to all children in order to determine the maximum number of items they could recall. They were found to have a memory span ranging from 3 to 5 words (M = 3.99, SD = 0.56). Second, to get a measure of their language skills, children were asked to participate in a narrative production task. This consisted of watching two short wordless cartoons (Die Maus, available at [http://www.wdrmaus.de](http://www.wdrmaus.de)) while the experimenter was doing something else and then recounting what had transpired in each cartoon to the experimenter.

These two preliminary tasks were followed by the experimental procedure proper. First, the experimenter showed and talked about the first two presentation slides to engage the children in the narrative context. She explained that there was a lively farm with lots of animals on it and that a sheep that lived on the farm wanted to find out what had happened to each of her six animal friends. Children were then asked to carefully listen to a series of stories about the sheep’s friends so that they could then help the sheep (“Do you want to help the sheep find out what happened to each animal?”). At this point, the third slide depicting the sheep and a question mark appeared. The fourth slide showed the two “farmers” (the storytellers) who would tell the stories. Before hearing each story, children were reminded to listen to the narratives attentively in order to help the sheep. The fifth slide contained the embedded video-recording of the narrative told in either the no-beat or beat condition, depending on the group to which the children had been assigned.

When each recording finished, a sixth slide appeared showing the sheep and a question mark again and children were asked the two questions intended to test their comprehension of the narrative (see Appendix E for English translations of all narratives and their corresponding comprehension ques-
tions). The first question checked children’s understanding of the problem that the animal character had faced (e.g., “Why did the duck get wet?”), and the second question checked their understanding of the strategy used by the character to solve this problem (e.g., “So how did the duck solve the problem?”). To conclude the sequence, the researcher gave the children some positive feedback (e.g., “You have done very well! The sheep is very happy!”). The recording plus comprehension questions sequence was repeated six times for each child, with the whole session lasting 10 to 15 min. Fig. 5 shows a summary of the experimental procedure.

Fig. 5. Experimental procedure for Experiment 2.
Results

Each child answered two comprehension questions about each of the six narratives, yielding a total of 12 responses per child and, therefore, 660 responses from the full group of 55 children. If children answered a comprehension question correctly, the response was coded as 1. The 660 responses were submitted together to a GLMM analysis, again using IBM SPSS Statistics 23. The dependent variable again was whether the participants gave a correct response or not (binomial distribution, logit link). Condition (no beat or beat) and memory span were set as fixed factors. A random intercept was specified for participant (not significant: $\beta = .293$, $SE = .204$, $p = .151$).

The GLMM results showed a significant main effect of condition, $F(1, 657) = 4.211$, $\beta = .572$, $SE = .279$, $p = .041$, $exp(\beta) = 1.772$, indicating that the odds of a participant tested in the beat condition responding correctly to the comprehension questions was about .772 times higher than those for a participant tested in the no-beat condition. Thus, there is evidence that the stories told with beat gestures helped the children comprehend the narratives as compared with stories told without beat gestures. Memory span was not found to be significant ($\beta = .568$, $SE = .304$, $p = .062$). Fig. 6 shows the mean proportion of correct comprehension scores in the two conditions.

General discussion and conclusions

The key findings of the two experiments presented here are that beat gestures help preschool and first-grade children to recall contrastively focused words and information related to them in a discourse, on the one hand, and to comprehend information in a narrative discourse, on the other.

![Fig. 6. Mean proportion of accurate comprehension responses in the two conditions, no-beat and beat. CI, confidence interval. * stands for $p \leq .05$.](image_url)
In Experiment 1, the results from the information recall task showed that children remembered both the target word and information related to the target word better when they were exposed to prominence in both speech and gesture than in the nonprominent and prosodic prominence-only conditions. The positive results of the beat gesture condition expand on previous results on the effects of beat gestures on the recall of information by children. At the same time, our results contrast with prior work that failed to find a facilitative effect of beat gestures in children (So et al., 2012) and adults (Feyereisen, 2006). The explanation for this discrepancy may lie in the fact that these studies presented participants with a sequence of isolated target sentences/words not related to any previous pragmatically relevant discourse context. In the current study, on the other hand, the recall task was presented within a pragmatically relevant task and target information was embedded within child-directed short discourse. In fact, our results are more consistent with Austin and Sweller’s (2014) experiment in which beat gestures seemed to facilitate 3- and 4-year-olds’ recall of how to get from one location to another, a more naturalistic context in which to examine the effect of beats. Thus, the findings of our experiment expand on previous investigations and confirm that preschool children can benefit from the use of beat gestures in pragmatically appropriate situations involving short discourses.

Surprisingly, the results of Experiment 1 showed no differences between the recall rates of the prominent and nonprominent prosody conditions. This contrasts with previous studies that reported beneficial effects of prosodic prominence on information recall by adults in comparison with nonprominent conditions (Fraundorf et al., 2010; Kushch & Prieto, 2016). For example, Fraundorf et al. (2010) showed that accented words associated with prominent prosody (L + H\textsuperscript{+}) were remembered better by adults than words that were less prominent (H\textsuperscript{+}). In our view, the null effects of prosodic prominence in our recall task can be explained by a lack of a saliency effect in the sense that the target items were not perceived to be prosodically different from the rest of the words in the child-directed discourse. It is well known that child-directed speech is characterized by a slower speech rate, higher pitch range, and larger F0 excursions than adult-directed speech (e.g., Zahner, Schönhuber, & Braun, 2016). In our materials, the short narratives presented to children contained a very systematic use of L + H\textsuperscript{+} pitch accentuation. In fact, 90.32% of all the content words occurring in the discourse materials (e.g., nouns, verbs, adjectives, adverbs) received a pitch accent. Thus, we can hypothesize that the target words in the prominent prosody condition (also produced with an L + H\textsuperscript{+} pitch accent) were insufficiently salient relative to the other words in the discourse.

In Experiment 2, our assessment of children’s answers to the comprehension questions revealed that the children better comprehended the stories performed with beat gestures (i.e., in the beat condition). To our knowledge, this constitutes the first behavioral evidence that beat gestures induce higher rates of discourse comprehensibility in children. These results contrast with the earlier study by Macoun and Sweller (2016). However, the reasons for this discrepancy may be related to differences in methodology. For one thing, the age ranges of participants in the two studies differed somewhat. The children in Macoun and Sweller’s study were between 3 and 5 years old, whereas our study (Experiment 2) reports results from older children. This suggests a need to examine children across a wider age range so as to be able to chart how the beneficial effects of gesture may change rapidly over the preschool and early school years. Another potential reason for the difference between our results and Macoun and Sweller’s results lies in the experimental materials. It is important to stress that prior to the carrying out the experiments reported here, we conducted a preliminary study to determine what kinds of beat gestures naturally accompany child-directed narratives and at what points in the narratives they are typically used. We then used the results of that study to design our experimental materials. This preliminary study also guided our placement of beat gestures within the narrative discourse. Whereas previous work has taken into account beat gestures mainly associated with focused positions, the results of our work suggest that it may be important to consider both focused positions and discourse markers (e.g., after, then). Moreover, the ratio of gesture to narrative length may be important as well given that the narratives used in Macoun and Sweller’s study were at least four times longer and contained fewer gestures than the narratives used here. In our view, a more natural use of beat gestures in discourse in terms of both their gestural form and their distribution within the narratives can help strengthen their positive effects on comprehension.
In sum, the results of the two experiments reported in this article show that beat gestures, like representational gestures, can be profitably used by children not only to recall information in pragmatically relevant discourse but also to comprehend narratives. Although beat gestures are nonreferential gestures, they nonetheless serve important linguistic functions such as focus marking and discourse structure marking (Dimitrova et al., 2016; Shattuck-Hufnagel et al., 2016) and, therefore, can crucially help enhance children’s recall and comprehension of discourse information. When children acquire their first language, they benefit from salient elements of speech such as prosodic cues and gestures, which boost their perceptual and attentional systems (de Diego-Balaguer, Martinez-Alvarez, & Pons, 2016). Beat gestures can be understood as pragmatic highlighters of focus positions and discourse markers in narratives, and children can profitably use them to recall and understand language when they are performed in pragmatically relevant ways.

Appendix A

Stories used in the storytelling task preliminary to Experiment 1

<table>
<thead>
<tr>
<th>Story 1</th>
<th>Olga was going for a walk on the beach and decided to go swimming. As she was getting out of the water, she saw red shells and green shells in the sand. She bent down and picked up some green shells</th>
</tr>
</thead>
<tbody>
<tr>
<td>Story 2</td>
<td>Laura’s sister gave her two tickets to go to the village circus. When she was at the circus, Laura saw clowns and animals. But what Laura liked most of all were the animals</td>
</tr>
</tbody>
</table>

Appendix B

Examples of outward open-palm beat gestures obtained in the recordings of the storytelling task preliminary to Experiment 1
Appendix C

Introductory script, discourses, and recall questions for Experiment 1

<table>
<thead>
<tr>
<th>Script for the introduction</th>
<th>Elmer the elephant is very active. He loves doing different activities such as going to the beach, walking through the forest, strolling around the city, or visiting museums. Elmer has a best friend called Núria who always travels with him. On their travels, they pick up lots of different kinds of objects. Sometimes Elmer forgets what he picked up or what he liked most. Now you are going to see some videos where Núria reminds Elmer what they did or what they picked up. Let’s have a look and see if you can help Elmer to remember all these things about their trips</th>
</tr>
</thead>
</table>
| Discourse 1 | NÚRIA: Elmer, do you remember our trip to the fields? In the early morning, we went for a walk in the fields. We walked for many hours! We noticed that there were roses and leaves by the lake, (2) and you gathered some roses, which soon wilted! (3) Later on, you saw that there were fishes and ducks in the lake, and you gathered some **fishes**. They were so small! We had a great time! RESEARCHER: Now help Elmer to remember what he picked up when he went to the field  
**Recall Question 1**: What did he gather in the field?  
**Recall Question 2**: What happened to the roses? What were the fish like? |
| Discourse 2 | NÚRIA: Elmer, do you remember our trip to the circus? One afternoon, we went to the circus and we saw lots of acts. The first act we saw was the one with animals. There were **seals** and zebras. You loved the seals. They were very friendly! Then we went to see the musicians and the magicians. You loved the **magicians**. They were always smiling. It was a very fun day! RESEARCHER: Now help Elmer remember the things he liked most at the circus  
**Recall Question 1**: What were the things that Elmer liked most at the circus?  
**Recall Question 2**: And what were these seals like? And what were these magicians like? |
| Discourse 3 | NÚRIA: Elmer, do you remember the day we went to see the children at your sister’s school? The children from the ladybird’s class had pencils and pens. They gave us **pens**. They were of many different colors. Then we went to see the rabbit’s class, and they were looking at pictures and **books**. They gave us some books. They were about princes and princesses. It was a very fun day! RESEARCHER: Now help Elmer remember what the children at the school gave him as a present  
**Recall Question 1**: What did the children at the school give Elmer as a present?  
**Recall Question 2**: And what were these pens like? And what were these books about? |
Appendix D

Examples of beat gestures obtained in the recordings of the storytelling task preliminary to Experiment 2

Shown are examples of the two beat gestures occurring most often in the storytelling recordings: the outward hand movement (left panel) and the inward hand movement (right panel).

Appendix E

Narratives and comprehension questions for Experiment 2

| First narrative | Once upon a time, a duck was walking to school. Suddenly, it started to rain, and the duck didn’t have an umbrella. In the end, he came up with a solution: He put his hood on his head to protect himself from the rain. | Comprehension Question 1: Why did the duck get wet?  
Comprehension Question 2: So how did the duck solve his problem? |
|-----------------|-------------------------------------------------------------------------------------------------|---------------------------------------------------------------|
| Second narrative | Once upon a time, a rabbit went for walk in the mountains. Suddenly, some cows started to walk toward him and he was scared. In the end, he found a solution: He stood still behind a tree until the cows left. | Comprehension Question 1: Why was the rabbit scared?  
Comprehension Question 2: So how did the rabbit solve his problem? |
| Third narrative  | Once upon a time, there was a horse that was hungry. Suddenly, he realized that there were no biscuits in the cupboard because he had eaten them all. In the end, he thought of a solution: He made biscuits in the oven. | Comprehension Question 1: Why didn’t the horse find any biscuits in the cupboard?  
Comprehension Question 2: So how did the horse solve his problem? |
| Fourth narrative | Once upon a time, there was a hen that was sleepy. Suddenly, she fell asleep on the sofa, but her alarm clock woke her up. She had forgotten that the following day was her birthday and that she had been planning to buy candles to celebrate it. In the end, she found a solution: She bought some enormous candles and was therefore able to celebrate her birthday. | Comprehension Question 1: Why couldn’t the hen celebrate her birthday?  
Comprehension Question 2: So how did the hen solve her problem? |
Fifth narrative

Once upon a time, a pig was playing football in the park. Suddenly, he realized that it was late and he had to go back home because otherwise his mother would get angry. In the end, he thought of a solution: He took a shortcut to get home. That way, he managed to not arrive late and his mother did not get angry.

Comprehension Question 1: Why did the pig have to go home back early?
Comprehension Question 2: So how did the pig solve his problem?

Sixth narrative

Once upon a time, a cat was staying at his grandparents' house in summer. Suddenly, he remembered that he had to do his homework because otherwise his grandparents wouldn't wait for him to go to the beach. In the end, he came up with a solution: He did the homework before his grandparents arrived, and that way he was able to go to the beach.

Comprehension Question 1: Why wouldn't the cat be able to go to the beach with his grandparents?
Comprehension Question 2: So how did the cat solve his problem?

Appendix F

Images from the stimulus video-recordings of Experiment 2

Shown are still images from the stimulus video-recordings of a storyteller reading one of the narratives off a teleprompter in two conditions: no-beat condition (left panel) and beat condition (right panel).

References
