Title:

Beat gestures improve word recall in 3- to 5- year- old children

Authors names and affiliations:

Alfonso Igualada^{1,2}, Núria Esteve-Gibert^{1,3}, Pilar Prieto^{4,1}

¹ *Universitat Pompeu Fabra, Dept. of Translation and Language Sciences, Roc Boronat, 138, 08018 Barcelona, Spain.*

² *Grup de Recerca en Cognició i Llenguatge (GRECIL), Universitat Oberta de Catalunya, Rambla del Poblenou, 156, 08018 Barcelona, Spain.*

³ *Aix Marseille Université, CNRS, LPL UMR 7309, 5 Avenue Pasteur, 13100 Aix-en-Provence, France*

⁴ *ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain*

alfonso.igualada@upf.edu nuria.esteve-gibert@blri.fr pilar.prieto@upf.edu

Corresponding author:

Alfonso Igualada

alfonso.igualada@upf.edu

Roc Boronat, 138. 08018. Barcelona

The word count (text + references): 8853 words.

ABSTRACT

While research has shown that adults can benefit from the presence of beat gestures in word recall tasks, studies have failed to conclusively generalize these findings to preschool children. This study investigates whether the presence of beat gestures helps children to recall information when these gestures have the function of singling out a linguistic element in its discourse context. One hundred and six 3-to-5-year-old children were asked to recall a list of words within a pragmatically child-relevant context (i.e., a story-telling activity) in which the target word was accompanied or not by a beat gesture. Results showed that children recalled the target word significantly better when it was accompanied by a beat gesture than when not, indicating a local recall effect. Moreover, the recall of adjacent non-target words did not differ depending on the condition, revealing that beat gestures seem to have a strictly local highlighting function (i.e., no global recall effect). Interestingly, even though the effect was weak, the analyses show that the children's ability to recall the target word improved with age only in the beat condition. These results demonstrate that preschoolers benefit from the pragmatic contribution offered by beat gestures when they function as multimodal markers of prominence.

Highlights:

Children's word recall improves when the target word is associated with a beat gesture

Children's ability to benefit from beat gestures develops during the preschool years

The presence of beat gestures in a target word does not improve the recall of adjacent words

Keywords:

Gesture-speech integration, beat gestures, word recall, discourse, pragmatics, gesture development

1 INTRODUCTION

In everyday communication, speakers use hand and body gestures to accompany speech. Beat gestures are a type of manual non-representational gesture which co-occurs with speech and which functions as a visual highlighter of information. In contrast with a representation gesture a beat gesture does not add propositional content to a given utterance (McNeill, 1992, 2005; Kendon, 1995) but rather is used to mark “the word or phrase it accompanies as being significant (...) for its discourse pragmatic content” (McNeill, 1992:15). Beat gestures have been defined as rhythmical movements of the hands which are timed together with prosodic prominence in speech (Loehr, 2012; Shattuck-Hufnagel, Ren, Mathew, Yuen & Demuth, 2016). Typically, the movements of hand gestures occur together with head and eyebrow movements, which together signal the privileged status of a given piece of discourse in a multimodal fashion (see, e.g., Cartmill, Demir & Goldin-Meadow, 2012; McNeill, 1992). In our study we investigate whether the use of beat gestures as a multimodal marker of prominence in a significant discourse context helps to improve language recall abilities in early childhood.

Research on gestures has extensively reported the beneficial results of representational gestures for various linguistic abilities, such as the improvement of narrative skills between the ages of 5 and 6 (Demir, Fisher, Goldin-Meadow & Levine, 2014) or the comprehension of complex syntactic abilities by 3- and 4-year-olds (Theakston, Coates & Holler, 2014). In parallel with this, the use of representational gestures has been shown to have cognitive benefits at different stages of children’s cognitive development. For example, the benefits of representational gestures have been proven for 4-5-year-old children in the recall of words in the first language (Church, Kelly & Lynch, 2000; So, Chen-Hui & Wei-Shan, 2012; Thompson, Driscoll & Markson, 1998), for 5-year-old children learning of words in a second language (Tellier, 2008), and for 9-year-olds children solving arithmetic operations (Goldin-Meadow, Cook & Mitchell, 2009). By contrast, the potential beneficial effects of beat gestures, which by definition do not carry semantic meaning, has not been investigated in depth, particularly in development.

Adults seem to benefit from observing beat gestures when asked to recall lexical information in a first language (Kushch & Prieto, 2016; So, et al., 2012) or learn novel words in a second language (Kushch, Igualada & Prieto, under revision). This data suggests that beat gestures highlighting a specific target in the discourse benefit the recall and learning of that target, though less is known about the impact of this highlighting function of beat gestures on the processing of the co-occurring discourse information.

What underlies the possible cognitive advantage offered by beat gestures? Interestingly, several studies measuring event-related brain potentials (ERP) in adults have shown neural evidence of the activation of language-related areas when beat gestures are perceived, suggesting that they have an attentional effect (Biau & Soto-Faraco, 2013; Holle Obermeier, Schmidt-Kassow, Friederici, Ward & Gunter, 2012; Wang & Chu, 2013). The functional neuroimaging study by Biau, Moris Fernandez, Holle, Ávila, and Soto-Faraco (2015) showed that different brain areas were activated depending on whether speech was synchronized with beat gestures or with other non-gestural stimuli (i.e., discs/dots moving on a screen). While beat gestures activated language-related areas of the brain, non-gesture stimuli activated visual-perception areas. Hubbard, Wilson, Callan, and Dapretto (2009) found that beat gestures, and not nonsense movements or still images, enhanced auditory processing of speech. All these data support the idea that beat gestures can be distinguished from other potential visual highlighters because of their direct integration in the language system rather than a more general visual perceptual system.

Moreover, from a linguistic point of view, beat gestures have been shown to serve a focus-marking function (Yasinnik, Renwick & Shattuck-Hufnagel, 2004; Jannedy & Mendoza-Denton, 2005; Loehr, 2012; Shattuck-Hufnagel, et al., 2016). In addition, adult listeners have shown an increase in prominence perception when words are produced together with hand gestures (Krahmer & Swerts, 2007) and head/facial beat gestures (Moubayed, Beskow & Granström, 2010). Apart from the abovementioned physiological and linguistic evidence, the positive cognitive effects of beat gestures are consistent with the embodied cognition paradigm, which underlines the relevance of the

body movements and multimodal supporting channels in cognition and in favoring memory traces (see Paivio, 1990; and Barsalou et al., 2003, Bersalou, 2008).

Though there is strong evidence that beat gestures are related to language and cognitive abilities in adults, the benefits of beat gestures in childhood are less clear. First, research into the first uses of beat gestures during language development is fairly limited. Some production studies point out that children start producing beat gestures around their fifth year of life, as their narrative skills develop (e.g., McNeill, 1992; Stefanini, Bello, Caselli, Iverson & Volterra, 2009; Mathew, Yuen, Shattuck-Hufnagel, Ren & Demuth, 2014; Shattuck-Hufnagel et al., 2016). Regarding perception, to our knowledge only two studies—So et al. (2012) and Austin and Sweller (2014)—have investigated the effects of beat gestures on the recall of information in childhood, with opposite results.

With regard to the first of these studies, So et al. (2012) found that adults benefited from both beat and representational gestures in a word recall task, while children benefited from co-speech representational gestures but not from beat gestures. In a first experiment, adults were shown a video presentation of a list of 10 verbs in three conditions (accompanied by a representational gesture, a beat gesture, or no gesture). In the two conditions containing gestures, each verb co-occurred either with a gesture representing some semantic feature of the action described (the representational condition) or with a simple open-palm downward beat gesture (the beat condition). The results showed that adults recalled a greater number of verbs in either the representational or the beat gesture conditions than in the no gesture condition. The second experiment replicated the first study but with 4- to 5-year-old children as subjects. The procedure was similar except for the number of words presented in the list, which was reduced from 10 to 5 in order to accommodate the shorter mnemonic span of the children. The results revealed that the children ability to recall words benefited from the presence of representational gestures but not from the presence of beat gestures. There are two possible reasons for these negative results. First, in the experiment every word in the list was accompanied by a beat gesture, which may have

reduced the highlighting function of the beat gestures. Typically, beat gestures in natural speech do not appear in with every word but rather are used to make particular items stand out from surrounding non-prominent elements (Terken, 1991; Wagner, Origlia, Avesani, Christodoulides, Cutugno, D'Imperio, Escudero Mancebo, Gili Fivela, Lacheret, Ludusan, Moniz, Ní Chasaide, Niebuhr, Rousier-Vercruyssen, Simon, Šimko, Tesser & Vainio, 2015). Moreover, the list of target words was presented in isolation and without a child-relevant discourse context in which the presence of beat gestures might have been pragmatically motivated. From a perspective of discourse assessment, a task in which linguistic units are presented without contextual support may lack pragmatic motivation (e.g., Ito, Jincho, Minai, Yamane & Mazuka, 2012).¹

With regard to the second study, Austin and Sweller (2014) found that beat gestures did facilitate the recall of spatial directions in 3- to 4-year-old children. Participants first visually examined a toy representation of a landscape. Then, the experimenter provided different directions that the children had to recall afterwards. These directions consisted of locations and actions accompanied by either a co-speech gesture, a beat gesture, or no gesture. In the no gesture condition, spatial directions were described verbally without gestures; in the co-speech gesture condition, different types of gestures (i.e., beat, deictic, metaphoric, and iconic gestures) were produced with affiliated target words; and in the beat condition, target words co-occurred with a beat gesture. Results showed that, on average, children recalled the information about the spatial directions better in either the co-speech gesture condition or the beat condition than in the no-gesture condition. In this experiment, the visual accessibility of the referents through the toy representation might have played a facilitating role in the recall of both location and action target words. Indeed, other studies assessing children's comprehension of prosodic prominence have demonstrated that the presence of visual stimuli facilitates the recognition of target items by children (Ito, et al., 2012; Ito, 2014). However, given that in many natural contexts the visual referents associated with beat gestures are typically not present in the immediate context, it would be of interest to generalize and extend

¹ Moreover, in her chapter about methodologies to explore the processing of prosodic focus by children, Ito (2014) recommends the use of a “conversational task that gives salience to contrastable referential candidates” (p. 206).

these findings to a task which does not involve a concomitant visual presence of the referents. This is what we will endeavor to do in the present study.

The first goal of our study is to investigate preschoolers' general ability to use beat gestures in a word recall task. In order to address the methodological issues that we feel were raised by the two earlier studies, the experimental task was designed with the following features: (a) beat gestures will function as local highlighters of target words which will contrast with adjacent words produced without beat gestures; (b) the task will be embedded in a discourse context that is pragmatically relevant for 3- to 5-year-old children; and (c) the target words will not refer to objects that are visually present in the experimental setting. We expect that children will interpret words associated with salient beat gestures as pragmatically more relevant than the others, which will enhance their recall of them. Thus, we hypothesize that children aged from 3 to 5 will make use of the highlighting function of a beat gesture when having to recall the word affiliated with that gesture.

A second aim of our study is to disentangle whether this potentially enhanced recall effect impacts not only on the target word associated with the beat gesture (i.e., a local recall effect) but also on the recall of adjacent words in the list which are not associated with beat gestures (i.e., a global recall effect). If beat gestures work exclusively as local highlighters of information conveyed through speech, adjacent words in the list should not be affected by the presence of a beat gesture highlighting a target word. This would be consistent with the results of previous studies showing language-related brain responses temporally associated with the words produced with a beat gesture (Hubbard et al. 2009; Biau & Soto-Faraco, 2013; Holle et al., 2012; Wang & Chu, 2013; Biau et al. 2015). If, by contrast, beat gestures induce a global recall effect, a beat gesture co-occurring with a target word will enhance the recall of adjacent words in the list.

Third, while previous studies have reported that children start producing beat gestures at around 5 years of age (e.g., McNeill, 1992; Stefanini et al. 2009; Mathew et al. 2014; Shattuck-Hufnagel, et al. 2016), and that they process other markers of prominence (like pitch accentuation) in an adult-like way around 6 years of age (Ito et al. 2014), we would like to assess the developmental pattern of these effects in younger children aged 3–5. If the effect of beat gestures on memory recall abilities is sensitive to age

differences, this study will provide important data on how multimodal markers of prominence are processed during development.

2 METHODS

2.1 PARTICIPANTS

One hundred and thirteen 3-to 5-year-old children were initially recruited for the study. However, seven of these children had to be excluded from the study: one child was diagnosed with language pathology by the school services, four children did not want to collaborate during the experiment, and two children were not tested because their memory span was greater than the length of the lists prepared for the experiment (see section 2.3.1 below). The total sample of the study thus comprised 106 children aged 3, 4, and 5. Table 1 offers details of the subject sample broken into groups according to age, gender, and memory span. All the participants were preschoolers enrolled at three public nursery schools located in the province of Barcelona, Spain.² In these nursery schools, the main language of instruction is Catalan. Parents were informed about the experiment's goal and signed a participation consent form. Furthermore, language exposure questionnaires (based on Bosch and Sebastián-Gallés, 2001) were administered to the caregivers in order to ensure that children were predominantly exposed to Catalan at home on a daily basis (mean exposure to Catalan as a percentage of all daily language exposure: $M = 87.30$, $SD = 13.09$). Parental questionnaires reported that children were healthy and had normal hearing.

² Escola Sant Martí in Arenys de Munt, Escola La Farigola del Clot in Barcelona, and Escola Pública Dr. Estalella Graells in Vilafranca del Penedès.

Table 1

Children	Number of subjects	Age in months	Girls	Boys	Memory span in number of words
3-year-olds	34	$M = 41.74$ ($SD = 3.58$)	12	22	$M = 2.68$ ($SD = .64$). 2 words ($N = 14$); 3 words ($N = 17$); 4 words ($N = 3$)
4-year-olds	40	$M = 53.93$ ($SD = 3.79$)	18	22	$M = 3.20$ ($SD = .56$). 2 words ($N = 3$); 3 words ($N = 26$); 4 words ($N = 11$)
5-year-olds	32	$M = 64.91$ ($SD = 3.16$)	17	15	$M = 3.84$ ($SD = .45$). 2 words ($N = 1$); 3 words ($N = 3$); 4 words ($N = 28$)

Table 1. The sample population broken down into age groups, showing Mean (M) and Standard Deviation (SD) for age in months and memory span in number of words, as well as gender.

Table 2

Memory span	N	Age in months
2 words	18	$M = 42.78$, $SD = 7.08$, Median = 41.50, Min = 35, Max = 64
3 words	46	$M = 50.15$, $SD = 7.03$, Median = 50.00, Min = 36, Max = 67
4 words	42	$M = 61.33$, $SD = 6.86$, Median = 63.00, Min = 42, Max = 71

Table 2. The sample population separated into groups according to memory span in number of words and showing Mean (*M*), Standard Deviation (*SD*), median, minimum (Min) and maximum (Max) age in months, broken down by memory span.

2.2 MATERIALS

The materials consisted of the audio-visual recordings of sentences produced by a Catalan female actor. Each sentence consisted of an introductory phrase followed by a list of nouns that an elephant named ‘Elmer’ (taken from the children’s book series about *Elmer the Patchwork Elephant* by David McKee) was supposed to remember (see Figure 1 for an example). All nouns in each list belonged to the same semantic context (things to buy at the market, animals to visit at a zoo, objects to tidy up in a room, or pictures to be drawn) and were controlled for frequency of usage. All the nouns included in the lists appeared in the Catalan version of the MacArthur-Bates Communicative Development Inventory (CDI) for children aged 16-30 months (Serrat, Sanz-Torrent, Badia, Aguilar, Olmo, Lara, Andreu & Serra, 2010) (see Appendix 1 for a complete list of sentences). They were all disyllabic with word stress either word-initial or word-final. The audiovisual recording was performed using a PMD660 Marantz professional portable digital video recorder and a Rode NTG2 condenser microphone. Adobe Premiere Pro CS6 editing software was used to splice together the various video sequences (introductory sentence followed by list) and also to embed a drawing of an elephant into the video image. Finally, each video was embedded in a PowerPoint presentation.

Figure 1

Example: “Elmer! Before leaving for your vacation, go to the zoo and say goodbye to the...

BEAT CONDITION: ducks, birds, **PARROTS**, horses, bears

NO-BEAT CONDITION: ducks, birds, parrots, horses, bears

Figure 1. Example of a five-word list in both beat and no-beat conditions in which the central target word is underlined. The word in bold and capital letters was emphasized with a beat gesture in the accompanying video recording.

In a within-subject experimental design, the test materials were presented to the children in two different audiovisual conditions, a beat condition and a no-beat condition. In the beat condition, in order to create a contrast of prominence among the nouns in the list only one word (the target word) was accompanied by a beat gesture. In the no-beat condition, the target noun was not accompanied by a beat gesture. In order to have a target list that was appropriate for the memory span of each particular child, three types of test lists were created, namely three-noun lists, four-noun lists, and five-noun lists, (see section 2.3.1). Table 1 above shows memory span expressed as the number of words a child could recall, broken down by age in years, and table 2 shows the age in months, broken down by memory span ability. For both conditions we controlled for serial sequential effects (i.e., the tendency to remember more easily the first or last items in a list) by placing the target word in the central position in the list (as seen in Figure 1). The critical items were placed in the second position in a three-word length list, in the second position in a four-word length list, and in the third position in a five-word length list.

The beat gesture consisted of a downward hand movement associated with a head nod, opened eyes, and raised eyebrows. More specifically, the hands were raised from a low position near the waist to chin level, with an open-palm vertically oriented hand shape. They were then lowered to chest level as the actor nodded, opened her eyes widely, and raised her eyebrows. Finally, the actor returned her hands to the initial rest position (Figure 2). The specific type of beat gesture used in the materials was defined after conducting a Discourse Completion Task³ (Billmyer & Varghese, 2000) with 10 adult

³ The Discourse Completion Task is a data elicitation method based on discourse contexts which has been applied for many years in research on pragmatics and sociolinguistics (see, e.g., Prieto and Roseano, 2010).

Catalan-speakers who were asked to imagine they had to prompt a 6-year-old child to remember a list of three words by saying something like *Joan! Hem de comprar pomes, iogurt, cebes* ‘John! We need to buy apples, yogurt, onions’ This was done in two ways. First, the adults were given no explicit instructions about how to form the sentence they would tell the child, and second, they were told to emphasize the target words in the list. Results of the Task showed that the adults produced significantly more beat gestures when asked to highlight the target words ($t(238) = -2.696, p < .01$), and that the most commonly occurring beat gesture was an up-down arm movement (100% of instances) with palms open and fingers spread (45% of instances; palm open with fingers curled in 27.8%, okay gesture shape 10%, palm open with fingers touching 10%, and pointing with index finger 2.5%). Typically, this manual gesture was accompanied by a head nod (75.9%) (see Figure 2, right panel). This combination of up-down arm movement, open palm with fingers spread, and head nod was therefore the beat gesture used in all instances by the actor in the video recording.

The recording of each target word was first produced in the beat condition followed by the recording of the no-beat condition in order to avoid variance across conditions. All words in all conditions were produced with the same flat and high pitch contour (H*). Measurements of the recorded signal showed that the pitch range (measured in hertz) of the words did not change across conditions: ($t(38) = .824, p = .415$; beat condition: $M = 53.55, SD = 21.21$; no-beat condition: $M = 48.97, SD = 12.91$). By contrast, the duration of the stressed vowel ($t(38) = 2.092, p < .05$; beat condition: $M = 301.29, SD = 69.54$; no-beat condition: $M = 256.85, SD = 64.75$) and duration of the word ($t(38) = 3.643, p = .01$; beat condition: $M = 1050.85, SD = 237.82$; no-beat condition: $M = 843.75, SD = 89.84$) changed across conditions. Throughout each recording, the gaze of the actor producing the sentence was directed at the bottom left corner of the screen to make it appear as though she was addressing Elmer the elephant, a colorful drawing of which was embedded into this area of the screen.

In the recall experiment, the children were exposed to a total of six trials: two practice trials aimed at familiarizing the child with the task (i.e., one trial per condition) and four experimental trials (i.e., two trials per condition). A total of 36 PowerPoint presentations were created by varying the following factors: (a) the number of words per list (three, four, or five words); (b) three random changes of the word order position in the lists to prevent words from always occurring in the same position; (c) two counterbalanced orders of condition presentation (i.e., one group of children first received an experimental trial in the beat condition followed by an experimental trial in a no-beat condition, this order being repeated across all trials, and the other group of children received all trials in the opposite order of presentation); and (d) a counterbalanced order of presentation of the semantic contexts to prevent a potential semantic priming effect (i.e., some children were exposed first to market and zoo contexts, others first to room and picture drawing contexts). In sum, a total of 36 PowerPoint presentations ($3 \text{ list lengths} \times 3 \text{ word orders in the list} \times 2 \text{ orders of conditions} \times 2 \text{ orders of semantic contexts} = 36 \text{ PowerPoint presentations}$) were produced.

Figure 2

Figure 2. Still images from the stimulus video showing the actor producing the target word in the no-beat (left panel) and beat condition (right panel) while addressing Elmer the elephant.

2.3 PROCEDURE

The children were tested individually in a quiet classroom at their nursery school. The full session lasted approximately 15-20 minutes and consisted of two tasks, a memory span task and a word recall task as detailed below. The memory span task allows us to assign each child to the list type appropriate to their memory span, namely three, four or five word list lengths. Then they were tested with the word recall task.

2.3.1. Memory span task

First, each child was asked to play a game intended to measure their memory span in which they were supposed to repeat a list of words spoken by the experimenter. Following Henry, Messer, Luger-Klein and Crane's (2012) procedure, memory span was measured in terms of the maximum number of words from the list that the child could recall. For all participants, the memory span test started with a list of one item, which was followed by a list of two, a list of three, and so on. This procedure continued until the child could no longer succeed in recalling all the words in the list. Once the maximum list length seemed to have been reached, the child was told four lists of this length but consisting of different words. This number of words was regarded as the child's memory span if all the items were recalled in at least three out of the four lists. The memory span threshold thus measured was what determined the length of the list in the subsequent word recall task, such that if the child's memory span was equal to two words, children were presented with three-word lists in the word recall task. Thus, word span ability was used as a control measure to adjust the demands of the word recall task to participants' memory abilities. Within those parameters, children were randomly assigned to one of the PowerPoint presentations previously prepared (see section 2.2).

2.3.2. Word recall task

The word recall task was performed during the narration of a story about an elephant that enjoys travelling, in which children were told that they would have to recall a list of things that the elephant has to do before travelling. The word recall task consisted of two phases: a presentation phase, where the characters and plot of the story were introduced, and a test phase, which involved a repetitive sequence of trials of three sub-

phases: a word list exposure phase, a word list recall phase, and a story-resumption or concluding scene (see Figure 3 for a schematic diagram and Appendix 1 for a detailed script of the word recall task).

In the presentation phase, the experimenter used the initial slide of a PowerPoint presentation to introduce Elmer the elephant and his friends, a group of elephants and a female human (here a photo of the actor amid drawings of elephants appeared). The experimenter then went on to explain that Elmer always forgot things but was very lucky because his good human friend always helped him to remember things. The child was encouraged to help Elmer too by repeating the list of things that the friend would remind him to do.

After this background to the story was presented to the child, the test phase began. In each trial the plot of the story continued with a set of sequences that all followed the same pattern: (a) a word list exposure phase, in which a video embedded within the PowerPoint showed the actor described a context and said a list of words; (b) a word list recall phase, in which the experimenter asked the child a prompt question and the child attempted to repeat the word list; and (c) a story-resumption or concluding scene with a distractor scene, which was intended to refresh the child's memory load and motivate the resumption of the narrative with a drawing of the character in a new scene (e.g., the beach, mountains, or desert), or the scene that brought the story to a close, with Elmer playing with his elephant friends because he had managed to finish everything he needed to do before leaving on his trip.

Figure 3

Figure 3. Schematic representation of the procedure of the word recall task.

2.4 CODING

An assistant was seated behind the child to code the responses. Children's responses were systematically coded in an answer sheet which included the lists of words in the same order of appearance than in the presentation. Before each target word there was a small check box in which the assistant indicated with a check mark a word that was mentioned by the child. The experimenter double-checked the words with the assistant after each word list recall phase by repeating what the child had said in the same order.

3 RESULTS

3.1 Local word recall effects

In order to assess the effect of beat gestures on the children's ability to recall the target word from among the items in each word list, a Generalized Linear Mixed Model (GLMM) was applied to the data. The dependent variable was the number of target

words recalled in each trial (1–recalled, 0–not recalled) by children during the test phase. Gesture condition (two levels: no-beat gesture and beat gesture), age (three levels: 3-, 4- and 5- year-olds), and all their possible interactions were set as fixed factors; trial, condition order, and participant were set as random factors. Bonferroni pairwise comparisons were carried out for the significant main effects and interactions.

The results of the GLMM analysis showed a main effect of condition ($F(1, 418) = 4.009, p < .05$), with a better recall of the target item in the beat condition than in the no-beat condition. The mean (M) and standard deviation (SD) values of the recall of the target item were as follows: no-beat condition: $M = .38, SD = .48$; beat condition: $M = .49, SD = .5$. The results did not show a main effect of age ($F(2, 418) = 2.804, p = .062$). No significant interaction between gesture condition and age was found ($F(2, 418) = .112, p = .894$). Figure 4 shows that at all ages, target items accompanied by a beat gesture were recalled better than target items not accompanied by a beat gesture.⁴

⁴ Following a reviewer's suggestion to test for a potential gender difference in the ability to recall the target word, a GLMM analysis was applied to the dependent measure of the ability to recall the target word. Gender (two levels: male and female), condition (two levels: no-beat and beat gesture), age (three levels: 3-, 4- and 5- year-olds), and all their possible interactions were set as fixed factors; trial, condition order and participant were set as random factors. Bonferroni pairwise comparisons were carried out for the significant main effects and interactions. The results of the GLMM did not show a main significant effect of gender ($F(1, 412) = 0.273, p = .602$), nor an interaction effect between gender and condition ($F(1, 412) = 1.587, p = .208$), nor a triple interaction effect between gender, condition and age ($F(2, 412) = 0.424, p = .654$). These results back up our decision not to include gender as a fixed factor in our main GLMM model.

Figure 4. Mean number of items recalled in the target position as a function of condition and age.

3.2 Global word recall effects

In order to assess what we call the global effects of beat gestures, that is, the potential effect of the presence of beat gestures on the recall of the non-target items in the list, another GLMM analysis was run. The dependent variable was calculated as a proportion of the number of non-target words recalled in each trial divided by the maximum number of items in the list. Gesture condition (two levels: no-beat gesture

and beat gesture), age (three levels: 3-, 4- and 5-year-olds), and all their possible interactions were set as fixed factors; trial, condition order and participant were set as random factors. Bonferroni pairwise comparisons were carried out for the significant main effects and interactions.

The results of the analysis showed that recall of non-target items was not significantly affected by the gesture condition ($F(1, 418) = .165, p = .695$). A main effect of age was found ($F(2, 418) = 21.697, p < .001$), and Bonferroni pairwise comparisons revealed that 5-year-old children recalled the items in non-target positions better than 3-year-olds ($p < .001$) and 4-year-olds ($p < .01$), and 4-year-old children recalled non-target items better than 3-year-old children ($p < .01$). No significant interaction was found between condition and age ($F(2, 418) = 1.505, p = .223$). The mean (M) and standard deviation (SD) values for the recall of non-target items were as follows: for 3-year-olds, no-beat condition: $M = .35, SD = .23$; beat condition: $M = .30, SD = .21$; for 4-year-olds, no-beat condition: $M = .42, SD = .19$; beat condition: $M = .40, SD = .18$; and for 5-year-olds, no-beat condition: $M = .48, SD = .16$; beat condition: $M = .52, SD = .16$.

3.3 Correlation analysis between word recall ability and age in months

In the previous set of results in section 3.1, no main effect of age nor a significant interaction between age and condition was found on the ability to recall the target word. Nevertheless, visual inspection of Figure 4 and the fact that the effect of age was near-significant ($F(2, 418) = 2.804, p = .062$) signals the importance of age in the ability to

recall words and suggests that this developmental pattern is stronger in the beat condition than in the no-beat condition.⁵

In order to explore these effects, two separate pairwise correlation analyses were run with children's age in months in relation to the ability to recall the target item in first the beat condition and then in the no-beat condition (that is, a separate correlation analysis for each gesture condition). The dependent variable was the average recall score for each child per condition. This measure was calculated by adding the total number of words that a child recalled in the two trials associated with one condition and then dividing that figure by two (the number of trials), resulting in possible averages of 0, .5, or 1.

The results of the first correlation analysis showed that the ability to recall the target word in the beat condition was positively and significantly correlated with age in months $r(104) = .234$ ($p < .05$). Conversely, children's age in months was not significantly correlated with recall of the target word in the no-beat condition $r(104) = .155$ ($p = .112$). These results confirm that the ability to recall words in the beat condition develops more strongly with age than the ability to recall words in the no-beat condition.

4 DISCUSSION AND CONCLUSIONS

⁵ The mean and standard deviation values for recall of the target item were as follows: for 3-year-olds, no-beat condition: $M = .31$, $SD = .46$; beat condition: $M = .40$, $SD = .49$; for 4-year-olds, no-beat condition: $M = .4$, $SD = .49$; beat condition: $M = .49$, $SD = .5$; and for 5-year-olds, no-beat condition: $M = .44$, $SD = .5$; beat condition: $M = .58$, $SD = .49$.

In this study we set out to investigate whether the presence of beat gestures increases children's word recall in a list of words when beat gestures function as multimodal highlighters in a child-relevant discourse context. Second, we investigated whether the impact of beat gestures on recall is limited to words that co-occur with a beat gesture (i.e., it has a merely local effect), or whether this effect extends also to adjacent words in the discourse (a global effect). Moreover, we were interested in investigating whether children's age influences their ability to benefit from perceiving beat.

The results of a word recall task performed by one hundred and six 3-to-5-year-old children showed that the children recalled the target words significantly better when they were accompanied by a beat gesture than when they were not. This demonstrates that preschool children benefit from the presence of beat gestures when the gesture marks an item as being more salient than others. This evidence represents a valuable addition to the hitherto contradictory results reported in the literature on the benefits of beat gestures in children. On the one hand, So et al. (2012) reported results showing that preschoolers recalled more words when they were associated with iconic gestures than when they were not, but this facilitating effect was not found with beat gestures. The clear effects of the beat condition in our study, in comparison with the negative effects found in So et al. (2012), might have been influenced by a set of factors: (a) in our materials, beat gestures functioned as local highlighters of a target word which contrasted with adjacent words produced without beat gestures; (b) the task was embedded in a discourse context that was pragmatically relevant for preschoolers; and (c) we followed a naturalistic approach in the design of the materials and the beat gestures included hand movements together with other gestural markers of prominence, such as head and eyebrow movements. It is well known that in natural communication,

beat gestures typically co-occur with other prosodic and gestural markers of prominence such as hand and head movements, and specific facial cues. In So et al.'s experiment, beat gestures were limited to hand movements only. In our study, on the other hand, beat gestures were multimodal. This multimodality may have boosted the children's perception of prominent elements within the discourse, thus explaining the difference in our results.

With respect to this multimodal encoding of beat gestures, future research will be needed to address the question of whether certain visual markers of prominence (i.e., particular head, eyebrow, or hand movements) are more powerful than others, or act like prosodic markers of prominence. As is well known, there is a tight temporal synchrony between beat gestures and prosodic prominence (Leonard & Cummins, 2010; Loehr, 2012; Shattuck-Hufnagel, et al., 2016). However, it is also possible that other non-manual multimodal markers of prominence like head movements worked together with the hand beat gestures to trigger the word recall effect. Recent results seem to suggest that adults learn novel words better when they are synchronized with a beat gesture, but only if the words are accompanied by prosodic prominence (Kushch, Igualada & Prieto, 2015). Moreover, while prosodic prominence triggers attentional processes, only visual prominence (e.g., beat gestures) facilitates semantic processing of the co-occurring words during the perception (Wang & Chu, 2013).

In contrast with So et al.'s results, and in line with Austin and Sweller (2014), we found that preschoolers recalled significantly more target words when they were accompanied

by a beat gesture. As noted above, in Austin and Sweller (2014) the visual presence of referents might have played a facilitating role in the recall task. In this regard, our study expands on their results by showing that beat gestures strongly favor word recall even when the referent is not visually present in the conversational context. Presumably the ability to access the highlighting function of beat gestures is developing in children of this age range, and therefore accompanying every word with a beat gesture (as in So et al.'s experiment) makes it more difficult for children to access the highlighting and contrasting function of such gestures. Importantly, in spontaneous discourse some words are accompanied by beats and others are not (Terken, 1991; Austin & Sweller, 2014; Wagner, et al., 2015), and this saliency feature of beat gestures might be important for the child to identify what is the important information in a discourse.

Our results also point to a local effect of beat gestures, since we found no overall increase in the children's recall of adjacent words in the beat condition. These results are in line with recent neurophysiological data showing a clear local time alignment between brain responses and target speech associated with beat gestures (Hubbard et al., 2009; Wang and Chu, 2013; Biau, et al., 2015). Beat gestures seem to act as attentional local highlighters that help children focus their attention on a particular piece of information, which in turn helps them improve their performance in a recall task.

With respect to the developmental factor, our results indicated that 3- to 5-year-old children performed similarly in the recall task regardless of age when age was expressed in years. A more fine-grained analysis showed a weak but significant positive correlation between age expressed in months and the ability to recall a target word in the beat condition. The correlation between age in months and the ability to recall target

words in the no-beat condition, however, was not significant. Interestingly, the ability to recall words associated with beat gestures seems to appear earlier than children's own first productions of beat gestures, which takes place around the age of 5 (McNeill, 1992; Stefanini et al., 2009; Mathew et al., 2014). Future research could further explore how beat gestures develop in parallel with other cognitive and linguistic abilities, and whether the sensitivity to beat gestures at age 3 can predict a greater use of these gestures (and other grammatical markers of saliency) at later ages. Oral strategies to express saliency in discourse (by means of, for example, degree modifiers, syntactic word order, or contrastive pitch accent) seem to develop at around 4-5 years of age (Chen, 2011; Ito et al., 2012; Järvikivi, Pyykkönen-Klauck, Schimke, Colonna & Hemforth, 2015; Tribushinina, 2004), and our results show a successful interpretation of beat gestures as early linguistic highlighting devices at even younger ages, pointing to a potential scaffolding role for beat gestures as multimodal markers of prominence in a discourse.

In sum, the main novelty of our study is the evidence it provides that children can benefit from the presence of beat gestures functioning as multimodal markers of prominence in a pragmatically appropriate context. Why is it that beat gestures help listeners recall the target words that they accompany? We think the answer lies in the fact that gestures mark referents as being prominent in a multimodal way, and this saliency increases the attention that a listener pays to a particular piece of information. In fact, the ability to selectively attend to specific elements of speech while disregarding others (i.e., temporal attention) has recently been argued to facilitate language development in its early stages (de Diego-Balaguer, Martinez-Alvarez & Pons, 2016). In the realm of applied linguistics, beat gestures could be used as a teaching strategy to

cue relevant information in a discourse context in both educational and therapeutic settings.

ACKNOWLEDGMENTS

We would like to thank the staff at the Escola Sant Martí, Escola La Farigola del Clot, and Escola Pública Dr. Estalella Graells for granting us access to and organizing the meetings with preschoolers. We also would like to thank the children and their families for their participation in the experiment. Furthermore, many thanks to Iris Hübscher for participating in the collection of the data, to Santiago González for his help in designing the study, and to Joan Borràs-Comes for assistance with the statistical analyses. Finally, we are grateful to Laura Bosch, Thierry Nazzi, Marianne Gullberg, Sotaro Kita, Susan Goldin-Meadow, Mili Mathew, Catherine So Wing-chee, and Elisabeth Austin for thoughtful ideas and comments at different international meetings. This research has been funded by grants awarded by Recercaixa 2013-2015 and the Spanish Ministry of Science and Innovation (FFI2015-66533 BFU2012-31995 “Intonational and gestural meaning in language”), and also by a grant awarded by the Generalitat de Catalunya (2014SGR-925) to the Prosodic Studies Group.

REFERENCES

- Austin, E. E. & Sweller, N. (2014). Presentation and production: The role of gesture in spatial communication. *Journal of Experimental Child Psychology*, 122, 92-103.
- Biau, E. & Soto-Faraco, S. (2013). Beat gestures modulate auditory integration in speech perception. *Brain & Language*, 124, 143-152.
- Biau, E., Moris Fernandez, L., Holle, H., Avila, C., & Soto-Faraco, S. (2015). Spontaneous beat gestures as prosody: an asynchrony with speech affects language processing. In Biau, E. *Beat gestures and speech processing: When prosody extends to the speakers hands* (Doctoral thesis), 89-119.
- Billmyer, K., & Varghese, M. (2000). Investigating instrument-based pragmatic variability: effects of enhancing discourse completion tests. *Applied Linguistics*, 21(4), 517–552.
- Barsalou, L. W., Simmons, W.K., Barbey, A. K., and Wilson, C. D. (2003) Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, 7(2), pp. 84-91. doi:10.1016/S1364-6613(02)00029-3
- Barsalou, L. W. (2008) Grounded cognition. *Annual Review of Psychology*, 59(1), pp. 617-645.
- Bosch, L., & Sebastián-Galles, N. (2001). Evidence of Early Language Discrimination Abilities in Infants from Bilingual Environments. *Infancy*, 2, 29–49.
- Cartmill, E. A., Demir, O. E., & Goldin-Meadow, S. (2012). Studying gesture. In Hoff, E. (Ed.), *Research methods in child language: a practical guide*. Blackwell Publishing.
- Chen, A. (2011). The developmental path to phonological focus-marking in Dutch. In S. Frota, E. Gorka, & P. Prieto (Eds.), *Prosodic categories: Production, perception and comprehension* (pp. 93-109). Dordrecht: Springer.

- Church, R., Kelly, S., & Lynch, K. (2000). Immediate memory for mismatched speech and representational gesture across development. *Journal of Nonverbal Behavior*, 24, 151–174.
- de Diego-Balaguer, R., Martinez-Alvarez, A., & Pons, F. (2016). Temporal attention as a scaffold for language development. *Frontiers in Psychology*, 7(44), 1- 15.
- Demir, Ö. E., Fisher, J. A., Goldin-Meadow, S., & Levine, S. C. (2014). Narrative processing in typically developing children and children with early unilateral brain injury: Seeing gesture matters. *Developmental Psychology*, 50(3), 815–28.
- Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. A. (2009). Gesturing gives children new ideas about math. *Psychological Science*, 20 (3), 267–272.
- Henry, L.A., Messer, D., Luger-Klein, S., & Crane, L. (2012). Phonological, visual, and semantic coding strategies and children’s short-term picture memory span. *The Quarterly Journal of Experimental Psychology*, 65(10), 2033-2053.
- Holle, H., Obermeier, C., Schmidt-Kassow, M., Friederici, A.D., Ward, J., & Gunter, T. C. (2012). Gesture facilitates the syntactic analysis of speech. *Frontiers in Psychology*, 3(74), 1-12.
- Hubbard, A. L., Wilson, S. M., Callan, D. E., & Dapretto, M. (2009). Giving speech a hand: gesture modulates activity in auditory cortex during speech perception. *Human Brain Mapping*, 30(3), 1028–37.
- Ito, K. (2014). Children’s pragmatic use of prosodic prominence. In McNeill, D. (ed.). *Pragmatic development in first language acquisition*. John Benjamins Publishing Company, 199-218.
- Ito, K., Jincho, N., Minai, U., Yamane, N., & Mazuka, R. (2012). Intonation facilitates contrast resolution: Evidence from Japanese adults & 6-year olds. *Journal of Memory and Language*, 66 (1), 265-284.
- Jannedy, S., & Mendoza-Denton, N. (2005). Structuring information through gesture and intonation. *Interdisciplinary Studies on Information Structure*, 3, 199–244.

- Järvikivi, J., Pyykkönen-Klauck, P., Schimke, S., Colonna, S., & Hemforth, B. (2015). Information structure cues for 4-year-old and adults: Tracking eye movements to visually presented anaphoric referents. *Language, Cognition and Neuroscience*, 29(7), 877-892.
- Kendon, A. (1995). Gestures as illocutionary and discourse structure markers in Southern Italian conversation. *Journal of Pragmatics*, 23, 247-279.
- Krahmer, E., & Swerts, M. (2007). Effects of visual beat gestures on prosodic prominence: Acoustic analyses, auditory perception and visual perception. *Journal of Memory and Language*, 57, 396-414.
- Kushch, O., Igualada, A. & Prieto, P. (under revision). Do beat gestures and prosodic prominences help when acquiring novel words in a second language? *Presentation at Eurosla 2015 conference*. Aix-en-Provence, France, August 26 – 29.
- Kushch, O., & Prieto, P. (2016). The effects of pitch accentuation and beat gestures on information recall in contrastive discourse. *Proceedings of Speech Prosody*. Boston, USA, May 31 – June 3.
- Leonard, T., & Cummins, F. (2010). The temporal relation between beat gestures and speech. *Language and Cognitive Processes*, 26, 1295-1309.
- Loehr, D. (2012). Temporal, structural, and pragmatic synchrony between intonation and gesture. *Laboratory Phonology*, 3(1), 71-89.
- Mathew, M. M., Yuen, I., Shattuck-Hufnagel, S., Ren, A., & Demuth, K. (2014). The use of prosodic gesture during parent-child discourse interactions. *Australian Linguistic Society Annual Conference*. The University of Newcastle, Australia.
- McNeill, D. (1992). *Hand and Mind*. Chicago: University of Chicago Press.
- McNeill, D. (2005). *Gesture and thought*. Chicago: University of Chicago Press.
- Moubayed, S.A., Beskow, J. & Granström, B. (2010). Auditory visual prominence. *Journal of Multimodal User Interfaces*, 3, 299-309.

- Paivio, A. (1990). *Mental representations: A dual coding approach*. New York, NY: Oxford University Press.
- Prieto, P., & Roseano, P. (2010). *Transcription of Intonation of the Spanish Language*. München: Lincom.
- Serrat, E.; Sanz-Torrent, M.; Badia, I., Aguilar, E., Olmo, R., Lara, M.F., Andreu, L. y Serra, M. (2010). La relación entre el aprendizaje léxico y el desarrollo gramatical. *Infancia y Aprendizaje*, 33(4), 435-448.
- Shattuck-Hufnagel, S., Ren, A., Mathew, M., Yuen, I., & Demuth, K. (2016). Non-referential gestures in adult and child speech: Are they prosodic? *Proceedings of Speech Prosody*. Boston, USA, May 31 – June 3.
- So, W. C., Chen-Hui, C. S., & Wei-Shan J. L. (2012). Mnemonic effect of iconic gesture and beat gesture in adults and children: Is meaning in gesture important for memory recall? *Language and Cognitive Processes*, 27(5), 665-681.
- Stefanini, S., Bello, A., Caselli, M. C., Iverson, J. M., & Volterra, V. (2009). Co-speech gestures in a naming task: Developmental data. *Language and Cognitive Processes*, 24(2), 168–189.
- Tellier, M. (2008). The effect of gestures on second language memorization by young children. *Gesture*, 8(2), 219-235.
- Terken, J. (1991). Fundamental frequency and perceived prominence. *Journal of the Acoustical Society of America* 89(4), 1768–1776.
- Theakston, A. L., Coates, A., & Holler, J. (2014). Handling agents and patients: Representational co-speech gestures help children comprehend complex syntactic constructions. *Developmental Psychology*, 50(7), 1973-1984.
- Thompson, L., Driscoll, D., & Markson, L. (1998). Memory for visual–spoken language in children and adults. *Journal of Nonverbal Behavior*, 22, 167–187.
- Tribushinina, E., (2014). Comprehension of degree modifiers by pre-school children: What does it mean to be ‘a bit cold’? *Folia Linguistica*, 48(1), 255-276.

- Wagner, P., Malisz, Z., & Kopp, S. (2014). Gesture and speech in interaction: An overview. *Speech Communication*, 57, 209–232.
- Wagner, P., Origlia, A., Avesani, C., Christodoulides, G., Cutugno, F., D'Imperio, M., Escudero Mancebo, D., Gili Fivela, B., Lacheret, A., Ludusan, B., Moniz, H., Ní Chasaide, A., Niebuhr, O., Rousier-Vercruyssen, L., Simon, A.C., Šimko, J., Tesser, F. & Vainio, M. (2015). Different parts of the same elephant: a roadmap to disentangle and connect different perspectives on prosodic prominence. *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, Scotland, August 10-14.
- Wang, L., & Chu, M. (2013). The role of beat gesture in pitch accent in semantic processing: An ERP study. *Neuropsychologia*, 51, 2847-2855.
- Yasinnik Y, Renwick M, Shattuck-Hufnagel S (2004). The timing of speech-accompanying gestures with respect to prosody. *Proceedings of Sound to Sense, MIT*. June 11-13.

Appendix 1

Scripts for narration of the two parts in the word recall task.

1 Presentation of the characters and plot of the story

1.1 Presentation of the main character and his friends by the experimenter.

- 1.1.1 Slide with a drawing of the main character hidden under a colored rectangle to draw children's attention to the story.

(Catalan) *Saps qui és aquest?* (English) *Do you know who this is?*

- 1.1.2 Slide appears with a picture of the main character, Elmer the elephant.

(Catalan) *És un elefant que es diu Elmer! L'Elmer és un elefant de colors i li agrada molt viatjar. Viatjar és el que més li agrada!* (English) *This is an elephant called Elmer! Elmer is a colorful elephant and he enjoys traveling very much. Traveling is what he likes most!*

- 1.1.3 Slide appears with a picture of other elephants and among them a photo of the actor, who is about to appear in the video.

Presentation of the main character's friends.

(Catalan) *Mira! Aquests són els amics de l'Elmer. Els seus amics són... (pausa) elefants! I aquesta noia? Qui és? (pausa). També és la seva amiga, i es diu Núria.* (English) *Look! These are Elmer's friends. His friends are... (pause) elephants! And this girl? Who is she? (pause). She's his friend too, and her name is Núria.*

1.2 Presentation of the plot of the story.

Same as previous slide. (Catalan) *L'Elmer i els seus amics volen marxar de viatge, però saps què passa? (pausa). L'Elmer sempre s'oblida de les coses. Té molta mala memòria, i sempre s'oblida de les coses que ha de fer. Sort que hi ha la Núria, que és molt maca i l'ajuda a recordar. Escolta. Que l'ajudem nosaltres també? (pausa). Va, anem a ajudar-lo a recordar el que ha de fer l'Elmer abans de marxar de viatge. Para molta atenció al què diu.* (English) *Elmer and his friends want to go on a trip but do you know what happens?*

(pause). *Elmer always forgets things. He has a very bad memory and he always forgets the things that he has to do. Luckily Núria is really nice and helps him to remember. Listen. Shall we help him too?* (pause). *Let's help Elmer to remember what he needs to do before going on the trip. Pay close attention to what Núria is saying.*

2 Repeating sequence of the plot consisting of the following three phases:

- 2.1 **Word list exposure phase.** Slide with a video recording of the actor. In the first part of the introductory sentence, the actor gazes straight at the camera and says “Hi!” to engage the child’s attention. Then, when a drawing of Elmer the elephant appears in the bottom left part of the screen, she shifts her gaze so that she appears to be looking at Elmer. She addresses Elmer by name and keeps her gaze directed at him for the remainder of this phase as she presents a particular context and reminds Elmer of the list of things he must do here.

Market context. (Catalan) *Hola! Elmer! Abans de marxar, has de anar al mercat i comprar pomes, iogurt, cebes, aigua, raïm.* (English) *Hi, Elmer! Before leaving, you have to go to the market and buy apples, yogurt, onions, water, grapes.*

Zoo context. (Catalan) *Hola! Elmer! Abans de marxar, has de anar al zoo has d'acomiar-te dels óssos, lloros, ànecs, cavalls, pardals.* (English) *Hi, Elmer! Before leaving, you have to go to the zoo and say goodbye to the bears, parrots, ducks, horses, birds.*

Room context. (Catalan) *Hola! Elmer! Abans de marxar, a l'habitació has d'endreçar llibres, nines, cotxes, papers, globus.* (English) *Hi, Elmer! Before leaving, in your room you have to tidy up your books, dolls, cars, papers, balloons.*

Drawing context. (Catalan) *Hola! Elmer! Abans de marxar, has de fer un dibuix de l'escola i que tingui portes, arbres, classes, taules, coixins.* (English)

Hi, Elmer! Before leaving, you have to make a drawing of the school with doors, trees, classrooms, desks, pillows.

2.2 Word list recall phase prompted by the experimenter.

Experimenter: (Catalan) *Què ha de fer l'Elmer?* (English) *What does Elmer have to do?*

2.3 Story resumption or concluding scene.

2.3.1 Distractor phase and motivation to link to next sequence. Slide with a picture of the elephant in different scenes.

2.3.1.1 Interaction between child and experimenter. (Catalan) *Mira! Ha anat a... Allà farà...* (English) *Look! He has gone to... There he will...*

2.3.1.2 Link to the next sequence of the plot

(Catalan) *Però la Núria li diu: "Elmer, encara has de fer més coses". L'Elmer s'ha oblidat de altres coses que ha de fer. Anem a ajudar-lo a recordar un altre cop. A veure què diu ara, la Núria. Para molta atenció.* (English) *But Núria is saying: "Elmer, you still have things to do". Elmer has forgotten about other things that he has to do. Let's help him again to remember. Let's see what Núria is saying now. Pay close attention.*

2.3.2 End of the story. Slide with a picture of the elephant playing with his elephant friends.

(Catalan) *Finalment, aquesta és l'última cosa que l'elefant havia de fer! Mira que content que està després de treballar tant. L'Elmer ara ja pot anar de viatge amb els seus amics.* (English) *Finally, that was the last thing that the elephant had to do! Look how happy he is after all this hard work. Elmer can now go on a trip with his friends.*