

# Talking to the beat: Six-year-olds' use of stroke-defined non-referential gestures

First Language

1–18

© The Author(s) 2017

Reprints and permissions:

[sagepub.co.uk/journalsPermissions.nav](http://sagepub.co.uk/journalsPermissions.nav)

DOI: 10.1177/0142723717734949

[journals.sagepub.com/home/fla](http://journals.sagepub.com/home/fla)



**Mili Mathew**

Macquarie University, Australia; St. Cloud State University, USA

**Ivan Yuen and Katherine Demuth**

Macquarie University, Australia

## Abstract

Children are known to use different types of referential gestures (e.g., deictic, iconic) from a very young age. In contrast, their use of non-referential gestures is not well established. This study investigated the use of *stroke-defined non-referential* 'beat' gestures in a story-retelling and an exposition task by twelve 6-year-olds, an age at which proficiency in discourse is beginning to develop. The goals of the study were to (1) establish if children this age use *stroke-defined beats*, (2) determine whether the two discourse types influence the incidence of stroke-defined beats, and (3) examine the extent to which stroke-defined beats co-occur with lexical words or pitch accents. The results showed that nine of the children produced at least one beat gesture with a well-defined stroke phase, and that the frequency of the stroke-defined beat gesture use did not differ significantly between the two tasks. Stroke-defined beats occurred more often on lexical words than function words, but they did not co-occur more often with a pitch accent, suggesting its *potential* link with pitch accents for emphasis. This study therefore provides support for children's 'prosodic' use of gesture – a function which may become more common as discourse abilities develop.

## Keywords

Discourse, emphasis, gestural development, gesture–speech relationship, stroke-defined beats

---

## Corresponding author:

Mili Mathew, Department of Communication Sciences and Disorders, St. Cloud State University, 720 4th Ave. S., St. Cloud, MN 56301, USA.

Email: [mmathew@stcloudstate.edu](mailto:mmathew@stcloudstate.edu)

## Introduction

Gestures are spontaneously-occurring visible actions of any part of the body, often used to accompany an utterance or to supplant part of an utterance (Kendon, 2004). Although gestures could have influenced the genesis of conventional sign languages (Marshall & Morgan, 2015), these co-speech hand gestures are not defined by convention (McNeill, 1992). Rather they are often classified in terms of hand configurations or movement characteristics into iconic, deictic, metaphoric, and beat gestures (Kendon, 2004; McNeill, 1985). Iconic, deictic, and metaphoric gestures are referred to as '*referential*', because they resemble aspects of the semantic or lexical content. Beats are considered abstract and '*non-referential*' because they do not convey semantic meaning and are often identified based on their movements. They are commonly assumed to serve an *emphatic/highlighting* function during discourse (Igalada, Esteve-Gibert, & Prieto, 2017; McNeill, 1992). Yet the movement characteristics of beats have been variably delineated, and they therefore tend to be classified on the basis of their discourse function. This mixed characterization of beats across studies poses a challenge to any investigation of the link between beats and their discourse role.

According to Kendon (1996), gestures are characterized by features such as excursion, peak, well-boundedness, and symmetry. Excursion describes an action moving from a 'rest' position and returning to the 'rest' position. Peak (also identified as stroke) refers to the 'centre' of the excursion (i.e., the purpose of the excursion). An action is 'well-bounded' with a clear onset and offset. An action is considered 'symmetrical' when its backward movement is a mirror image of its forward movement in time. Of these four features, 'peak/stroke' often constitutes the crux of a gesture, as it reflects maximum information density. Gestures are generally used in three communicative ways: (1) gestures as stand-alone utterances; (2) gestures replacing speech, but as components of the same utterance; and (3) gestures in conjunction with speech.

Previous research on adults has investigated deictic and beat gestures in conjunction with speech (e.g., Krahmer & Swerts, 2007; Rusiewicz, Shaiman, Iverson, & Szuminsky, 2013, 2014). Beat gestures have been found to modify the acoustic realization of speech in the same way as pitch accents (Krahmer & Swerts, 2007). For instance, duration of the vowel /a/ in Dutch was longer either with a pitch accent or a beat gesture than without. This indicates that beat gestures and pitch accents can serve a similar '*highlighting/emphasizing*' function. Similarly, listeners are able to use visual beat gestures to evaluate the strength of acoustic prominence. The gesture–speech link has also been reported for deictic gestures (Rusiewicz et al., 2013, 2014). For example, Rusiewicz and colleagues found that contrastive stress and speech with delayed auditory feedback increased total gesture duration. These findings suggest that gestures and speech are often temporally aligned and can influence one another. This synchrony with pitch-accented syllables is reflected in both elicited speech (Leonard & Cummins, 2011) and spontaneous speech (Loehr, 2012). It should be noted that the presence vs. absence of a pitch accent was manipulated for nouns in the above-mentioned experimental studies, since nouns are more likely to attract pitch accent.

As for children, gestures generally appear in early infancy, but tend *not* to be beats (Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979). Early gestures are often

referential in nature, directed towards an object or referring to a 'concrete' entity. From the age of 10 months, infants begin to communicate by using *deictic* gestures, such as pointing (Bates et al., 1979), and this often takes place before they are able to use meaningful words. At around 12 months, *iconic* gestures emerge (e.g., child flapping arms to represent a bird), coinciding with the appearance of first words (Bates et al., 1979). By 18 months, children produce gestures to complement or supplement the lexical content of their speech (Bates et al., 1979; Butcher & Goldin-Meadow, 2000; Iverson, Caprici, & Caselli, 1994).

However, only a few studies have examined the link between beats and language, and when this link develops in children's speech. Although a few studies have reported the use of beats in bilingual children as young as 2.5/3 years (Mayberry & Nicoladis, 2000; Nicoladis, Mayberry, & Genesee, 1999), beats in discourse tend to be reported for older children aged 6 and 10 (Blake, Myszczyzyn, Jokel, & Bebiroglu, 2008; Colletta, Pellenq, & Guidetti, 2010). Thus, compared to referential gestures, beats tend to appear late. Perhaps the divergence in their occurrence might be related to the different functions of these different types of gestures as referential vs. non-referential. Since beats are typically postulated to serve as discourse highlighters, the later appearance of beats might be related to children's developing language and the ability to generate narratives, which requires more discourse organization.

In a longitudinal (within subject) study of five French–English bilingual children aged 2.5–3 years, Nicoladis et al. (1999) observed a correlation of beats with mean length of utterance (MLU) in the three children who used them. Beats were found to occur in longer and more complex utterances. This suggests that use of beats might be related to increased language proficiency. Similarly, Colletta et al. (2010) reported that French-speaking 6- to 10-year-old children used more gestures (both referential and non-referential) for discursive functions in narratives (i.e., gestures accompanying the use of connectors; performing an anaphoric function or highlighting linguistic units). Similar trends have been observed in children across various linguistic and cultural backgrounds, although there were differences in the frequency with which these discursive gestures were used (Colletta et al., 2015; Nicolas, Guidetti, & Colletta, 2017). In Colletta et al. (2010), beats were included in these discursive gestures, though they did not report the proportion of beats vs. other gestures used by the children. These findings seem to suggest that use of beat gestures might increase with longer utterances and complex discourse structure of narratives.

As for the gesture–speech link in children, Blake et al. (2008) examined three gestures (iconics, deictics, and beats) in 15 children aged 6–7 years (children with specific language impairment vs. typically-developing children). They found that iconic gestures (i.e., referential) were used most frequently, followed by deictics (i.e., referential), and then beats (i.e., non-referential). These three gesture types were reported to interact with grammatical categories. Iconics co-occurred more often with a noun phrase subject and a verb. Deictics tended to co-occur with a noun phrase subject and a prepositional phrase. Beats were more likely to co-occur with a noun phrase, a verb or a prepositional phrase than a conjunction or an adverb, suggesting that beats co-occur quite often on open class lexical grammatical categories. However, if children use beats for emphasis, as evidenced in the adult literature, it might be expected that beats would be most likely to

co-occur with pitch-accented words, rather than merely open class lexical items, some of which might not have a pitch accent.

In order to better understand what types of gestures are being used when, and how these develop, it is important to understand what is meant by the term 'beats'. However, each of the studies discussed above has used a slightly different criterion for identifying a 'beat'. For example, Nicoladis et al. (1999) described beats as meaningless, biphasic, up-down movements of the arms that have an emphatic function, e.g., making a 'beat' gesture in the air while the child says 'no go bed'. On the other hand, Blake and colleagues (2008) characterized beats as short, quick movements used for emphasis, repair or introduction of characters, e.g., a child raises hand while slightly twisting it and dropping it back while saying 'the *big* field'. In contrast, Colletta and colleagues (2010) focused more on discourse function, and classified beats along with other referential gestures as 'discursive' when they served to highlight linguistic units or mark discourse cohesion by linking clauses. Thus, beats have been variably described on the basis of the hand movements or the linguistic contexts they appear in: there has been no consensus regarding the physical characteristics of 'beats'. This makes it difficult to disentangle the movement characterization of beats from their linguistic role, let alone understand how and when they develop. One potential repercussion is that discourse functions could be attributed inappropriately to use of beats and hand movements that are probably not intended as gestures. This lack of explicit movement characterization further hinders investigation of the development of beat–speech synchrony.

The diversity of characteristics used to define 'beats' in the child studies echoes the various terms used in the adult literature. Labels such as 'beats' (Kendon, 2004; McNeill, 1985), 'batons' (Ekman & Friesen, 1972), and 'rhythmic gestures' (McClave, 1991) have all been used to refer to the physical properties of the movements and their 'emphatic' function. In a study of adult prosodic gesture organization in a lecture-type scenario, Yasinik, Renwick, and Shattuck-Hufnagel (2004) demonstrated that beats can be identified *as intentional or purposeful movements that do not often reflect contextual meaning, but are characterized by a well-defined stroke, the obligatory phase of the gesture, where the movement dynamics are manifested with the greatest clarity* (cf. Kendon, 2004). Based on the analysis of their corpus of three adults, they showed that gesture strokes tended to be aligned with pitch-accented syllables, supporting a discourse–structure–signaling role for stroke-defined beats. This explicit characterization of beats allows us to more consistently identify purposeful hand movements which do not reflect contextual meaning, and to investigate whether stroke-defined beats are used for emphasis in children's early discourse.

Children's discourse is often comprised of narration and explanation. As young as 2 years of age, children can provide short narratives of routine events (Eisenberg, 1985; Miller & Sperry, 1988; Nelson, 1978) and begin to use explanatory speech acts (Colletta & Pellenq, 2009; Veneziano & Sinclair, 1995). They continue to expand on these discourse skills as they reach school-age (Alamillo, Colletta, & Guidetti, 2013; Colletta et al., 2015; Hicks, 1990). It has been demonstrated that use of co-speech gestures may vary based on the type of discourse in which children are engaged. For example, Riseborough (1982) observed that 7-year-olds produced many iconic gestures in descriptions of artwork and in descriptions of motor activities, but almost none in story-telling.

According to Riseborough, when the topic of the task was difficult to verbalize, such as in the artwork or motor activity descriptions, more iconic gestures were observed. In contrast, when telling a recalled story, where there was no such challenge, fewer iconic gestures were used. It is thus possible that the nature of the discourse type can influence the use of other gesture types, such as beats, in young school-aged children. However, Blake et al. (2008) found no difference in the frequency of beats and iconics across story-retelling and classroom description tasks, which were all followed by probe questions. Similarly, Alamillo et al. (2013) found no task effect on how often children used gestures performing a ‘representational’ function (facial expressions and body movements to represent a concrete object) vs. a ‘discursive’ function (gestures to structure discourse inclusive of accentuation). It therefore remains unclear if and how discourse types might influence children’s tendency to use beats.

Even less clear is whether children around 6 years of age, when becoming more proficient at composing longer narratives, use beats for emphasis. During discourse (at least in English), emphasis is generally realized through the use of pitch accent (Bolinger, 1958). If beats are used for emphatic purposes, it is more likely they may occur on words with pitch accents, rather than on those without. Previous child studies have suggested that beats are indeed emphatic in nature (Blake et al., 2008; Colletta et al., 2010; Nicoladis et al., 1999). However, these findings were either based on the movement characteristics of the gesture (e.g., Blake et al., 2008) or on the assignment of a discourse role to the gesture (Colletta et al., 2010), and not on the co-occurrence of gesture strokes and pitch-accented words. One of the aims of the present study was therefore to determine if children approaching school age use stroke-defined beats in conjunction with pitch accents to serve a common discourse function (i.e., ‘emphasis’).

The goals of the present study were therefore to (1) establish whether English-speaking 6-year-olds use non-referential *stroke-defined beat gestures*, as proposed by Kendon (2004), (2) examine the influence of different discourse types on the use of these stroke-defined beats (i.e., story-retelling vs. exposition), and (3) determine the extent to which stroke-defined beats co-occur with (a) open class lexical items, or (b) pitch accents.

## Predictions

We hypothesized that at least some of the children would produce stroke-defined beats (cf. Kendon, 2004). Given that young children predominantly use referential gestures (Bates et al., 1979; Blake et al., 2008; de Bot & Schrauf, 2009; Nicoladis et al., 1999), and the strict characterization of beats used in this study (Yasinik et al., 2004), we also expected that 6-year-olds might produce more referential gestures than stroke-defined beats. Based on preliminary suggestions from Riseborough (1982), Blake et al. (2008), and Alamillo et al. (2013), we also hypothesized that there would be an effect of discourse type on the use of stroke-defined beats, with more being used during exposition than the story-retelling task. Since lexical words such as nouns and verbs are more likely to co-occur with beats (Blake et al., 2008), we also hypothesized that stroke-defined beats would occur more often with lexical words than function words. Finally, if beats are used for the purpose of emphasis (Blake et al., 2008; Colletta et al., 2010; Nicoladis

et al., 1999), we expected stroke-defined beats to co-occur more often with pitch-accented lexical words as is the case for adults (Krahmer & Swerts, 2007; Loehr, 2012; Yasinik et al., 2004).

## Method

### Participants

Participants included 12 monolingual Australian English-speaking 6-year-olds (7 boys, 5 girls) (age range: 5;3–7;5 years, mean: 6;3 years) (see Table 1 for participant ages and number/types of stroke-defined beats used). There were three 5-year-olds, five 6-year-olds, and four 7-year-olds. All were typically-developing children (as reported by their parents) with no prior history of language or hearing disorders. The children (and their caregivers) were recruited through advertisements and were paid for their participation in the study. Informed consent of the children and the caregivers was obtained from the caregivers. The study was approved by the Ethics Committee at Macquarie University.

**Table 1.** Number of referential gestures and stroke-defined beats produced by children during story-retelling and exposition tasks.

Participant	Age	Gender	Story-retelling		Exposition		Total
			Referential	Beats	Referential	Beats	
1	6;1	Male	0	0	4	0	4
2	7;4	Male	0	0	7	0	7
3	6;8	Female	0	0	8	0	8
4	5;6	Female	1	0	5	1	7
5	6;1	Female	7	1	2	3	13
6	6;4	Male	2	0	13	1	16
7	7;5	Male	13	3	9	2	27
8	5;5	Female	5	2	18	3	28
9	7;5	Female	18	0	5	6	29
10	5;3	Male	6	5	14	5	30
11	7;5	Male	6	11	23	0	40
12	6;8	Male	22	19	21	6	68
Total			80	41	129	27	277

### Tasks

The children completed two consecutive experimental tasks: a story-retelling task and an exposition task, both of which have been shown to elicit gestures in school-aged children (Alamillo et al., 2013; Alibali, Evans, Hostetter, Ryan, & Mainela-Arnold, 2009; Colletta et al., 2010; Novack, Congdon, Hemani-Lopez, & Goldin-Meadow, 2014). After extensive piloting, it was determined that the caregiver's presence, and lack of experimenter in the room, facilitated child-initiated conversation.

For the story-retelling task, the child watched a 2-minute cartoon video clip called 'Pigeons', which featured three baby pigeons and their failed attempts to reach a cupcake on a window ledge. The video clip contained only background music, without any speech. While the child was watching the movie, the caregiver was seated facing the wall listening to music through head-phones (Sennheiser HD280 pro). The task was then carried out as a game. The child saw the movie and then narrated the story to the caregiver, who was also encouraged to ask probe questions in order to elicit information from the child. The caregiver then completed a picture-puzzle based on the movie to provide motivation for the clarification questions.

For the exposition task, the child was provided with a hypothetical scenario in which the family were declared winners of a lottery and instructed to plan a family holiday trip to any destination of their choice with the money they had won. Here again, the caregiver was encouraged to prompt the child in order to elicit information and maintain interaction (question-answer format), though the child was encouraged to take the lead. However, unlike the narration task, the children found it difficult to verbalize on their own and therefore, the sample was largely elicited through the question-answer format. A total of 111 minutes of data (44 minutes from the story-retelling task and 67 minutes from the exposition task) were collected from the 12 children and analyzed.

### *Procedure*

The caregiver-child dyads were invited into a sound-attenuated room and were audio and video recorded during the tasks. Two tripod-mounted video cameras (Go-Pro Hero3) were set up facing one another, with a ceiling microphone (AKG C535EB) located in the middle of the room. Two low chairs were placed facing each other at an equal horizontal distance of 50 cm from the microphone in order to ensure good sound capture. The audio recording was sampled at 22.1 KHz, using SFS (Speech Filing System) software and the video recording was captured at a sampling rate of 120 frames per second. The camera recording the child was placed approximately 25 cm in front of the chair where the caregiver sat with unobstructed gestural space, while the camera recording the caregiver was placed just behind the chair where the child sat. This set-up provided an unobstructed view for the video capture of both child and caregiver.

The story-retelling task was carried out first, as it was easier and helped the children relax. At the end of the task, participants were given a 5-minute break before continuing with the exposition task. Each procedure and task lasted approximately 10 minutes.

### *Coding*

Each video sample was extracted and annotated for gestures and speech separately, for both child and caregiver. The following gestural coding scheme was adopted to address the research questions regarding children's use of stroke-defined beat gestures.

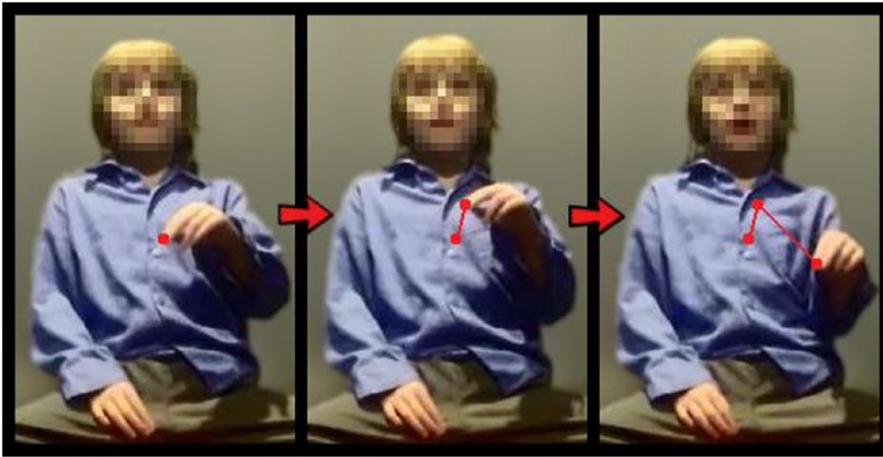


*Gestures.* One of the goals was to operationalize the identification of beat gestures in terms of gesture characteristics used in the adult literature, and to explore the potential discourse role of stroke-defined beats. Using ELAN (Lausberg & Sloetjes, 2009), the coder first viewed the video samples without audio input, to identify the regions containing gestures. For each identified gesture, the following gestural events were marked: (1) the beginning of the gesture, (2) the end of the gesture. To be consistent with the defining characteristics of stroke-defined beats in the adult literature (Yasinik et al., 2004), two additional gestural properties were included in annotating beats: (3) the stroke phase (beginning to end), and (4) the apex (end-point of maximum excursion). The stroke of a gesture is described as the phase of excursion in which the movement dynamics of ‘effort’ and ‘shape’ are clearly expressed (Kendon, 2004). In the case of stroke-defined beats, this was demarcated by the ‘effort’ and ‘cessation’ of movement. Similarly, as defined in Kendon (2004), ‘effort’ refers to clear movement trajectory; ‘cessation’ of movement refers to a sudden stop in the movement (cf. Yasinik et al., 2004). The stroke phase therefore ends with a peak in the movement excursion; this was coded as the apex of a beat gesture.

The visual identification of the stroke phase thus established the hand movement as a purposeful gesture. The stroke-defined gestures were then classified into types according to McNeill’s classification (1992). The audio input was switched on in this second stage of coding to distinguish stroke-defined beat gestures from referential gestures, based on the presence or absence of a semantic referent or contextual meaning. Therefore, if the stroke of the gesture indicated components of an accompanying referent, in a concrete or abstract manner, then the gesture was classified as one of the three referential types. For example, when a participant formed a bowl shaped hand gesture while saying ‘the nest was on the window sill’, this was considered an iconic gesture type. On the other hand, if the stroke did not reflect any meaning (e.g., sharp dip of hand as seen in Figure 1), then it was categorized as a stroke-defined beat. McNeill’s descriptions of the different gesture types are outlined below:

- a. Iconic: hand shapes/movements that reflect concrete relationship to the semantic content of speech, e.g., the hand makes several upward and downward movements while the child says ‘they jumped up and down to get the cupcake’.
- b. Metaphoric: hand shapes/movements that reflect abstract relationship to the semantic content of speech, e.g., open palms face the listener while the child says ‘just that’.
- c. Deictic: pointing movements that aid in locating entities/actions in space, e.g., a child raises his arm above his head and his index finger points to a location in right-side space while saying ‘it was up there’.
- d. Stroke-defined beats: intentional hand movements that do not reflect contextual meaning but have a well-defined stroke, e.g., both hands (held near the mid-point of the chest) move sharply down and then stop while the child says ‘a picture’.





**Figure 1.** Phases of a beat gesture produced by a participant with the rest, preparation, and stroke phases.

According to this coding, beats must meet two criteria: presence of ‘stroke’ and absence of semantic meaning. This coding yielded a total of 277 gestures for the 12 children. These four gesture types were subsequently grouped into (1) referential gestures ( $n = 209$ ), which included iconic ( $n = 133$ ), metaphoric ( $n = 27$ ), and deictic ( $n = 49$ ); and (2) stroke-defined beats ( $n = 68$ ). All the stroke-defined beat gestures produced by the children had four phases: rest, preparation, stroke, and recovery. Forty-one out of a total of 68 stroke-defined beats had a single stroke, and 27 had a hold phase resulting in multiple strokes. Thus 40% of these children’s stroke-defined beats had elaborate internal structure within the gestural phase. It was also noted that 5% of the referential gestures had multiple strokes, with one of the strokes having a beat component after the initial realization of the referential stroke component (e.g., a participant produced the stroke of an iconic gesture that resembled a bird nest which was followed by a hold phase, before the child produced a beat stroke, while maintaining the same hand shape throughout). Otherwise, the stroke-defined beat gestures were typically executed with relaxed, open, and spread finger hand-shapes.

Thirteen additional hand movements without a well-defined stroke phase (e.g., continuous circular movements) were annotated but excluded from the analysis, as these were deemed unintentional. Occasional gestures that occurred in segments of unintelligible speech, or during perceptually evident word search or dysfluency, were not coded. However, gestural realizations with a well-defined stroke phase that did not have accompanying speech (i.e., during a pause) were included in the analysis ( $n = 4$ ).

**Speech.** The audio recordings were extracted from the video and orthographically transcribed. The transcribed speech data were then divided into utterances (i.e., linguistic units defined by pauses at the beginning and the end). This was carried out with the aid

of speech waveforms and spectrograms using Praat (Boersma & Weenink, 2014). Acoustic criteria were used to delimit the utterance unit, and included: (1) continuous presence of spectral and intensity information within the utterance unit, and (2) absence of acoustic energy (i.e., no spectral and intensity information) with presence of unfilled pauses at the beginning and the end of the utterance unit.

The presence of pitch-accented words was identified only for those utterances where stroke-defined beat gestures occurred in each child sample. The apices of stroke-defined beat gestures were then checked to determine if they fell within a pitch-accented word. These pitch-accented words were then labeled according to the grammatical class they belonged to (i.e., content/lexical items [nouns/adjectives, verbs/adverbs] vs. function words [connectives, determiners, etc.]).

The speech and gesture coding was carried out for all samples by an experienced coder. A second coder then coded 23% of the speech samples for location of pitch accents and a third coder annotated 23% of gesture samples independently. Reliability between coders was 82% for the identification of pitch accents, 90% for the identification of beat gestures, and 87% for the identification of their strokes.

## Results

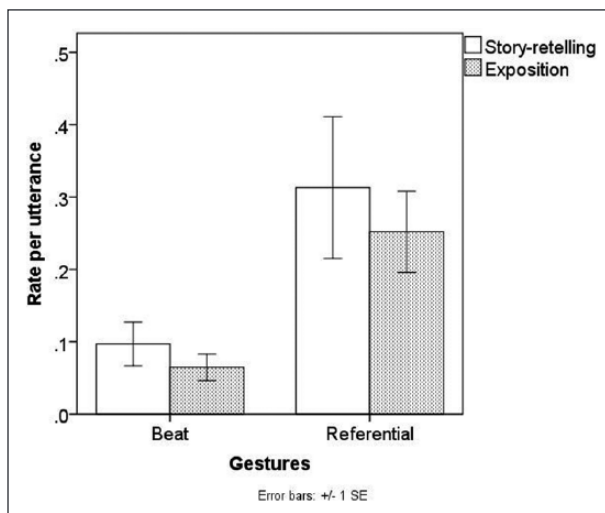
One of the main goals in this study was to investigate the possibility that, by the age of 6, at least some English-speaking children might be using beat gestures (as defined by a clear stroke phase) in either a story-retelling or an exposition task. This was indeed the case; out of the 12 children who participated in the study, nine produced stroke-defined beats. Table 1 presents the number of referential and stroke-defined beat gestures produced by all participants for each of the two tasks.

The subsequent analyses then focused on the subset of nine children who used some stroke-defined beats. We predicted that children might differ in their use of gestures, with more use of referential gestures than stroke-defined beat gestures. This was the case: out of the 258 gestures these nine children produced, most were referential; only 26% were stroke-defined beats ( $n = 68$ ). Of these nine children producing stroke-defined beats, 16 beats came from the 5-year-olds, 30 beats from the 6-year-olds, and 22 beats from the 7-year-olds. While the 5-year-olds and the 7-year-olds produced comparable stroke-defined beat gesture rates (0.072 and 0.074), the 6-year-olds produced the highest beat gesture rate of 0.12 (refer to Table 2 for individual beat gesture rates). Given the small number of participants ( $n = 3$ ) in each age bracket, we grouped all the participants together for statistical analysis. Non-parametric tests were used to analyze the data because they did not meet the normality criterion. Alpha was set at .05 for all analyses. To control for any effects of number of utterances on the frequency of gestures, we calculated the rate of occurrence for referential gestures and stroke-defined beats per utterance for each child (Figure 2). A Wilcoxon test was conducted to compare the rate of referential gestures and the rate of stroke-defined beat gestures. The analysis showed a significant difference between the two gesture types,  $Z = -2.666$ ,  $p = .008$ . Thus, the rate of referential gestures was higher than the rate of stroke-defined beat gestures, consistent with previous reports (Bates et al., 1979; Blake et al., 2008; de Bot & Schrauf, 2009; Nicoladis et al., 1999). To test whether the rate of stroke-defined beat and referential

gestures differed between the two discourse tasks, two Wilcoxon tests were then performed. No significant difference was found,  $Z = -.296$ ,  $p = .767$ , for stroke-defined beats, and  $Z = -.178$ ,  $p = .859$ , for referential gestures.

**Table 2.** Rate of beat gesture occurrence (per utterance) across participants. Participants are numbered with reference to Table 1.

Participant	Rate of occurrence
4	0.020
5	0.070
6	0.011
7	0.086
8	0.067
9	0.066
10	0.085
11	0.082
12	0.168

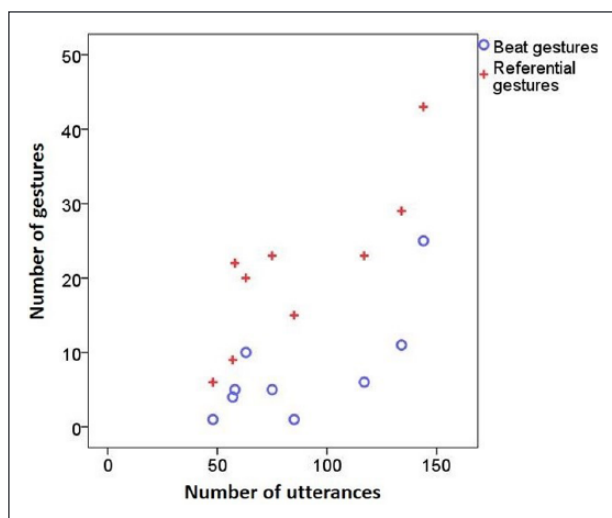


**Figure 2.** Rate of beat and referential gestures per utterance as a function of story-retelling vs. exposition task.

It is possible that the two different discourse tasks might yield a difference in the size of verbal output. Therefore, a Wilcoxon test was conducted to compare the number of utterances between the two tasks. The analysis again showed no significant difference between the number of utterances produced in the two tasks,  $Z = -1.718$ ,  $p = .086$ .

Although the overall number of utterances did not differ significantly between discourse tasks, the number of utterances could differ between individual child participants.

It is possible that children who spoke more might use more gestures and this could affect the amount of overall gestures (i.e., referential and beat gestures) in speech. Therefore, two Spearman's correlations were performed to determine whether referential and stroke-defined beat gestures co-varied with the amount of speech across individual children. There was a significant positive correlation between the number of utterances used and the number of referential gestures used ( $R_s(9) = .87, p = .002$ ), as well as a significant positive correlation between the number of utterances used and the number of stroke-defined beats used ( $R_s(9) = .714, p = .031$ ) (see Figure 3). This suggests that children who spoke more also used more gestures, with the implication that the frequency with which children use referential gestures and stroke-defined beats depends on the amount of overall speech.



**Figure 3.** Number of beat and referential gestures used as a function of the number of utterances produced per child.

Recall that Blake et al. (2008) found beats occurring most often with nouns and verbs (i.e., lexical forms). Beat strokes in the noun/adjective and verb/adverb categories were therefore collapsed into a broad 'lexical' category and those in connectives ('and', etc.) into a 'functional' category. A Wilcoxon test was conducted, showing a significant difference in the incidence of stroke-defined beats between lexical words and function words, with  $Z = -2.494, p = .013$ . As anticipated, there were more stroke-defined beats on the lexical words. However, as shown in Table 3, five out of nine children employed a total of 18 tokens of stroke-defined beats (26% of total stroke-defined beats) with functional categories as well.

Finally, we compared the co-occurrence patterns of stroke-defined beats with a pitch-accented word and those without one, using a Wilcoxon test. Although there were more stroke-defined beats with pitch accent ( $n = 43$ ) than without ( $n = 25$ ), no statistical

**Table 3.** Number (%) of beat strokes that co-occurred with pitch accent (PA) and lexical (LEX) words. Participants are numbered with reference to Table 1 and include only those who produced stroke-defined beats.

Participant	Total beats	With PA	LEX
4	1	1 (100)	1 (100)
5	4	0 (0)	4 (100)
6	1	1 (100)	0 (0)
7	5	3 (60)	5 (100)
8	5	3 (60)	5 (100)
9	6	3 (50)	4 (67)
10	10	10 (100)	6 (60)
11	11	9 (82)	7 (64)
12	25	13 (52)	18 (72)
Total	68	43 (63)	50 (74)

difference in the number of stroke-defined beats was found,  $Z = -1.723$ ,  $p = .085$ . This was counter to our prediction. Although stroke-defined beats did not seem to be associated with the use of pitch accent overall, they could be drawn to pitch accents in lexical words which are more likely to be assigned discourse emphasis. We therefore performed a further analysis only on stroke-defined beats in lexical words. A Wilcoxon signed-rank test showed that stroke-defined beats that aligned with lexical words did not occur more often with a pitch accent than without ( $Z = -1.761$ ,  $p = .078$ ). This finding was counter to our predictions that stroke-defined beats would be found to co-occur more often with pitch accent for emphatic purposes. It is possible that this is due to the protracted learning of when and how to use pitch accents (cf. Wells, Peppé, & Goulondris, 2004).

## Discussion

This study set out to determine if English-speaking 6-year-olds use *stroke-defined beats* during discourse tasks. It found that many did, using non-referential gestures that have a well-defined stroke phase, similar to that found in adults (cf. Yasinik et al., 2004). This is consistent with McNeill's (2005) proposal regarding the use of 'beats' by older children.

Overall, the children in this study produced stroke-defined beats less often than referential gestures. This is also consistent with findings from previous studies on both typically and atypically developing children (e.g., Blake et al., 2008; Nicoladis et al., 1999). This suggests that the 6-year-olds do not yet exhibit adult-like distributions of gesture use (i.e., 33%; McNeill, 1985). This is not surprising, as children have a long history of producing referential gestures before they are capable of meaningful verbal communication (Bates et al., 1979). This early tendency towards the use of referential gestures can be found when young children transition from the one-word stage of development to the multi-word stage in development (Özçalışkan & Goldin-Meadow, 2005). Referential gestures thus seem to lead/guide language output in early childhood. Perhaps the early

stages of language learning are concerned with learning new words (Rowe & Goldin-Meadow, 2009), and this biases children towards a higher percentage of referential gesture use in early speech.

However, when children become more competent communicators and produce longer and more complex sentences and discourses, there is a need to be able to use appropriate linguistic and gestural means to make the discourse coherent and cohesive. This is evident in Colletta et al.'s (2015) and Nicolas et al.'s (2017) findings, where children aged 10 years, who produced longer narratives than children aged 6 years, also produced more non-referential gestures. Although 6-year-olds used stroke-defined beats in the present study, they still exhibited abundant use of referential gestures when composing their discourse. This suggests that these children are still developing their skills in discourse structure at both the linguistic and gestural level. The fact that 26% of their gestures involved stroke-defined beats during child-to-adult conversation indicates that they are moving closer towards achieving the 33% distributional patterns seen in adults (McNeill, 1985).

Despite the small number of children in this study at each age, it is interesting that the 6-year-olds produced about 44% of the total stroke-defined beats, followed by the 7-year-olds (32%) and the 5-year-olds (24%). It appears that the 7-year-olds exhibited a distribution pattern of stroke-defined beats comparable to the adults (33% as reported in McNeill, 1985). This pattern of stroke-defined beat gesture use appears to be a function of the amount of verbal output, with the 7-year-olds producing a total of 309 utterances, the 6-year-olds a total of 286 utterances, and the 5-year-olds a total of 186 utterances. However, it must be kept in mind that this could be a random effect considering the small sample of children across these ages.

Our findings also indicate a significant correlation between spoken output and use of referential/stroke-defined beat gestures across children, similar to earlier reports (Colletta et al., 2015; Nicoladis et al., 1999; Nicolas et al., 2017). Children increasingly used gestures when they spoke more. This pattern is consistent with Levy and McNeill's (1992) suggestion that 'gestures and speech are different manifestations of one process in utterance generation. Thus, the principle of more quantity of expression applies to gestures as well as to speech, because it directly applies to this underlying process' (p. 299). It has also been reported that children use gestures in order to organize information in a multi-modal (verbal and non-verbal) format, such that it eases their cognitive load during the act of communication (Goldin-Meadow, 2000). This may help explain the findings in the present study, especially since the children were engaged in tasks that were spontaneous in nature and not well practiced. Alternatively, it might be the case that the children who talked more were more relaxed, and therefore tended to use more gestures, or perhaps they were simply more competent speakers.

Given previous reports suggesting that different discourse tasks might influence children's gesture use (Alamillo et al., 2013; Blake et al., 2008; Riseborough, 1982), we expected there might be more stroke-defined beat gestures in the exposition task than the story-retelling task. This is because the former reflected planning a future trip, explaining and analyzing (Berman & Nir-Sagiv, 2007; Mosenthal, 1985), whereas the latter reflected sequencing of a story they saw (McCabe & Bliss, 2003). However, once controlled for the overall number of utterances, the children used the *same* proportion of stroke-defined beats in both tasks. This is despite the fact that in the story-retelling task, children spoke

for a longer period of time without interruption from the conversational partner (the caregiver). In the exposition task, the caregivers frequently asked probe 'wh-' questions to which the children replied using shorter sentences. This might have alleviated the demand to plan and organize language and structure complex discourse during exposition. Therefore, it is important to take into account syntactic complexity and interlocutor interactions in future studies to identify possible effects of discourse type on the use of beat gestures to organize complex discourse. Perhaps it is possible to tease apart the effect of discourse type from interlocutor interactions by using a familiar caregiver vs. a non-familiar conversant as interlocutors in the same discourse interaction. Future studies could also examine the possibility that syntactic complexity varies with discourse types, and the extent to which this interacts with the use of stroke-defined beats.

It was found that three-quarters of the children in this study exhibited the use of stroke-defined beats, and that stroke-defined beats co-occurred significantly more often with open class lexical words than closed class function words, consistent with Blake et al. (2008). However, these stroke-defined beats did not always co-occur with pitch accents. Since pitch accent is an indicator of emphasis, this suggests that the link between stroke-defined beats and pitch accents might not yet be well established for these 6-year-old children, counter to McNeill's (2005) suggestion. Since open class lexical words (such as nouns, adjectives) are more likely to be the locus of contrastive/new information for emphasis and therefore attract the use of pitch accents and beats, it follows that beats will co-occur more often with a pitch accent in lexical words. Yet our findings did not support this. This suggests that children might not yet have established a close link between the use of stroke-defined beats (i.e., manual channel) and pitch accent (i.e., verbal channel) as a unified manifestation of the common discourse function/goal (i.e., 'emphasis').

This raises further questions for gestural models which assume an interface representation between gestures and speech, such as the Interface Hypothesis (Kita & Özyürek, 2003). According to this hypothesis, linguistic encoding will affect the type of information that gestures are organized to convey. For instance, Kita and Özyürek (2003) found that English speakers shaped an arc movement to accompany their use of the verb 'swing' in describing a scene where a character swung on a rope. But Turkish speakers did not use an arc movement or the verb 'swing' in their description of the same scene. The Turkish speakers used the verb 'go' instead, because Turkish does not have an agentive intransitive verb. If linguistic encoding in the verbal domain is going to influence how gestures are organized in the manual domain, as suggested in the Interface Hypothesis, we would have expected the use of pitch accent to co-occur necessarily with the beats to convey emphasis. As Kita and Özyürek's study (2003) focused on the iconic gestures and our study focused on the beat gestures, it is possible that the gesture-speech interface representation differs between the two gesture types. Perhaps pitch accent as a means to signal discourse meaning (e.g., emphasis) is related to the overall discourse structure, covering a longer stretch of speech, thereby increasing the cognitive load and influencing how beat gestures are generated in the manual domain. Since our findings showed that stroke-defined beats occurred more often with lexical words, this raises the possibility that gestures might lead/precede children's developing use of pitch accent for the purpose of 'emphasis', in a similar fashion as some gestures precede children's linguistic constructions (Özçalışkan & Goldin-Meadow, 2005). It would be interesting to probe the



developmental relationship between the use of stroke-defined beats and pitch accents in future studies to determine how and when these two channels become more consistently synchronized.

Blake et al. (2008) also reported higher instances of beats co-occurring with nouns, verbs, and prepositions, although there were differences in the coding schemes used between the two studies. In the present study we considered the co-occurrence of beat *apices* with words, whereas Blake et al. (2008) considered words that co-occurred with the *preparation (beginning) phase* of the gesture. This again raises the issue of how best to compare findings across studies. We have tried to show here that adopting a consistent definition for what constitutes a stroke-defined beat, and consistently coding for this across studies of both children and adults, is critical for better understanding of how and when children develop adult-like use of gesture.

## Conclusions

This study found that 6-year-olds use both referential and beat gestures when engaged in story-retelling and exposition tasks. Critically, it defined ‘beat’ as an *intentional movement with a well-defined stroke phase and no contextual meaning*. The occurrence of referential gestures and stroke-defined beats correlated with the number of utterances produced, suggesting that the use of both gesture types may become more frequent as children’s communicative competence develops. It was also found that 6-year-olds have not yet formed a link between stroke-defined beats and the use of pitch accent, despite the fact that stroke-defined beats were likely to occur with lexical words that tend to attract emphasis. These findings suggest several avenues for future research with a wider age range of children using a variety of tasks and interlocutors. It is hoped that the definition of *stroke-defined beats* adopted in this study will serve as an important benchmark, allowing for more effective comparison across studies, experimental conditions, and populations.

## Acknowledgements

We thank Stefanie Shattuck-Hufnagel and Pei Li Ren for helpful discussion and feedback. Preliminary versions of these results were presented at the 45th Annual Meeting of the Australian Linguistics Society (ALS), University of Newcastle, and at Speech Prosody 2016, Boston.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded, in part, by the following grants: NIH R01HD057606, ARC CE110001021, and ARC FL130100014.

## References

Alamillo, A. R., Colletta, J. M., & Guidetti, M. (2013). Gesture and language in narratives and explanations: The effects of age and communicative activity on late multimodal discourse development. *Journal of Child Language, 40*, 511–538.

- Alibali, M. W., Evans, J. L., Hostetter, A. B., Ryan, K., & Mainela-Arnold, E. (2009). Gesture–speech integration in narrative discourse: Are children less redundant than adults? *Gesture*, 9, 290–311.
- Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The emergence of symbols: Cognition and communication in infancy*. New York, NY: Academic Press.
- Berman, R. A., & Nir-Sagiv, B. (2007). Comparing narrative and expository text construction across adolescence: A developmental paradox. *Discourse Processes*, 43, 79–120.
- Blake, J., Myszczyzyn, D., Jokel, A., & Bebiroglu, N. (2008). Gestures accompanying speech in specifically language-impaired children and their timing with speech. *First Language*, 28, 237–253.
- Boersma, P., & Weenink, D. (2014). Praat: Doing phonetics by computer (Version 5.3.84) [Computer software]. Available from <http://www.praat.org/>
- Bolinger, D. L. (1958). A theory of pitch accent in English. *Word*, 14, 109–149.
- Butcher, C., & Goldin-Meadow, S. (2000). Gesture and the transition from one- to two-word speech: When hand and mouth come together. In D. McNeill (Ed.), *Language and gesture* (pp. 235–257). New York, NY: Cambridge University Press.
- Colletta, J. M., Guidetti, M., Caprici, O., Cristilli, C., Demir, O. E., Kunene-Nicolas, R. N., & Levine, S. (2015). Effects of age and language on co-speech gesture production: An investigation of French, American, and Italian children's narratives. *Journal of Child Language*, 42, 122–145.
- Colletta, J. M., & Pellenq, C. (2009). Multimodal explanations in French children aged 3 to 11 years. In N. Nippold & C. Scott (Eds.), *Expository discourse development and disorders: International perspectives* (pp. 63–97). New York, NY: Taylor & Francis.
- Colletta, J. M., Pellenq, C., & Guidetti, M. (2010). Age-related changes in co-speech gesture and narrative: Evidence from French children and adults. *Speech Communication*, 52, 565–576.
- de Bot, K., & Schrauf, J. (2009). *Language development over the lifespan*. New York, NY: Routledge.
- Eisenberg, A. (1985). Learning to describe past experiences in conversation. *Discourse Processes*, 8, 177–204.
- Ekman, P., & Friesen, W. V. (1972). Hand movements. *Journal of Communication*, 22, 353–374.
- Goldin-Meadow, S. (2000). Beyond words: The importance of gesture to researchers and learners. *Child Development*, 71, 231–239.
- Hicks, D. (1990). Genre skills and narrative development in the elementary school years. *Linguistics and Education*, 2, 43–73.
- Igualada, A., Esteve-Gibert, N., & Prieto, P. (2017). Beat gestures improve word recall in 3- to 5-year-old children. *Journal of Experimental Child Psychology*, 156, 99–112.
- Iverson, J. M., Caprici, O., & Caselli, M. C. (1994). From communication to language in two modalities. *Cognitive Development*, 9, 23–43.
- Kendon, A. (1996). An agenda for gesture studies. *The Semiotic Review of Books*, 8, 121–131.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge, UK: Cambridge University Press.
- Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, 48, 16–32.
- Krahmer, E. J., & Swerts, M. G. J. (2007). The effects of visual beats on prosodic prominence: Acoustic analyses, auditory perception and visual perception. *Journal of Memory and Language*, 57, 396–414.
- Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods, Instruments, & Computers*, 41, 841–849.
- Leonard, T., & Cummins, F. (2011). The temporal relation between beat gestures and speech. *Language and Cognitive Processes*, 26, 1457–1471.

- Levy, E. T., & McNeill, D. (1992). Speech, gesture and discourse. *Discourse Processes*, 15, 277–301.
- Loehr, D. (2012). Temporal, structural, and pragmatic synchrony between intonation and gesture. *Laboratory Phonology*, 3, 71–89.
- Marshall, C. R., & Morgan, G. (2015). From gesture to sign language: Conventionalization of classifier constructions by adult hearing learners of British Sign Language. *Topics in Cognitive Science*, 7, 61–80.
- Mayberry, R. L., & Nicoladis, E. (2000). Gesture reflects language development: Evidence from bilingual children. *Current Directions in Psychological Science*, 9, 192–196.
- McCabe, A., & Bliss, L. (2003). *Patterns of narrative discourse: A multicultural, life span approach*. New York, NY: Pearson Education.
- McClave, E. (1991). *Intonation and gesture* (Unpublished doctoral dissertation). Washington, DC: Georgetown University.
- McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review*, 92, 350–371.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.
- McNeill, D. (2005). *Gesture and thought*. Chicago, IL: University of Chicago Press.
- Miller, P., & Sperry, L. (1988). Early talk about past: The origins of conversational stories of personal experience. *Journal of Child Language*, 15, 293–315.
- Mosenthal, P. B. (1985). Defining the expository discourse continuum. *Poetics*, 14, 387–414.
- Nelson, K. (1978). How children represent knowledge of their world in and out of language: A preliminary report. In R. S. Siegler (Ed.), *Children's thinking: What develops?* (pp. 255–273). Mahwah, NJ: Lawrence Erlbaum.
- Nicoladis, E., Mayberry, R., & Genesee, F. (1999). Gesture and early bilingual development. *Developmental Psychology*, 35, 514–526.
- Nicolas, R. M., Guidetti, M., & Colletta, J. M. (2017). A cross-linguistic study of the development of gesture and speech in Zulu and French oral narratives. *Journal of Child Language*, 44, 36–62.
- Novack, M. A., Congdon, E. L., Hemani-Lopez, N., & Goldin-Meadow, S. (2014). From action to abstraction using the hands to learn math. *Psychological Science*, 25, 903–910.
- Özçalışkan, S., & Goldin-Meadow, S. (2005). Gesture is the cutting edge of early development. *Cognition*, 96, 101–113.
- Riseborough, M. G. (1982). Meaning in movement: An investigation into the interrelationship of physiographic gestures and speech in seven-year-olds. *British Journal of Psychology*, 73, 497–503.
- Rowe, M. L., & Goldin-Meadow, S. (2009). Early gesture selectively predicts later language learning. *Developmental Science*, 12, 182–187.
- Rusiewicz, H. L., Shaiman, S., Iverson, J. M., & Szuminsky, N. (2013). Effects of prosody and position on the timing of deictic gestures. *Journal of Speech, Language and Hearing Research*, 56, 458–470.
- Rusiewicz, H. L., Shaiman, S., Iverson, J. M., & Szuminsky, N. (2014). Effects of perturbation and prosody on the coordination of speech and gesture. *Speech Communication*, 57, 283–300.
- Veneziano, E., & Sinclair, H. (1995). Functional changes in early child language: The appearance of references to the past and of explanations. *Journal of Child Language*, 22, 557–581.
- Wells, B., Peppé, S., & Goulandris, N. (2004). Intonation development from five to thirteen. *Journal of Child Language*, 31, 749–778.
- Yasinik, Y., Renwick, M., & Shattuck-Hufnagel, S. (2004). *The timing of speech-accompanying gesture with respect to prosody*. *Proceedings of the International Conference: From sound to sense* (pp. 97–102). Cambridge, MA: MIT.