



HHS Public Access

Author manuscript

Gesture (Amst). Author manuscript; available in PMC 2016 January 04.

Published in final edited form as:

Gesture (Amst). 2009 ; 9(3): 290–311. doi:10.1075/gest.9.3.02ali.

Gesture–speech integration in narrative:

Are children less redundant than adults?

Martha W. Alibali^a, Julia L. Evans^b, Autumn B. Hostetter^c, Kristin Ryan^a, and Elina Mainela-Arnold^d

^aUniversity of Wisconsin — Madison

^bSan Diego State University

^cKalamazoo College

^dPennsylvania State University

Abstract

Speakers sometimes express information in gestures that they do not express in speech. In this research, we developed a system that could be used to assess the redundancy of gesture and speech in a narrative task. We then applied this system to examine whether children and adults produce non-redundant gesture–speech combinations at similar rates. The coding system was developed based on a sample of 30 children. A crucial feature of the system is that gesture meanings can be assessed based on form alone; thus, the meanings speakers express in gesture and speech can be assessed independently and compared. We then collected narrative data from a new sample of 17 children (ages 5–10), as well as a sample of 20 adults, and we determined the average proportion of non-redundant gesture–speech combinations produced by individuals in each group. Children produced more non-redundant gesture–speech combinations than adults, both at the clause level and at the word level. These findings suggest that gesture–speech integration is not constant over the life span, but instead appears to change with development.

Keywords

gesture; narrative; development; gesture–speech integration

Gesture and speech encode meaning in different ways. As McNeill (1992) has argued, gestures express meanings holistically, relying on images that are idiosyncratic and constructed by individual speakers at the moment of speaking. Speech, in contrast, expresses meanings in a linear, segmented fashion, utilizing words and grammatical devices that are

Authors' addresses Martha W. Alibali Department of Psychology University of Wisconsin — Madison 1202 West Johnson Street Madison, Wisconsin 53706 USA mwalibali@wisc.edu
Julia L. Evans School of Speech, Language and Hearing Sciences San Diego State University 5500 Campanile Drive San Diego, CA 92182–1518 USA jevans@mail.sdsu.edu
Autumn B. Hostetter Department of Psychology Kalamazoo College 1200 Academy Street Kalamazoo, Michigan 49006 USA autumn.hostetter@kzoo.edu
Kristin Ryan Department of Psychology University of Wisconsin — Madison 1202 West Johnson Street Madison, Wisconsin 53706 USA kristine.ryan@gmail.com
Elina Mainela-Arnold Department of Communication Sciences and Disorders Pennsylvania State University 401K Ford Building University Park, PA 16802 USA ezm3@psu.edu

socially codified. As a consequence, gesture and speech rarely express information that is completely redundant.

The degree of semantic overlap between gesture and speech can be conceptualized as falling along a continuum (Goldin-Meadow, 2003). At one end of the continuum are cases in which the information expressed in the two modalities overlaps a great deal. For example, consider a speaker who says, “I was holding a big box” and produces a gesture that mimes holding a big box. In this case, both modalities express the same idea, so the degree of redundancy between gesture and speech is high. The gesture also expresses additional nuances of meaning, such as information about the position of the hands as they hold the box, but the semantic information expressed in the two modalities is largely overlapping.

At the other end of the continuum are cases in which there is little or no semantic overlap between the two modalities. In one often-cited example, a speaker describing a scene from a Sylvester and Tweety cartoon said, “she chases him out again” while swinging her arm as if wielding a weapon (McNeill, 1992). In fact, the speaker was describing a scene in which Granny chases Sylvester while swinging an umbrella. In this example, the speaker expresses an aspect of the scene in gesture (swinging the umbrella) that she does not express at all in speech. Thus, in this case, the degree of redundancy between gesture and speech is low.

In past research, gesture–speech combinations in which the two modalities express largely non-overlapping semantic information have been referred to as “complementary” (McNeill, 1992), “supplementary” (Özçaliskan & Goldin-Meadow, 2005), “mismatching” (Church & Goldin-Meadow, 1986; Perry, Church, & Goldin-Meadow, 1988), or “non-redundant” (Alibali, Kita, & Young, 2000). Such combinations have been taken as evidence that gesture and speech work together in an integrated fashion to express speakers’ intended meanings (Kendon, 2004; McNeill, 1992). The aim of the present research is to investigate how gesture and speech work together to express meanings, and to examine whether there is change over development in how gesture and speech are integrated.

Why might speakers produce non-redundant gesture–speech combinations? Several ideas have been proposed. One possibility is that speakers produce non-redundant gesture–speech combinations simply because the two modalities have different expressive possibilities (see Kendon, 2004, for discussion of this issue). Some types of semantic information are readily expressed in gestures, but would be difficult or cumbersome to express in words. For example, consider information about spatial relationships (e.g., the relative positions of pieces of furniture in a living room), about irregular shapes (e.g., the shape of the state of Wisconsin), or about manner or path of motion (e.g., the path an animal runs in attempting to escape a predator). In each of these examples, a verbal description would be difficult to formulate, as well as difficult for a listener to understand. However, in each of these examples, the information could be readily expressed in gestures. Because the modalities have different expressive possibilities, speakers may utilize them to express different types of information.

Indeed, speakers appear to design their gesture–speech combinations so that each modality expresses content that is both most appropriate given its expressive potential, and most

effectively communicated in that modality. Several studies have shown that listeners glean important information about shape, size and spatial relations from speakers' gestures (Beattie & Shovelton, 1999a, 1999b; Melinger & Levelt, 2004). Moreover, when information about size is of high importance in the overall meaning being conveyed, it tends to be expressed uniquely in gesture, in non-redundant gesture–speech combinations (Beattie & Shovelton, 2006). The same is true when the listener does not share the speakers' knowledge about size (Holler & Stevens, 2007). Thus, speakers seem to deploy information across modalities in sensible ways, given the expressive and communicative potential of each modality.

Another possible reason why speakers produce non-redundant gesture–speech combinations is that they may have *multiple ideas activated* at the moment of speaking, and they express one in speech, and the other(s) in gesture. For example, when children explain Piagetian conservation tasks or mathematical equivalence problems, they sometimes express one problem-solving strategy in gesture and another in the accompanying speech (Church & Goldin-Meadow, 1986; Perry et al., 1988). Children who frequently produce such gesture–speech “mismatches” when explaining problems are thought to be in a transitional knowledge state in which they have multiple, unintegrated ideas about the concept underlying the problems (Goldin-Meadow, Alibali, & Church, 1993). The simultaneous activation of multiple ideas leads them to produce non-redundant gesture–speech combinations.

A similar argument has been made about adults' production of non-redundant gesture–speech combinations during problem solving. When multiple possible solution paths are considered at the same time, as at “decision points” in solving the Tower of Hanoi problem, adults frequently express different information in speech and gesture when they explain their reasoning aloud (Garber & Goldin-Meadow, 2002).

A final possibility is that speakers produce non-redundant gesture–speech combinations when they have *difficulties expressing ideas in words*. These difficulties may arise for a variety of reasons, such as because the information is not easily “packaged” in the linear, sequential fashion required for speech (e.g., Hostetter, Alibali, & Kita, 2007; Kita, 2000), or because the speaker has difficulty accessing appropriate lexical items (e.g., Krauss, 1998). If the target information is not successfully packaged in speech, or if a desired lexical item is not found, the intended information may be expressed in gesture instead. Thus, speakers may have difficulties in expressing ideas *in words*, but they do not necessarily have difficulties in *communicating* the intended information. Indeed, de Ruiter (2006) has argued that when speakers encounter problems in speech, they sometimes *compensate* for these problems by communicating information in gesture.

This possibility suggests that individuals who have smaller vocabularies or weaker verbal skills should be likely to rely on gestures to express some of the semantic content they wish to express. Thus, individuals with smaller vocabularies or weaker verbal skills should more frequently produce gestures that are not redundant with speech than individuals who have larger vocabularies or better verbal skills. Note that this does not require that speakers with smaller vocabularies or weaker verbal skills be more dysfluent when speaking than those

with larger vocabularies or better verbal skills. Such speakers might express their ideas in fluent, non-redundant gesture–speech combinations that reflect the fact that the two modalities have different possibilities for expressing meanings.

If non-redundant gesture–speech combinations stem from difficulties expressing ideas in speech, then it may be the case that children are more likely to produce such non-redundant combinations than adults. Children may not have the vocabulary to express all of the ideas they wish to express, or they may not be able to produce the complex grammatical forms needed to express some of their ideas. As a consequence, they may rely on gesture to do some of the “work” of communicating the ideas they wish to express. Indeed, there is evidence that, in the early stages of language acquisition, children who can produce only single words in speech often combine those words with gestures that express other ideas (e.g., pointing to a cup and saying “Mommy”) (Butcher & Goldin-Meadow, 2000). If the child cannot produce the necessary word (in this example, “cup”), or is not yet able to produce a two-word utterance, the child may rely on non-redundant gestures to express part of the intended meaning.

Along similar lines, older children may also frequently rely on gesture to express some aspects of the meanings they wish to communicate. Little is known, however, about how older children integrate gesture and speech, or about how children’s gesture–speech integration compares to that of adult speakers. One goal of the present research is to systematically compare the frequency of non-redundant gesture–speech combinations in children and adults.

We hypothesize that children may produce more non-redundant gesture–speech combinations than adults. However, we do not mean to imply that speakers with full mastery of the language should *always* produce redundant gesture–speech combinations. Indeed, as noted above, speakers express content in each modality that is appropriate given the expressive potential of that modality. We also do not mean to imply that fully redundant gesture–speech combinations reflect mastery in language production. We only wish to assert the possibility of the converse, namely, that non-redundant combinations *may* stem from difficulties expressing ideas in speech, as one of several possible sources. We expect that adult speakers will produce some non-redundant gesture–speech combinations; however, we expect that children will produce them more frequently than adults.

One challenge in studying gesture–speech redundancy across the lifespan, however, is the lack of a task and associated coding system suitable for evaluating gesture–speech integration in individuals of different ages. The tasks that have been utilized to date are primarily problem solving tasks, including Piagetian conservation tasks (e.g., Church & Goldin-Meadow, 1986), mathematical equivalence problems (e.g., Perry et al., 1988), balance scale problems (Pine, Lufkin, & Messer, 2004; Thurnham & Pine, 2006), Piagetian bending rods problems (Stone, Webb, & Mahootian, 1991), constant change problems (Alibali, Bassok, Solomon, Syc, & Goldin-Meadow, 1999), factoring problems (Wagner, Nusbaum, & Goldin-Meadow, 2004), and the Tower of Hanoi problem (Garber & Goldin-Meadow, 2002). Systems for coding gesture–speech integration in problem explanations have been developed for each of these tasks. The basic procedure used in developing a

coding system is to (1) identify problem solving strategies for the tasks, (2) identify the key features of gestures that commonly occur during explanations of each strategy, and (3) assign each gesture a meaning based on the context in which it typically occurs. The system development is carried out with one sample, and the system is then applied to a new sample. The meaning of the gestures produced by each participant in the new sample can be evaluated without access to the accompanying speech, based on similarity in form to the gestures identified in the system based on the original sample. Finally, the gestures' meanings can be compared to the semantic content of the co-occurring speech to determine whether gesture and speech express redundant information.

Although this process has yielded useful information about gesture–speech integration in problem explanations for the tasks listed above, these tasks cannot be used to compare the gesture–speech redundancy of samples that differ substantially in terms of age, education, or problem-solving ability. For instance, the coding system developed by Church and Goldin-Meadow (1986) for identifying gesture–speech mismatches in explanations of Piagetian conservation tasks would likely reveal very few mismatches if used with an adult sample. Although this difference could be due to the older age and better vocabularies of the adult sample, a more likely cause would be the adults' better understanding of the task itself. It seems then, that, in order to gain a better understanding of how gesture–speech redundancy changes over the lifespan, we require a task that can be understood and completed in a similar way by individuals of different ages.

A narrative task in which participants retell the events of a cartoon or story is a good candidate for such a task. Narrative tasks have been used extensively by researchers interested in gesture (e.g., McNeill, 1992; Rauscher, Krauss, & Chen, 1996), and such tasks have been used successfully to elicit gestures from children as young as 2 1/2 years old (McNeill, 1992). Furthermore, there are reports of individuals producing non-redundant gestures during narrative tasks (such as the “chasing” example provided at the outset of this paper). However, most of the evidence on non-redundant gestures in narrative has involved isolated examples; as yet, there is no systematic procedure for coding non-redundant gesture–speech combinations in any narrative task. Thus, a primary goal of this research was to develop a coding system that could be used to reliably evaluate the relationship between gesture and speech in a narrative task.

In sum, then, this research had two primary goals: (1) to develop a coding system that can be used to evaluate the semantic integration of gesture and speech in a narrative task, and (2) to compare the frequency of non-redundant gesture–speech combinations in children and adults. In the following section, we describe the development and application of the coding system. In the section that follows, we examine whether the frequency of non-redundant gesture–speech combinations differs in children and adults.

Development of the coding system

Method

Participants—Participants were 30 children (ages 5;0–10;0).

Stimulus—The experimental stimulus was a 90-second episode of the German children’s cartoon, *Die Sendung mit der Maus*, which features an elephant and a mouse. The episode was one of those used by Alibali and Don in their research on gesture in children’s narratives (Alibali & Don, 2001). In the cartoon, the mouse jumps onto a high bar, swings back and forth, flips around, and then lands on the ground. The elephant tries to jump onto the bar, and in doing so, he bends the bar down. The mouse then tries to fix the bar, but is unable to do so. A leprechaun with a tall hat enters and walks under the bar. As the leprechaun’s hat passes under the bar, it pushes up on the bar and fixes it. The cartoon includes music but no words. See Table 1 for a complete list of all the cartoon’s events.

Procedure—Each child viewed the cartoon twice, and then narrated what happened in the cartoon to a listener. Children’s speech was transcribed. All gestures were identified and the words that coincided with each gesture were noted. Gestures were defined as movements of the hands or arms that were produced as part of the effort of speaking. These included (a) *representational gestures*, which are movements that depict semantic content via the shape, placement, and/or motion trajectory of the hands or arms (e.g., a circular motion of the index finger to depict SPIN), (b) *deictic gestures*, which are movements that indicate objects or locations (e.g., touching the head to indicate HAT), and (c) *beat gestures*, which are motorically simple, rhythmic gestures that do not depict semantic content related to speech (e.g., moving the hand, palm up, up and down while speaking). In most cases, the hand(s) returned to rest position after each gesture. When multiple gestures were produced in succession without the hand(s) returning to rest position, the boundaries between gestures were determined based on changes in the handshape, motion, or placement of the hands.

Development of the gesture coding system

To develop the system for coding whether gesture–speech combinations were redundant or non-redundant, we followed the procedure that was used by Church, Perry and Goldin-Meadow to develop systems for coding the meanings of gestures in conservation and mathematical equivalence tasks (Church & Goldin-Meadow, 1986; Perry et al., 1988). We used this procedure to develop a lexicon of gesture forms and their associated meanings for the cartoon narrations.

In the first step of this procedure, we characterized the handshape, orientation, and motion trajectory of each gesture that accompanied children’s verbal narrations. We then provided a tentative interpretation for each gesture in light of the speech that accompanied the gesture, and in light of our knowledge about the events that occurred in the cartoon (see Table 1). Thus, in this first step, gesture was coded *with speech*. For example, one gesture involved moving the index finger across neutral space in a back-and-forth, arcing trajectory, and the accompanying speech described swinging. This gesture was interpreted as meaning “swing”.

In the second step of the procedure, we identified sets of gestures that had been assigned the same meaning in the first step. For all sets that included at least five gestures, we then extracted the common features that characterized most or all of the gestures. For example, there were eight gestures in the database that had been interpreted as “swing” in the first step of coding. We examined this subset of gestures and determined that all eight of them

included a back-and-forth trajectory. Some used a point handshape and others used an open handshape, but all involved a back-and-forth trajectory.

The common features that were identified for each subset of gestures were used to create a lexicon of gestures that express particular meanings, characterized in terms of gesture form. For example, SWING gestures involve a back-and-forth motion.

Results

A total of 13 gesture categories were defined using this procedure (see Table 2). Importantly, these gestures can be interpreted *without reference to the accompanying speech*, based on their physical form alone. Thus, gesture and speech can be coded *independently*, and the semantic content expressed in each modality can be compared in order to assess the relationship between speech and gesture.

Note that our ability to code gesture and speech independently hinges on the fact that different speakers use gestures that are largely similar to express similar meanings. This suggests that speakers may rely on a repertoire or “vocabulary” of gestures that are readily recognizable. Indeed, some of the gestures in our system are likely to be readily recognized regardless of context (e.g., a circular spinning motion to mean SPIN). However, many of the other gestures in our system are readily interpreted only within the constraints provided by the context, specifically, by the particular cartoon we used. For example, a grasping handshape hovering in neutral space is readily interpreted as “grab bar” when it is produced within a narration of this particular cartoon, but this same gesture would be interpreted differently in many other contexts. Our ability to code gesture and speech independently depends on the systematicity and recognizability of the gestures that speakers produce *when narrating this particular cartoon*.

Experiment

Our next step was to apply this coding system to examine whether children and adults differ in the frequency with which they produce non-redundant gesture–speech combinations. Note that the sample of children included in this experiment was an entirely new sample (different from the sample used for development of the coding system).

Method

Participants—Participants were 17 children, ages 5;6–10;0 (8 boys, 9 girls), and 20 college students (5 male, 15 female). The college students were randomly chosen from a larger sample of adults who completed the narrative task for another study (Hostetter & Alibali, 2007). Children (or their parents, depending on the parents’ preference) were paid for their participation, and college students received extra credit in Introduction to Psychology.

Procedure—Each child retold the cartoon to an experimenter who waited outside the room while the child viewed the cartoon. The children were told that the experimenter had not seen the cartoon, in an effort to encourage them to include more information in their narrations. Adults retold the cartoon to an experimenter who busied herself with other tasks

(paperwork) while the adult viewed the cartoon. Both children and adults completed the cartoon narration task as part of a larger series of tasks. Children were told that the researchers were studying “how children think and talk”, and adults were told that the study addressed how people remember and communicate information.

Each participant viewed the cartoon stimulus twice, and then retold the story to an experimenter. Children received four prompts to encourage them to tell more about the story. The prompts were: (1) tell a little bit more about what happened when the mouse was first on the bar, (2) tell a little bit more about when the mouse’s friend tried to jump on the bar, (3) tell a little bit more about what the mouse did to try to fix the bar, (4) tell a little bit more about the “magic man” (we used this term because we expected some children to be unfamiliar with the word “leprechaun”). We did not use prompts with the adults, because their narratives were typically thorough enough that such prompts would have been redundant. Participants’ speech was transcribed. All gestures were identified and the words that coincided with each gesture were noted.

Coding gestures—Gestures that met the criteria for one of the 13 categories in the coding system were identified as *gestures with codable meanings*; these gestures were further coded in terms of the relationship between gesture and speech (see below). The gestures that did not have codable meanings included (a) *beat* gestures, which are motorically simple gestures that do not express semantic content, (b) representational gestures that conveyed meanings not represented in our lexicon (e.g., TRUNK or HANG),¹ and (c) representational gestures that expressed meanings from our lexicon (as inferred based on the accompanying speech) but that did so in a way that did not meet our formal criteria for interpreting the gestures in the absence of speech. Table 3 presents the frequency of beat and representational gestures (with and without codable meanings) for children and adults.

Coding the gesture–speech relationship—For gestures with codable meanings, we assessed the relationship between gesture and speech at two levels. At the *word* level, we examined whether the *exact words* that accompanied the gesture expressed the meaning assigned to the gesture. For example, one child said, “He went around and around and around the bar” while producing a gesture meaning *SPIN*, which coincided with the words “he went around and around”. In this example, both the gesture and the co-occurring words convey the meaning “spin”, so this gesture was coded as redundant with speech at the word level. At the *clause* level, we examined whether the *clause* in which the gesture occurred expressed the meaning that was coded in the gestures. For example, one child said, “He tried to push the bar back up” and produced a gesture meaning *PUSH BAR UP*, which coincided with the words “he tried”. Because the meaning of the gesture appeared in the clause as a whole, even though it did not appear in the exact words that accompanied the gesture, this gesture was coded as redundant with speech at the clause level. Note that all gestures that are redundant at the word level are also redundant at the clause level. Examples are presented in Table 4.

¹Included in this subcategory are metaphoric gestures that used locations in space to represent characters from the cartoon; these were fairly rare and they occurred only among the adults.

To assess reliability of the gesture coding procedures, the gestures of four children and seven adults were recoded by another coder. For children, agreement was 95% for identifying gestures from the stream of manual movement, and 98% for determining whether or not the meaning of each gesture was codable within our coding system. For gestures that both coders agreed could be assigned meanings, agreement was 90% for assigning one of the 13 meanings from our lexicon to each gesture. Agreement was 88% for categorizing gestures with codable meanings as redundant or non-redundant with the accompanying words, and 97% for categorizing gestures with codable meanings as redundant or non-redundant with the accompanying clause. For adults, agreement was 85% for identifying gestures from the stream of manual activity, and 93% for determining whether or not the meaning of each gesture was codable in our coding system. For gestures that both coders agreed could be assigned meanings, agreement was 96% for assigning one of the 13 meanings to each gesture. Agreement was 90% for categorizing gestures with codable meanings as redundant or non-redundant with the accompanying words, and 95% for categorizing gestures with codable meanings as redundant or non-redundant with the accompanying clause.

Coding dysfluencies in speech—Because we hypothesized that non-redundant gesture–speech combinations might be produced because of difficulties expressing ideas in speech, we also coded dysfluencies in participants’ speech. We identified several types of dysfluencies (Levelt, 1983), including: (a) filled pauses, such as “um” and “uh”; (b) repetitions, in which one or more words were simply repeated (e.g., “he was, he was swinging in the air”), (c) repairs, in which one or more words within a given syntactic frame were altered (e.g., “he started going, flipping real fast over that”), (d) fresh starts, in which the speaker shifted to a new syntactic frame (e.g., “and then this, then he was mad”). For each participant, we tallied the total number of dysfluencies, divided by the total number of words produced, and multiplied by 100 to yield a rate of dysfluencies per 100 words.

Results and discussion

Before examining the relationship between gesture and speech, we first compared the overall gesture rate (including both gestures with codable meanings and gestures without codable meanings) for children and adults. Past work on this issue suggested that adults gesture more than children. For example, Mayberry, Jacques and DeDe (1998) compared the proportion of spoken words accompanied by gestures in a small sample of 11-year-old boys and adults, and found that the adults produced more than three times as many words with gestures as the children. As seen in Figure 1, in this study, adults gestured at a higher rate than did children, but this difference was not significant, $t(35) = 1.59, p = .12$.

We then turned to the main question of interest: Did children produce non-redundant gesture–speech combinations at a higher rate than adults? To address this question, we calculated the proportion of gestures with codable meanings that were non-redundant with the accompanying speech, both at the word level and at the clause level. We chose to examine the data at both levels, in part because findings about whether gestures coincide with speech may hinge on how verbal phrases are constructed (e.g., if a speaker said “he pushed the bar up” with a PUSH gesture on “pushed”, this would be coded as redundant, but

if the speaker said “he tried to push the bar up” with a PUSH gesture on “tried”, this would be coded as non-redundant at the word level but redundant at the clause level). Note that participants who produced no gestures with codable meanings were excluded from this analysis, due to having a denominator of zero. Five adults and six children were excluded on this basis. Of these, one child and one adult were excluded because they produced only gestures that did not have codable meanings in our system, and five children and four adults were excluded because they did not produce any gestures at all during their narratives.²

As seen in Figure 2, children produced a greater proportion of non-redundant gesture–speech combinations than adults, both at the word level, $t(24) = 2.88, p < .01$, and at the clause level, $t(24) = 2.92, p < .01$. Thus, children were more likely than adults to convey information in their gestures that they did not also explicitly convey in speech.

Were the types of non-redundant gesture–speech combinations produced by children similar to those of adults, or was the non-redundancy different in some important way? To find out, we examined each gesture–speech combination that was non-redundant at the clause level, and classified it into one of five categories: (1) gesture is produced in silence (e.g., one child produced a SPIN gesture in silence *after* having described the character spinning on the bar in speech); (2) gesture creates a referent for deictic speech (e.g., “he just goes like that” with a gesture meaning GRAB BAR); (3) gesture makes the meaning of the accompanying speech more specific (e.g., “he does some gymnastic moves” with a gesture meaning SPIN; “this man in green comes by” with a gesture meaning WALK); (4) gesture elaborates the meaning of the accompanying speech, providing additional detail (e.g., “his top hat just bumps the bar a little bit” with a gesture meaning UP); (5) gesture expresses information that differs substantially in meaning from the accompanying speech (e.g., “[he] tried pushing it back into place” with a gesture meaning JUMP; “the mouse gets all angry” with a gesture meaning GRAB BAR). Table 5 presents the proportion of non-redundant gestures that fell into each category for children and for adults. For adults, most non-redundant gestures either elaborated or specified the meaning of the accompanying speech. For children, the nature of non-redundant gestures was more diverse.

For each gesture–speech combination that was non-redundant at the clause level, we also examined whether the information expressed uniquely in gesture was expressed in the immediately following speech (i.e., in one of the two following clauses). For adults, 75% of gestures that were non-redundant at the clause level were “resolved” in the following speech in this way (50% in the next clause, 25% two clauses later). The comparable proportion for children was only 28% (all in the next clause). Thus, adults eventually expressed in speech most of the information that they initially expressed uniquely in gesture, but this did not hold for children.

²It is of interest whether speech differed in the speakers who used no gestures, relative to participants who did use gestures (see, e.g., Melinger & Levelt, 2004). Among both children and adults, those who used no gestures produced fewer words than those who produced gestures. The difference was not significant for adults, $M = 112.5$ vs. $M = 179.6, t(18) = 1.62$, but it was for children, $M = 106.2$ vs. $M = 178.9, t(15) = 2.38, p < .05$. The narratives of the non-gesturers tended to have a list-like character, listing the events that occurred in the cartoon, without much additional commentary or interpretation. However, the sample of non-gesturers in each age group was too small to draw firm conclusions about differences in narrative structure or content.

Finally, we considered whether non-redundant gesture–speech combinations might be associated with difficulties expressing ideas in speech. Children produced dysfluencies at a significantly higher rate than adults ($M = 4.63$ vs. $M = 2.58$), $t(35) = 2.70$, $p < .01$ (one-tailed). However, the correlation between dysfluency rate (per 100 words) and proportion of gestures that were non-redundant was small and not significant for the children ($r(11) = .12$). The comparable correlation could not be tested in the adult sample, because there was little variability in the proportion of non-redundant gestures for adults.

We also examined whether children produced more dysfluencies with clauses accompanied by non-redundant gestures than with clauses accompanied by redundant gestures. There were 7 children who produced gestures of both types; these children were indeed more dysfluent with clauses with non-redundant gestures than with clauses with redundant gestures ($M = 0.75$ vs. $M = 0.16$ dysfluencies), $t(5) = 1.86$, $p = .06$, one-tailed.

General discussion

This research had two primary goals. First, we sought to develop a coding system that could be used to examine gesture–speech integration in a narrative task. The crucial feature of the system is that gesture meanings can be assessed without the accompanying speech. Second, we sought to compare patterns of gesture–speech integration in children and adults. We found that children produced more non-redundant gesture–speech combinations than adults, both at the clause level and at the word level.

Why did children produce more non-redundant gesture–speech combinations than adults? As noted in the introduction, there are at least three, non-mutually-exclusive reasons why speakers might produce such combinations. First, speakers may have *multiple ideas activated* at the moment of speaking, and they may express one idea in speech, and the other(s) in gesture. Some of the non-redundant gesture–speech combinations we observed had this character; for example, one child said, “[the mouse] tried pushing it back into place” with a gesture meaning JUMP. This child seemed to be simultaneously thinking about jumping up to reach the bar and pushing the bar up. However, many of the non-redundant gesture–speech combinations we observed did not reflect multiple ideas in any obvious way (e.g., combinations in which gesture elaborates the meaning of the accompanying speech or makes it more specific). Furthermore, at first glance, this explanation seems incompatible with the finding that children and adults differ in their production of non-redundant gesture–speech combinations. It is possible that children are more likely to activate multiple ideas at one time than are adults; however, other data that corroborate this point are needed.³

Second, speakers may produce non-redundant gesture–speech combinations because the two modalities have different expressive possibilities (see, Kendon, 2004). It is possible that the

³Studies of gesture–speech “mismatches” in problem solving tasks have demonstrated that frequent mismatches are an indicator that a child is “ready to learn” about the task, and thus in a transitional or unstable knowledge state with respect to the task (Alibali & Goldin-Meadow, 1993; Church & Goldin-Meadow, 1986; Perry et al., 1988). Along similar lines, in the narrative task, frequent production of non-redundant combinations that reflect multiple ideas might suggest that the narrator is in a transitional or unstable cognitive state regarding his or her narration. The narrator may have multiple, competing ideas about what to focus on at various points in the story. It is worth noting, however, that in the present study, we observed relatively few gesture–speech combinations that clearly manifested multiple, competing ideas.

expressive possibilities of the verbal modality may change with the skill level of the speaker. With children's more limited verbal skills, the expressive possibilities of the verbal channel may be more limited, and children may rely more frequently on the expressive potential of gesture in order to communicate their intended meanings (see de Ruiter (2006) for a related argument regarding aphasic speakers).

A third, related possibility is that speakers may produce non-redundant gesture–speech combinations because they have *difficulties expressing ideas in words*. For participants in both groups, knowledge about the cartoon was acquired non-verbally (i.e., by watching the video), and the task required participants to convert that knowledge into verbal form. Relative to adults, children may have more difficulty expressing their ideas in verbal form; in support of this idea, we found that children produced speech dysfluencies at a higher rate than adults. Furthermore, children were more dysfluent when they produced non-redundant gestures than when they produced redundant gestures, providing suggestive evidence that difficulties in the verbal channel are associated with non-redundant gestures. Given children's greater difficulty expressing their ideas in the verbal channel, it seems unsurprising that they would more frequently express aspects of their intended meanings in gestures. The age difference in non-redundant gestures is natural from this perspective.

Current theories of gesture production suggest at least two possible loci at which children's difficulties with verbal expression may lead to production of non-redundant gesture–speech combinations. First, children have smaller vocabularies than adults, so they may not have lexical items that fully express their intended meanings. They may produce non-redundant gesture–speech combinations as they (unsuccessfully) attempt to access appropriate lexical items (see Krauss, 1998). Second, children have less experience than adults at “packaging” visuo-spatial information into verbal form. Such packaging may be challenging for children, and they may produce non-redundant gesture–speech combinations as they attempt (not entirely successfully) to package information into speech (see Alibali et al., 2000; Kita, 2000).

If difficulties with verbal expression lead children to produce non-redundant gesture–speech combinations, then younger children should produce more such combinations than older children. The sample of children in the present study is too small to allow a direct test of this hypothesis, but we intend to examine this issue in future research. It also seems likely that children with developmental language impairments characterized by difficulty with verbal expression may be more likely than age-matched peers to produce non-redundant gesture–speech combinations. Similarly, one might expect that individuals who have strong spatial visualization skills but weak verbal skills would also be more likely than speakers with other combinations of skills to produce non-redundant gesture–speech combinations. Future research is needed to test these hypotheses.

Of course, some limitations of this work must be acknowledged. First, the system was developed on children, and later applied to both children and adults. A system that was developed on adults might allow more of the adults' gestures to be assigned meanings. However, we do not expect that a more inclusive system would reveal a dramatically different picture of adults' gesture–speech integration, because our data suggest that most

adult gestures express ideas that are also expressed in the accompanying speech. A second limitation is that our sample was small, and because some participants did not produce any gestures, the sample for the final comparisons was smaller still. The fact that we found significant group differences even with such a small sample is striking.

In brief, we have shown that children produce more non-redundant gesture–speech combinations than adults in their narrative retellings of a cartoon story. These findings indicate that gesture–speech integration is not constant over the life span. Instead, patterns of gesture–speech integration appear to change, in the direction of greater redundancy between modalities, as children’s language skills develop. Our data suggest that children are indeed less redundant than adults.

Acknowledgements

This research was supported by grants from the Spencer Foundation (S133-DK59; Julia Evans & Martha Alibali, Co-PIs) and the National Institute on Deafness and Other Communication Disorders (R01 DC 6560–01, Julia Evans, PI).

We thank Lisbeth Simon, Maureen Kaschak, and Tracy Barrett for assistance collecting the child data, Kristi Kalmoe for assistance developing the coding system, and Julia Lin, Le’ Andrea Vernon, Nancy Demerdash, and Liza Meredith for assistance with data transcription and coding. We also thank Kerri O’Brien for her contributions to the pilot work that led up to this study, and Marisa Sizemore and Donna Thal for helpful discussions of this research.

About the authors

Martha W. Alibali is Professor of Psychology at the University of Wisconsin, Madison. Her research investigates mechanisms of knowledge change in cognitive development and learning, with a special focus on the role of gesture in processes of change. She is also interested in the role of gesture in communication, particularly in educational settings. She is the author of more than 40 journal articles and book chapters, and her work has been or is currently funded by the National Science Foundation and the Institute of Education Sciences.

Julia L. Evans is Professor in the School of Speech, Language, and Hearing Sciences, Faculty in the Joint Doctoral Program (San Diego State University & University of California, San Diego), and Director of the Child Language and Cognitive Processes Lab at San Diego State University. Her research focuses specifically on language and cognitive processing in children with Specific Language Impairments (SLI) and the extent to which language and cognitive representations in SLI differ from those in typically developing children. Her research assumes that knowledge is graded in its representation and emerges developmentally within an embodied cognitive framework that encompasses social, emotional, perceptual, and motor representations. Her work indicates that both language and cognitive representations in children with SLI are under-specified, require more resources to maintain, and are more vulnerable to external processing demands than those of typically developing children. She is the author of more than 40 journal articles and book chapters, and her work has been or is currently funded by the Spencer Foundation and the National Institutes of Health.

Autumn B. Hostetter is Assistant Professor in the Psychology Department at Kalamazoo College. Her research examines the cognitive underpinnings of speech-accompanying gestures, particularly as they relate to theories about the embodiment of cognition. She is also interested in the sources of individual differences in both producing and comprehending gestures. Her recent publications include: Visible embodiment: gestures as simulated action (*Psychonomic Bulletin & Review*, 2008); I see it in my hands' eye: representational gestures reflect conceptual demands (*Language and Cognitive Processes*, 2007); and Raise your hand if you're spatial! Relations between verbal and spatial skills and representational gestures (*Gesture*, 2007).

Kristin Ryan received a B.S. degree in Communicative Disorders from the University of Wisconsin, Madison, where she worked with Julia Evans and Martha Alibali on this research and several other related projects. She went on to earn a B.S. degree in Nursing, and she now works as a registered nurse.

Elina Mainela-Arnold is Assistant Professor in the School of Communication Sciences and Disorders at the Pennsylvania State University. Her main interests are in the cognitive mechanisms involved in language acquisition and use, in both typically developing children and children with language impairments. Her recent publications include: Understanding conservation delays in children with SLI: Task representations revealed in speech and gesture (*Journal of Speech, Language, & Hearing Research*, 2006); and Lexical representations in children with SLI: evidence from a frequency-manipulated gating task (*Journal of Speech, Language, & Hearing Research*, 2008).

References

- Alibali, Martha W.; Bassok, Miriam; Solomon, Karen O.; Syc, Sharon E.; Goldin-Meadow, Susan. Illuminating mental representations through speech and gesture. *Psychological Science*. 1999; 10:327–333.
- Alibali, Martha W.; Don, Lisa S. Children's gestures are meant to be seen. *Gesture*. 2001; 1:113–127.
- Alibali, Martha W.; Goldin-Meadow, Susan. Transitions in learning: what the hands reveal about a child's state of mind. *Cognitive Psychology*. 1993; 25:468–523. [PubMed: 8243044]
- Alibali, Martha W.; Kita, Sotaro; Young, Amanda. Gesture and the process of speech production: we think, therefore we gesture. *Language and Cognitive Processes*. 2000; 15:593–613.
- Beattie, Geoffrey; Shovelton, Heather. Do iconic hand gestures really contribute anything to the semantic information conveyed by speech? An experimental investigation. *Semiotica*. 1999a; 123(1/2):1–30.
- Beattie, Geoffrey; Shovelton, Heather. Mapping the range of information contained in the iconic hand gestures that accompany spontaneous speech. *Journal of Language and Social Psychology*. 1999b; 18:438–462.
- Beattie, Geoffrey; Shovelton, Heather. When size really matters. *Gesture*. 2006; 6:63–84.
- Butcher, Cynthia; Goldin-Meadow, Susan. Gesture and the transition from one- to two-word speech: when hand and mouth come together. In: McNeill, David, editor. *Language and gesture: window into thought and action*. Cambridge University Press; Cambridge: 2000. p. 235–257.
- Church, Ruth Breckinridge; Goldin-Meadow, Susan. The mismatch between gesture and speech as an index of transitional knowledge. *Cognition*. 1986; 23:43–71. [PubMed: 3742990]
- de Ruiter, Jan-Peter. Can gesticulation help aphasic people speak, or rather, communicate? *Advances in Speech Language Pathology*. 2006; 8:124–127.
- Garber, Philip; Goldin-Meadow, Susan. Gesture offers insight into problem solving in adults and children. *Cognitive Science*. 2002; 26:817–831.

- Goldin-Meadow, Susan. *Hearing gesture: how our hands help us think*. Harvard University Press; Cambridge, MA: 2003.
- Goldin-Meadow, Susan; Alibali, Martha W.; Ruth Breckinridge Church. Transitions in concept acquisition: using the hand to read the mind. *Psychological Review*. 1993; 100:279–297. [PubMed: 8483984]
- Holler, Judith; Stevens, Rachel. The effect of common ground on how speakers use gesture and speech to represent size information. *Journal of Language and Social Psychology*. 2007; 26:4–27.
- Hostetter, Autumn B.; Alibali, Martha W. Raise your hand if you're spatial: relations between verbal and spatial skills and gesture production. *Gesture*. 2007; 7:73–95.
- Hostetter, Autumn B.; Alibali, Martha W.; Kita, Sotaro. I see it in my hands' eye: representational gestures reflect conceptual demands. *Language and Cognitive Processes*. 2007; 22:313–336.
- Kendon, Adam. *Gesture: visible action as utterance*. Cambridge University Press; Cambridge: 2004.
- Kita, Sotaro. How representational gestures help speaking. In: McNeill, David, editor. *Language and gesture: window into thought and action*. Cambridge University Press; Cambridge: 2000. p. 162-185.
- Krauss, Robert M. Why do we gesture when we speak? *Current Directions in Psychological Science*. 1998; 7:54–60.
- Levelt, Willem J. M. Monitoring and self-repair in speech. *Cognition*. 1983; 14:41–104. [PubMed: 6685011]
- Mayberry, Rachel I.; Jacques, Joselynn; DeDe, Gayle. What stuttering reveals about the development of the gesture–speech relationship. In: Iverson, Jana M.; Goldin-Meadow, Susan, editors. *The nature and function of gesture in children's communication*. Jossey-Bass; San Francisco, CA: 1998.
- McNeill, David. *Hand and mind: what gestures reveal about thought*. University of Chicago Press; Chicago: 1992.
- Melinger, Alissa; Levelt, Willem J. M. Gesture and the communicative intention of the speaker. *Gesture*. 2004; 4:119–141.
- Özçaliskan, Seyda; Goldin-Meadow, Susan. Gesture is at the cutting edge of early language development. *Cognition*. 2005; 96:B101–B113. [PubMed: 15996556]
- Perry, Michelle; Church, Ruth Breckinridge; Goldin-Meadow, Susan. Transitional knowledge in the acquisition of concepts. *Cognitive Development*. 1988; 3:359–400.
- Pine, Karen J.; Lufkin, Nicola; Messer, David. More gestures than answers: children learning about balance. *Developmental Psychology*. 2004; 40:1059–1067. [PubMed: 15535756]
- Rauscher, Frances H.; Krauss, Robert M.; Chen, Yihsiu. Gesture, speech, and lexical access: the role of lexical movements in speech production. *Psychological Science*. 1996; 7:226–231.
- Stone, Addison; Webb, Rebecca; Mahootian, Shahrzad. The generality of gesture speech mismatch as an index of transitional knowledge: evidence from a control-of-variables task. *Cognitive Development*. 1991; 6:301–313.
- Thurnham, Angela J.; Pine, Karen J. The effects of single and dual representations on children's gesture production. *Cognitive Development*. 2006; 21:46–59.
- Wagner, Susan M.; Nusbaum, Howard; Goldin-Meadow, Susan. Probing the mental representation of gesture: is handwaving spatial? *Journal of Memory and Language*. 2004; 50:395–407.

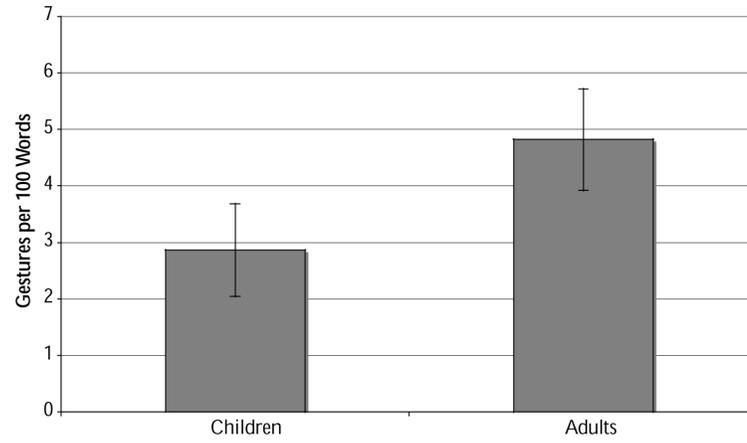


Figure 1. Rate of gestures per 100 words for children and adults. The error bars represent standard errors.

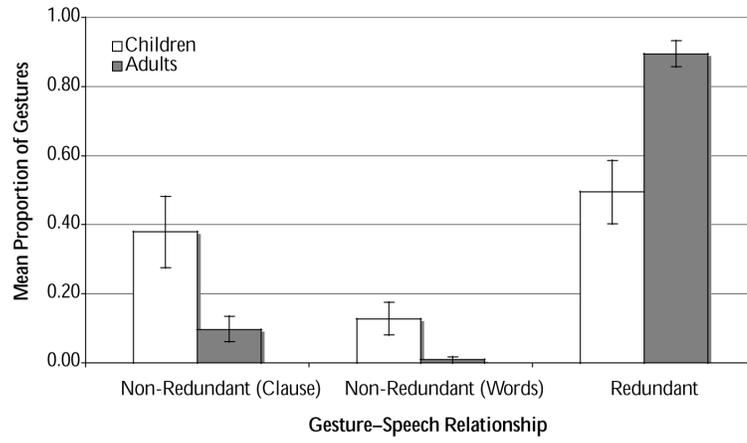


Figure 2. Proportion of gestures with codable meanings with each type of gesture–speech relationship. The error bars represent standard errors.

Table 1

Events in the cartoon

-
1. Mouse appears.
 2. Mouse walks up to bar.
 3. Mouse blinks.
 4. Mouse stands underneath bar.
 5. Mouse stretches / warms up.
 6. Mouse jumps up and grabs bar.
 7. Mouse swings.
 8. Mouse makes grunting noises.
 9. Mouse spins around bar.
 10. Elephant enters.
 11. Elephant watches mouse on bar.
 12. Mouse jumps off.
 13. Elephant “claps”.
 14. Mouse turns around and sees elephant.
 15. Elephant walks up to bar.
 16. Elephant jumps, trying to get on bar.
 17. Elephant grabs bar with trunk.
 18. Elephant’s trunk bends the bar.
 19. Mouse gets upset.
 20. Mouse scolds elephant.
 21. Elephant hangs for a couple seconds.
 22. Elephant gets off the bar.
 23. Mouse gestures for elephant to get back.
 24. Mouse tries to fix the bar.
 25. Leprechaun with top hat appears.
 26. Mouse gives up and steps back.
 27. Leprechaun walks under bar.
 28. Bar is magically repaired by leprechaun’s hat.
 29. Leprechaun walks off.
 30. Elephant laughs.
 31. Mouse pouts and looks embarrassed.
-

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Gesture lexicon: Meanings and descriptions of gesture forms

Gesture meaning	Description of gesture form
Swing	Gesture that includes a back and forth motion; may be produced with hands or legs
Spin	Gesture that includes a circular motion, typically repeated and in neutral space; may be produced with one or both hands
Bar	Gesture that traces or takes the form of the bar; one or both hands may point, flatten, or form O's to represent round shape of bar
Stand	Gesture that traces or takes the form of the stand; one or both hands (typically both) point or flatten to represent upright stand for bar
Bar + stand	Gesture that traces or takes the form of the bar + stand; typically produced with both hands; points may trace shape of stand and bar, or hands (with fingertips together) bend at knuckles or wrists so fingers represent bar and palms or arms represent stand
Grab bar	Gesture in which hands hover in parallel, sometimes with grasping motion; typically produced with both hands, either in neutral space or above head
Bent bar	Gesture in which hands trace shape of bent bar <i>or</i> half of bent bar, <i>or</i> gesture in which hands hover while holding shape of bent bar; may be produced with one or both hands
Dismount	Gesture in which hand makes a downward arcing motion; may include a slight upward motion before the downward motion; typically produced with one hand
Hat	Gesture made on or above head, in which hands either trace hat shape, form hat shape with hands or point to (imaginary) hat; may be produced with one or both hands using either points or flat handshapes
Jump	Gesture in which hands move up and down several times; typically produced in neutral space or in lap; may be produced with one or both hands, using either flat or curved open handshapes
Push bar up	Gesture in which both hands, palms face up or out, move up; typically produced in high neutral space or above head
Up	Gesture in which one hand moves up, in either point, flat, or curved open handshape; typically produced in neutral space
Walk	Gesture that includes alternate stomping motion; can be produced with feet or hands

Table 3

Frequency of representational and beat gestures for children and adults

Gesture Type	Frequency	
	Children	Adults
Beat	5	27
Representational (total)	95	130
Meaning in lexicon (codable)	66	59
Meaning not in lexicon (not codable)	29	71

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4

Examples of different patterns of gesture–speech integration in children and adults

Participant Group	Speech	Gesture	Gesture Meaning	Redundant with words?	Redundant with clause?
Adult	And there's a uh like a bar	Right hand point makes circles	SPIN	No	No
Adult	And his top hat just bumps the bar a little bit and pushes it up	Right hand palm toward front, moves up	UP	No	Yes
Adult	But instead of trying to push it upwards	Right hand palm up, moves up	PUSH UP	Yes	Yes
Child	And then he finally got up and started going faster	Right hand point makes circles	SPIN	No	No
Child	The um elephant was too heavy and the bar made a u-shape	Both hands palms up, wrists touching to make a U-shape	BENT BAR	No	Yes
Child	And he put it up with the top of his hat	Right hand touches head	HAT	Yes	Yes

Note. Words in bold represent the words being spoken at the moment the gesture was produced.

Table 5

Proportion of non-redundant gestures that occurred in each type of gesture– speech combination for children and adults

Combination Type	Proportion of Non-redundant Gestures	
	Children	Adults
Gesture produced in silence	.11	.00
Gesture creates referent for deictic speech (e.g., “like this”)	.17	.00
Gesture makes meaning of speech more specific	.28	.13
Gesture elaborates meaning of speech	.28	.63
Gesture and speech express substantially different meanings	.17	.25
N	18	8

Note. Table values are for gesture–speech combinations that are non-redundant at the clause level.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript