

**The Emerging Gesture-Speech Relationship in Preschoolers Who Do and Do Not
Stutter**

Laiiah Factor
Lisa Gershkoff-Stowe
Julie D. Anderson

Indiana University

The Emerging Gesture-Speech Relationship in Preschoolers Who Do and Do Not Stutter

When children speak, they often use deictic (pointing) and iconic (representational) gestures to support their spoken message. Much research has centered on what role gesture plays in language learning and how gesture and speech evolve to form a single, integrated communicative system (Capone & McGregor, 2004; McNeill, 1992). Far from being a mere compensatory mechanism, gesture can be supplemental and redundant to speech, helping to convey meaning across multiple modalities and providing insight into a child's underlying cognitive processes across development (Goldin-Meadow, 2000; 2009). By providing a window into development, gesture can shed light on the course of typical development, but it can also reveal the first signs of when development has gone awry. As such, the study of gesture can provide a new understanding of both typical and atypical development.

In mastering spoken language, it is not uncommon for typically-developing children to experience occasional disfluencies during speech production. However, some children produce more disfluencies than others and when these disfluencies are predominantly "stuttering-like" in nature, these children are typically diagnosed with developmental stuttering. To our knowledge, only one study to date has examined the relationship between gesture and speech in adults and children who stutter (Mayberry, Jaques, & DeDe, 1998). The results of this study suggest that the presence of chronic speech disfluencies affect the relationship between gesture and speech in that people who stutter produced significantly less gestures overall and the production of these gestures "waited" for fluent speech; however, these conclusions were primarily based on a sample of 12 adults (6 adults who stutter) as well as four 11-year-old children (2 children who stutter).

In the present study, we examine what happens to gesture in the presence of disfluencies in preschool children's spontaneous speech. Of particular interest is how developmental stuttering may affect the temporal and semantic relationship between gesture and speech. We know from the adult literature that gestures are often coordinated with speech in both timing and meaning (Wagner, Malisz, & Kopp, 2014). When adults speak, their gestures are synchronized in time such that gesture production typically occurs just ahead of or simultaneous to the spoken message. Evidence from Mayberry and her colleagues (1998) suggests that when adults and older children are chronically disfluent, the timing of gesture production changes to accommodate for the lack of speech fluency. If this also holds true for preschool-aged children, then it stands to reason that any disruption in the forward flow of speech ought to affect the temporal and semantic relationship between gesture and speech. That is, we would expect to see differences in this relationship during disfluent versus fluent speech. Given that children who stutter (CWS) tend to produce more and different types of disfluent speech than children who do not stutter (CWNS) (Yairi & Ambrose, 2005), we also expect to see between-group differences in the temporal and semantic alignment of gesture and speech. In particular, we predict that CWS will produce more non-overlapping, or supplementary, gesture-speech combinations than CWNS, based on the assumption that children who have difficulty with fluent speech production depend more heavily on gesture to express meaning (Alibali, Evans, Hostetter, Ryan, & Mainela-Arnold, 2009).

Method

Participants

Participants were 18 children, 3;4 to 5;9 years of age, with no history of neurological, speech-language (other than stuttering), hearing, or intellectual problems. CWS and CWNS

were matched by age (± 5 months), gender (2 girls, 7 boys per group) and socio-economic status. The mean age for CWS was 51.4 months ($SD = 8.93$) and 53.3 months ($SD = 9.04$) for CWNS, $t(16) = -0.45, p = .66$. All children scored $> 16^{\text{th}}$ percentile on speech-language tests and passed a hearing screening. Children were classified as CWS or CWNS based on the frequency of stuttered disfluencies (part-word repetitions, single-syllable word repetitions, sound prolongations, and/or blocks) produced during a parent-child conversational interaction. A Mann-Whitney test revealed that CWS exhibited significantly more stuttered disfluencies than CWNS, $z = -3.49, p < .001$.

Procedure

Each child viewed a 7-minute cartoon (*Canary Row*, Warner Brothers) twice, first as a whole and then broken down into 18 brief segments. Immediately following each 20-30 second segment, the child narrated what happened in the cartoon to their parent who could not see the screen. Feedback was not given other than to encourage the child to stay on task and no mention of gesture was made by the experimenter.

The spontaneous speech and gestures produced by the children during the story narration were transcribed and analyzed using ELAN software. Speech was coded as either fluent or disfluent. Disfluent speech included both “normal” and “stuttering-like” disfluencies as well as a mix of both types. Normal disfluencies included interjections (e.g., “and um it broke”), repetitions (e.g., “and then he-and then he bumped his head), and revisions (e.g., “he hit-she hits her with the umbrella”). Stuttering-like disfluencies included part-word repetitions (e.g., “a-a-and it was stuck”), single-syllable word repetitions (e.g., “he-he-he rolled down”), and sound prolongation (e.g., “I’ll fffind it”) (Yairi & Ambrose, 2005).

Gestures were isolated and assigned a meaning gloss based on hand shape, location, and/or movement (McNeill, 1992). Both deictic and iconic gestures were analyzed. Deictic gestures involved the use of the index finger, palm, or other body parts to indicate reference. Iconic gestures consisted of hand movements that depicted images of concrete entities or actions (Kendon, 2004).

Gesture and speech were coded for whether they were co-expressive in both time and meaning (McNeill, 1992). Gestures were temporally synchronous to speech when the onset of the stroke phase coincided with the onset of the corresponding verbal utterance; they were asynchronous in time when the stroke phase of the gesture either preceded the co-occurring speech or followed it (as measured in milliseconds). Additionally, gestures were semantically supplementary, or overlapping, when their meaning matched that of speech; they were non-overlapping when the content of the gesture added new information to the verbal message.

Results

We first report the mean number of fluent words, disfluencies, and gestures produced by CWS and CWNS during the narrative task. The data are presented in Table 1. As expected, we found that CWS were significantly less fluent in their speech than CWNS, $t(16) = 3.04, p < .01$ (one-tailed). Additionally, CWS produced significantly more “stuttering-like” disfluencies ($t(16) = 3.40, p < .005$ one-tailed) compared to CWNS, who produced more normal disfluencies ($t(16) = 1.70, p < .05$ one-tailed).

Following Mayberry et al. (1998), we also expected CWS to gesture significantly less than their typically developing peers. This prediction was not confirmed, however. Although CWS produced many fewer gestures, on average, than CWNS, this difference was not reliable, $t(16) = 1.18, p = .25$. Instead, we found considerable variability in the rate of gesture production,

particularly among CWS. Further analysis revealed that the majority of the gestures were iconic gestures, as shown in Table 2; this was true for both groups of children (CWS = $p < .005$; CWNS = $p < .001$). Deictic gestures, although occurring less frequently, were relatively more common among CWS than CWNS. However, this difference was only marginally significant ($p = .08$).

In the series of analyses that follow, we investigate possible differences in the temporal and semantic relationship between gesture and speech in CWS and CWNS. We first examine the timing of gestures during fluent and disfluent speech, collapsed across disfluency type. We then examine whether gesture and speech carry the same or different semantic content.

The Temporal Relationship between Gesture and Speech

As previously described, there are three ways in which the timing between gesture and speech might occur: gesture before speech, gesture simultaneous to speech, and gesture after speech. A consistent finding in the adult literature is that speakers generally produce gesture just before or concurrent to speech (Wagner et al, 2014). To test this prediction in CWS and CWNS, we compared the mean proportion of gestures occurring after speech (“Speech First”) to gestures occurring either before or simultaneous to speech (“Not Speech First). Interestingly, and contrary to adult findings, a 2 (Group) x 2 (Order) mixed analysis of variance (ANOVA) revealed a significant Group by Order interaction, $F(1,16) = 6.39, p = .02$. Specifically, for CWS, “Speech First” gestures ($M = .77, SD = .16$) were more frequent than “Not Speech First” gestures ($M = .23, SD = .16$). Although we found no main effect of Group, there was a trend towards a similar temporal ordering of gesture and speech for CWNS (“Speech First”: $M = .57, SD = .16$; “Not Speech First”: $M = .43, SD = .17$).

Additional within-subject comparisons of temporal order were conducted to examine possible effects of fluency for both CWS and CWNS. One question that arises is whether

normal disfluencies might impact the temporal ordering of gesture and speech in CWNS. However, a 2 (Order) x 2 (Fluency) repeated ANOVA revealed no main effects or interaction (see Figure 1). The other key question is whether chronic stuttering-like disfluencies affect the timing of gesture and speech. Based on Mayberry et al.'s (1998) findings, we expected the disfluent speech produced by CWS to disrupt the temporal execution of their gestures. Although we found the expected main effect of Order for CWS, $F(1, 7) = 20.25, p = .003$, surprisingly, there was no effect of Fluency (see Figure 2).

The Semantic Relationship between Gesture and Speech

An additional question we address is whether the information conveyed by speech and gesture varied as a function of chronic disfluency. We predicted that CWS would rely more heavily on gestures that supplemented (as opposed to reinforced) their speech compared to CWNS, possibly as a way to compensate for chronic disruptions to the forward flow of speech. Gestures were coded as either “overlapping” (i.e., redundant to speech) or “non-overlapping” (i.e., non-redundant or supplemental to speech). A 2 (Group) x 2 (Overlap) mixed ANOVA indicated a main effect of Overlap, $F(1, 16) = 13.17, p = .002$ but no main effect of Group or a Group by Order interaction. Both groups of children produced more semantically non-overlapping gestures (CWS $M = .63, SD = .20$; CWNS $M = .65, SD = .11$) than overlapping gestures (CWS $M = .37, SD = .20$; CWNS $M = .35, SD = .11$). This result is consistent with the findings reported by Alibali and her colleagues (2009), who also found that typically developing children rely more on gestures to supplement their speech than adults who use gesture to reinforce their speech.

As before, we conducted subsequent analyses to compare fluent to disfluent speech for CWS and CWNS, separately. We predicted that when CWS experienced a disruption in the flow

of speech, they would rely on non-overlapping gestures to convey meaning. A 2 (Overlap) x 2 (Fluency) repeated ANOVA conducted on the mean proportion of semantically overlapping and non-overlapping gesture-speech combinations for CWS indicated only a marginally significant effect of Overlap, $F(1, 7) = 4.33, p = .08$. The data are shown in Figure 3. A similar analysis conducted on CWNS also found a main effect of Overlap, $F(1, 7) = 4.33, p = .03$, but no effect of Fluency (see Figure 4).

Discussion

The purpose of this study was to explore the temporal and semantic relationship between gesture and speech in chronically disfluent preschool-aged children. Compared to their typically developing peers, we predicted that CWS would demonstrate a disruption in the temporal coordination between gesture and speech when narrating the contents of a cartoon to their parents. However, contrary to predictions, we found that the onset of speech preceded the onset of gesture in both groups of children, regardless of fluency. This result fails to replicate Mayberry et al. (1998) who found that chronic stuttering-like disfluencies interrupted the timing of gesture relative to speech. However, differences in the age and severity of their participants may account for some of the disparity in the findings. Whereas our focus was on young preschoolers during episodes of fluent and disfluent speech, Mayberry et al. (1998) specifically emphasized stuttering-like disfluencies in children who were considerably older. As a result, these older children may have developed compensatory habits that affected the temporal coordination of gesture and speech.

The temporal pattern of CWS and CWNS in the present study is also markedly different from that of adult speakers who most often produce gestures that precede or are simultaneous to speech (Wagner et al., 2014). In light of the adult literature, our results suggest that the

integration of gesture and speech emerges over time, possibly in response to developments in the ability to retrieve and order information within a narrative context.

Our final major prediction was that CWS would rely more heavily on gestures to supplement their disfluent speech relative to their more fluent peers. Alibali, Evans, et al. (2009) reported finding that typically developing children, aged 5-10 years of age, produced more non-overlapping speech-gesture combinations compared to adults. They concluded that younger speakers may have difficulty expressing their ideas or accessing words from lexical memory and thus would be more likely to rely on non-redundant information to convey meaning. Although our initial hypothesis was not supported, the present findings extend Alibali et al.'s results by showing that CWS as well as CWNS produced gestures that were non-redundant, not only during moments of fluent speech but during disfluent speech as well. Taken together the results suggest that the semantic function of gesture may shift over the course of language development as the structure and operation of the lexicon changes.

Although additional studies that include larger numbers of CWS are needed to replicate and extend the present findings, the results suggest intriguing new insights into the development of the gesture-speech system and the complimentary role of gesture in the expression of verbal meaning.

References

- Alibali, M. W., Evans, J. L., Hostetter, A. B., & Mainela-Arnold, E. (2009). Gesture-speech integration in narrative discourse: Are children less redundant than adults?. *Gesture*, 9(3), 290-311.
- Capone, N. C. & McGregor, K. K. (2004). Gesture development: A review for clinical and research practices. *Journal of Speech, Language, and Hearing Research*, 47, 173-186.
- Goldin-Meadow, S. (2000). Beyond words: The importance of gestures to researchers and learners. *Child Development*, 71(1), 231-239.
- Goldin-Meadow, S. (2009). How gesture promotes learning throughout childhood. *Child Development Perspectives*, 3, 106-111.
- Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge, MA: Harvard University Press.
- Mayberry, R., Jaques, J., & DeDe, G. (1998). What stuttering reveals about the development of the gesture-speech relationship. *New Directions for Child Development*, 79, 77-87.
- McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press.
- Wagner, P., Malisz, Z., & Kopp, S. (2014). Gesture and speech in interaction: An overview. *Speech Communication*, 57, 209-232.
- Yairi, E., & Ambrose, N. (2005). *Early Childhood Stuttering*. Austin, TX: Pro-Ed.

Table 1

Mean number of gestures, fluent words and disfluencies by CWS and CWNS

	Gestures	Fluent Words**	SLD Disfluencies**	ND Disfluencies*	Mix Disfluencies
CWS	22.11 (18.14)	260.11 (77.67)	39.78 (25.64)	24.22 (15.68)	12.44 (23.18)
CWNS	32.33 (18.53)	413.56 (130.14)	9.56 (5.75)	38.11 (18.88)	.78 (.97)

* $p < .05$ ** $p < .01$

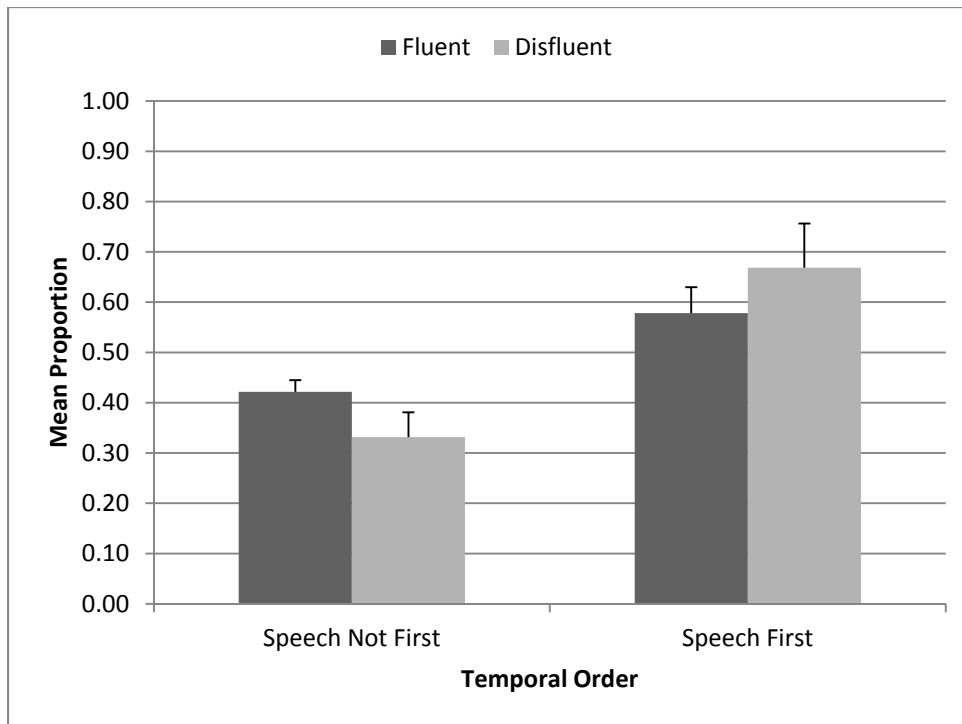
Table 2

Mean frequency and proportion of iconic and deictic gestures by CWS and CWNS

	Total Gestures	Iconic	Deictic
CWS	199	148 (.74)	51 (.26)
CWNS	291	268 (.92)	23 (.08)
Total	490	416 (.85)	74 (.15)

Figure 1

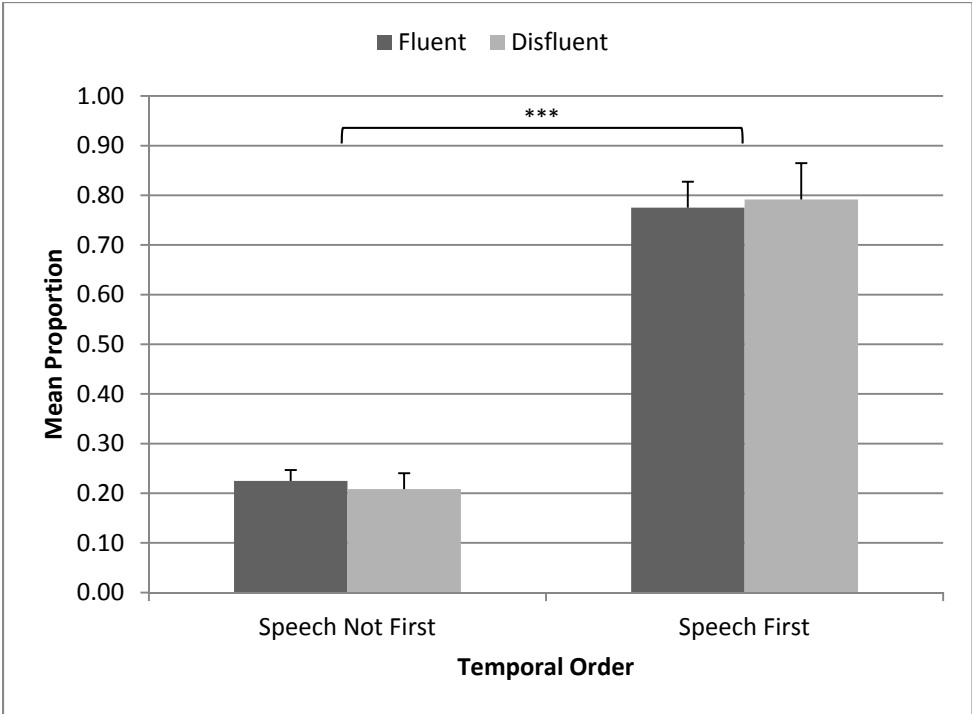
Temporal order as a function of speech fluency for CWNS



Note. Bars represent standard error of the mean

Figure 2

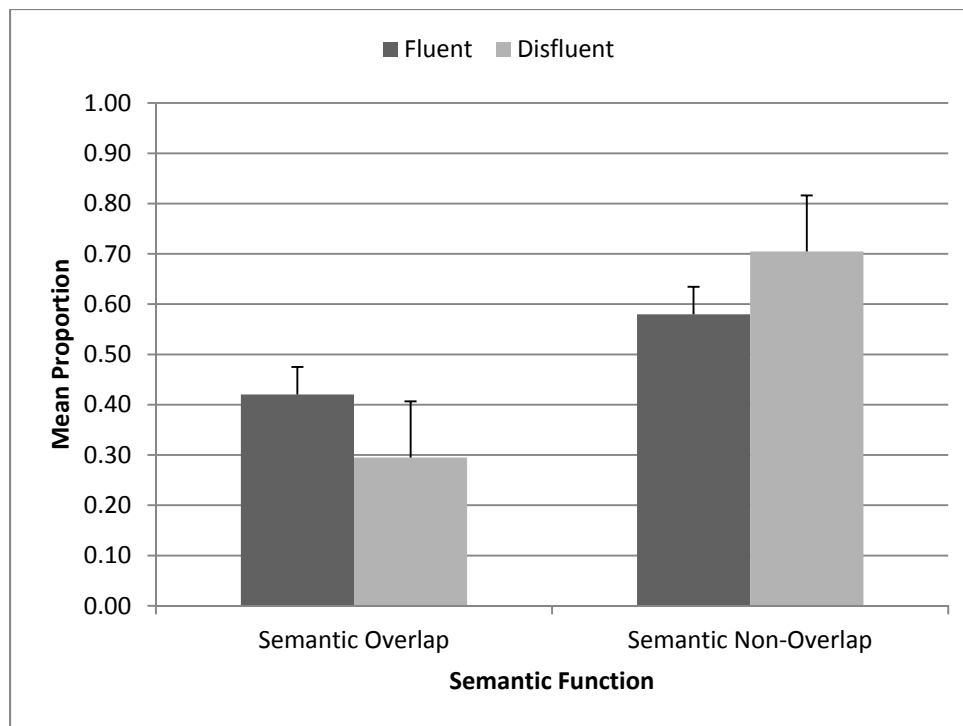
Temporal order as function of speech fluency for CWS



Note. Bars represent standard error of the mean
*** $p < .005$

Figure 3

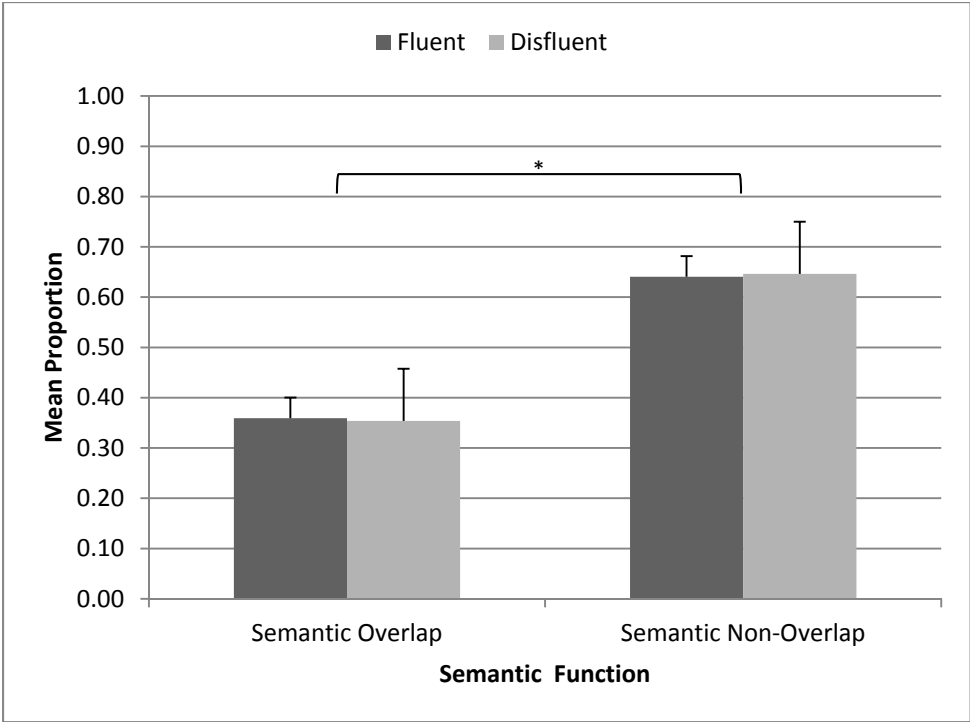
Semantic overlap as a function of speech fluency for CWS



Note. Bars represent standard error of the mean

Figure 4

Semantic overlap as a function of speech fluency for CWNS



Note. Bars represent standard error of the mean
* $p < .05$