First Language

Associations among play, gesture and early spoken language acquisition

Suzanne Hall, Lisa Rumney, Judith Holler and Evan Kidd

The online version of this article can be found at:

Published by:

§SAGE

Additional services and information for *First Language* can be found at:

Email Alerts: http://fla.sagepub.com/cgi/alerts

Subscriptions: http://fla.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations: http://fla.sagepub.com/content/33/3/294.refs.html

>> Version of Record - Jun 19, 2013

What is This?

Associations among play, gesture and early spoken language acquisition

Suzanne Hall
La Trobe University, Australia

Lisa Rumney
La Trobe University, Australia

Judith Holler
Max Planck Institute for Psycholinguistics, The Netherlands

Evan Kidd
The Australian National University, Australia

Abstract
The present study investigated the developmental interrelationships between play, gesture use and spoken language development in children aged 18–31 months. The children completed two tasks: (i) a structured measure of pretend (or 'symbolic') play and (ii) a measure of vocabulary knowledge in which children have been shown to gesture. Additionally, their productive spoken language knowledge was measured via parental report. The results indicated that symbolic play is positively associated with children's gesture use, which in turn is positively associated with spoken language knowledge over and above the influence of age. The tripartite relationship between gesture, play and language development is discussed with reference to current developmental theory.

Keywords
Development, gesture, play, spoken language

Corresponding author:
Evan Kidd, Research School of Psychology (Building 39), The Australian National University, Canberra, ACT 0200, Australia.
Email: evan.kidd@anu.edu.au

The developmental relationships between (i) play and spoken language and (ii) gesture and spoken language have been well established over decades of research on child language acquisition. Both play and gesture have been shown to positively predict the acquisition of spoken language. For instance, McCune-Nicolich (1981) and McCune (1995) showed that complexity in play is associated with subsequent analogous development in language production. Similarly, early work on gesture and language by Bates and colleagues showed that deictic gestures predict the onset of verbal language (Bates, 1976; Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979; Bates, Camaioni, & Volterra, 1975). The tightly coupled relationship between gesture and language continues once children begin to speak: in their second year of life children begin to produce iconic gestures that subsequently predict the acquisition of spoken words and multi-word speech (e.g. Acredolo & Goodwyn, 1988; Capirci, Contaldo, Caselli, & Volterra, 2005; Capirci, Iverson, Pizzuto, & Volterra, 1996; Iverson & Goldin-Meadow, 2005; Özçalişkan & Goldin-Meadow, 2005; for a review see Capone & McGregor, 2004). Although it has been studied less, a link has been identified between play and gesture (e.g. Acredolo & Goodwyn, 1985; Bates et al., 1979; Namy, Acredolo, & Goodwyn, 2000; Namy, Vallas, & Knight-Schwarz, 2008), suggesting that these three systems are interrelated in development.

Play (particularly pretend, or 'symbolic' play), gesture and language simultaneously highlight the symbolic and social nature of our species. In the past their interrelationships have been taken as evidence of a *semiotic function*: a foundational cognitive ability that gives rise to fundamental symbolic skills that at least initially proceed to develop in a domain-general manner (Piaget, 1962; Werner & Kaplan, 1963). Beyond the claim that these behaviours are reflections of a common underlying representational system, there have been few comprehensive developmental accounts of their interrelationships. From a socio-pragmatic perspective, Nelson (2007) suggests that play and gesture allow the child to externalise mental representations shared between child and caregiver in joint activity, which helps to develop symbolic understanding, a crucial prerequisite for successful language acquisition. In the current study we provide additional empirical evidence for the existence of this tripartite developmental relationship between the three systems.

Empirical demonstrations of the links between play, gesture and language

Numerous studies have identified a closely coupled relationship between play and early spoken language development (e.g. Bornstein, Vibbert, Tal, & O'Donnell, 1992; Kelly & Dale, 1989; Lewis, Boucher, Lupton, & Watson, 2000; McCune-Nicolich, 1981; Ogura, 1991; Shore, O'Connell, & Bates, 1984). In the most comprehensive study of its type to date, McCune (1995) reported on a large cross-sectional and a smaller longitudinal study that investigated the interrelationships between play and language in children aged 8–24 months. Following previous studies, McCune showed that development in play coincided or was closely followed by developments in language in both the cross-sectional and longitudinal samples. For a majority of the sample, children's play developed with increasing complexity, from pre-symbolic play schemes through to symbolic combinatorial and

hierarchical play. These developments were associated with milestones in early language development, from the production of first words to the production of multi-word speech.

Similarly, there are numerous demonstrations of the predictive developmental relationship between gesture and language (e.g. Acredolo & Goodwyn, 1988; Bates et al., 1979; Capirci et al., 1996, 2005; Iverson, Capirci, & Caselli, 1994; O'Reilly, Painter, & Bornstein, 1997; Özçalişkan & Goldin-Meadow, 2005; Pizzuto & Capobianco, 2005; Rowe & Goldin-Meadow, 2009; Rowe, Özçalişkan, & Goldin-Meadow, 2008). Most research has concentrated on early language development in infants, where it has been shown that gesture use precedes and predicts the onset of spoken language. Although the uses and functions of gestures may change throughout development, children do not abandon the use of gesture once they begin to talk. For instance, Mayberry and Nicoladis (2000) showed that gesture frequency increases with increasing utterance length in children aged 2–3.5 years, and that their use of iconic gestures preceded their use of spoken labels for concepts (see also Kidd & Holler, 2009). These studies support the idea that gesture and language form an integrated communicative system, and that speech and co-speech gesture *together* constitute human language (Bavelas & Chovil, 2000; Clark, 1996; Kendon, 2000, 2004; McNeill, 1992).

Play may also provide an incidental learning environment whereby children learn gestural symbols. Acredolo and Goodwyn (1985) suggested that children derive many of their first symbolic gestures from action sequences observed in caregiver–child play. In a longitudinal study of three children, Capirci et al. (2005) showed that children's iconic gestures and/or words were commonly first expressed as action schemas. For example, the (play-based) action of bringing an empty spoon to the child's lips became an iconic gesture, where the child raised an empty hand to the lips to signify eating. This was later replaced by a referential word for the action. Children appear to acquire these schemas in object-focused activities (see also Volterra & Erting, 1994). Namy et al. (2000) showed that, whereas there are few empty-handed symbolic gestures produced by caregivers in their interactions with children aged 12–15 months, they produced many 'in-hand gestural labels', which are object-specific play routines with objects in hand, such as moving a toy car. Namy et al. interpreted these data to suggest that the early equal weighting placed on both spoken language and gesture in interaction guides children to treat words and gestures as equivalent forms of symbolic reference (see also Namy & Waxman, 1998, but see Marentette & Nicoladis, 2011; Puccini & Liszkowski, 2012). That is to say, since the social dynamic affords the use of verbal and non-verbal behaviours, children may use both sources of input to extract information about the symbolic nature of the world. Building on these data, Namy et al. (2008) showed that caregivers' in-hand gestures during pretend play predicted their children's gesture vocabulary as measured by parental report, although only caregivers' out-of-hand gestures predicted children's gesture production in free play.

In contrast to the wealth of literature that has identified coupled developmental relationships between play, gesture and language, there has been comparatively little research that has investigated the tripartite relationship between the three systems. Bates et al. (1979) followed 25 children longitudinally between the ages of 9 and 13 months. The children completed a range of tasks that assessed play, gesture, spoken language and other general abilities (e.g. imitation, motor development). A detailed

discussion of this large study is well beyond the scope of this review; however, there are two results that are particularly relevant to the current study. First, play predicted gesture better than it predicted language. In particular, symbolic and combinatorial play best predicted gesture development. These play measures were also associated with language, but to a lesser extent. Second, gesture predicted language development. In particular, communicative pointing best predicted language development at this early stage, in addition to other gesture types such as giving, reporting and ritual requests. Importantly, non-communicative pointing was also positively associated with language development. The dual influence of communicative and non-communicative pointing was argued to reflect two different developmental processes. The association between communicative pointing and language was argued to reflect the fact that communicative pointing is a gestural, sensorimotor form of naming, which functions to share reference with an interlocutor. In contrast, the association between non-communicative pointing and language was explained with reference to Werner and Kaplan (1963), who argued that 'pointing for self' is related to the development of 'objects of contemplation'. That is, Werner and Kaplan suggested that non-communicative pointing facilitates children's understanding of reference by helping them to distinguish between 'the knower and the known'; that is, themselves and the object of interest. Finally, the children's use of gesture was found to continue as they acquired more spoken language, suggesting that gestures are not replaced by spoken language in early language development. Rather, gesture, and in particular pointing, is used alongside spoken language as children enter their second year and beyond.

The Bates et al. (1979) study showed that children's play behaviour predicts their gesture use, which subsequently predicted early language development. This suggests a common underlying mechanism that supports development in these three domains. Piagetian theorists have interpreted these commonalities to reflect the existence of the semiotic function or sets of interrelated developmental schemata (*local homologies*), but such explanations privilege internal cognitive structure over social interaction. Recent developments in socio-pragmatic theory suggest fundamentally social foundations to development in play, gesture and language. Based on observations that each skill develops out of interaction with a competent other, the Cultural Learning approach to symbolic development argues that all three skills are emergent developments from the uniquely human skills of cooperation and shared intentionality (Tomasello, Carpenter, Call, Behne, & Moll, 2005; Tomasello, Carpenter, & Liszkowski, 2007; Tomasello & Rakoczy, 2003).

Rakoczy (2006, 2008) has argued persuasively for a rich, social interpretation of play. He argues that play constitutes the first unambiguous forms of shared intentionality in ontogeny; that through social interaction infants engage in joint attention, shared action and imitative cultural learning. Through pretend play they gain the added benefit of learning about representational relationships (i.e. X stands for Y in context C). This capacity for shared intentionality leads to a motivation to share experiences with social partners, which preverbally includes manual gestures such as deictic and iconic gestures, subsequently leading to the emergence of spoken language. Although this is largely consistent with the results of Bates et al. (1979), those data are in need of corroboration. The data were drawn from observations of parent–child interaction and from parent-report.

Although the two sources of data were largely consistent, the use of naturalistic observational data raises the possibility that children's symbolic skills were overestimated due to parental scaffolding and child imitation. Furthermore, their sample size was small ($N =$ 25), raising the possibility that some of the correlations they reported may have been inflated.

The current study

The current study revisited the tripartite relationship between play, gesture and spoken language in a cross-sectional sample of children aged 18–31 months. Whereas most previous studies have relied on naturalistic observation in parent–child interaction to measure play, gesture and language, the current study made use of independent measures of each domain. This enabled us to minimise the possibility that children's performance was supported by differential amounts of parental support via scaffolding. Following the developmental patterns reviewed above, it was hypothesised that all three systems would be associated with each other. Following Bates et al. (1979), it was hypothesised that play would be more strongly associated with gestural development than with spoken language development. It was also hypothesised that gesture use would be a better predictor of spoken language knowledge than would play.

Method

Participants

Fifty ($N = 50$) typically developing children (17 males, 33 females) aged 18–31 months ($M = 24.52$ months, $SD = 3.48$) were recruited through personal contacts and mother's groups across metropolitan Melbourne, Australia. Children of this age range were recruited so that we could capture a broad range of abilities across a 12-month age-span. Since we were testing the children on tests designed to elicit play, gesture and language, we set the bottom end of the age range at 18 months. This is the age at which a majority of children are typically able to engage in some form of symbolic play (Nielsen & Dissanayake, 2004). Additionally, at this age their frequency of gesture use and vocabulary knowledge is increasing (Fenson et al., 2007; Rowe et al., 2008). We do not claim to capture the emergence of these behaviours; rather, we are investigating the developmental relationships between these domains at a point in infancy where we should find large individual differences.

The children had no pre-existing language or cognitive impairments and were acquiring English as their native language. The majority of participants were born to parents from a middle-class background.

Materials

The children were tested on two measures: (i) the Test of Pretend Play (Lewis & Boucher, 1997) and (ii) the *Parole in Gioco* ('Word Games') vocabulary inventory (Bello, Caselli, Pettenati, & Stefanini, 2010), which was also used to elicit gestures. Finally, their spoken language knowledge was measured using the MacArthur Communicative Development

Inventory (Words and Sentences, Fenson et al., 2007), a parental report questionnaire. Each is described in turn.

The Test of Pretend Play. The Test of Pretend Play (ToPP) (Lewis & Boucher, 1997) is a standardised measure of pretend play for children aged 1–6 years. The test contains a standard set of toys, including everyday household objects (a bowl and spoon), representational toys (a teddy and a doll) and non-representational materials (a yellow top, red cloth, white counter, black box and a brown stick). Children are asked to copy the researcher's play or invent their own original play. The non-verbal version of the ToPP was used, as is recommended for children under 3 years. The ToPP has four sections, which increase in their representational complexity as the test progresses. Section 1 introduces the child to the task. Sections 2–4 measure three types of play. Section 2 assesses children's *object substitutions* using non-representational materials (e.g. using a disk as a doll's hat). Section 3 assesses *representational alone play*: children's ability to attribute imagined objects and properties in play scenes (e.g. pretending a teddy is driving a car or feels sad, without the aid of toy props) and Section 4 measures *self-alone play* without the aid of any toys (e.g. pretending to be a tree or eating an ice-cream). Sections 2–4 measure partially independent constructs (in the current sample all rs < .4). Therefore, although children receive an overall global score based on their performance on all four sections, we analysed the contribution of each type of play to gesture and language separately in our analyses using the raw scores for each section.

The ToPP has good concurrent validity with another test of play – the Symbolic Play Test (Lowe & Costello, 1988, $r = .60$). The test-retest reliability is .87. Lewis and Boucher (1997) reported internal consistencies across sections and age groups ranging from .55 to .94; for the current sample the internal consistency for each section was fairly stable. although not particularly high (Section 2: $\alpha = .61$, Section 3: $\alpha = .64$, Section 4: $\alpha = .64$).

Word Games Vocabulary Task. The 'Word Games' Vocabulary Task (Bello et al., 2010) measures young children's spoken language comprehension and production, as well as their use of gesture. The task was chosen because it has been shown to elicit many gestures from young children without any explicit elicitation or modelling (Stefanini, Bello, Caselli, Iverson, & Volterra, 2009; Stefanini, Caselli, & Volterra, 2007). The task consists of 88 coloured pictures which are divided into two subtests: (a) a nouns subtest which includes 44 pictures of different categories (body parts, animals, objects/tools, food and clothing); and (b) a predicates subtest which includes 44 pictures of actions and other predicate relationships that can be described by the use of adjectives or prepositions. Each subtest has a comprehension and production component. The task was translated into English by the final author. An initial inspection of the pictures used to elicit spoken and gestural responses showed that no concepts were culturally specific to Italy, and were therefore deemed appropriate to use with Australian children.

The subtests of the Word Games task show reasonable concurrent validity when compared to other vocabulary measures, with correlations ranging between .32–.53 and .28–.30. Data on other types of reliability and validity are not available, as only an Italian-language version of the task is currently normed.

The MacArthur–Bates Communicative Development Inventory. The MacArthur–Bates Communicative Development Inventory (MB-CDI) (Fenson et al., 2007) is a widely used measure of children's early language knowledge. The words and sentences form, designed for use with 16- to 30-month-olds, consists of two scales: words and sentence complexity. Part A of the form requires the parent/caregiver to indicate whether the child is able to say each item from a list of words which includes nouns, verbs, adjectives, pronouns, prepositions, quantifiers, articles and connectors. Part B assesses several aspects of morphology and syntax, including use of plurals (*-s*), possessives (*-'s*), progressive (*-ing*) and past tense (*-ed*), as well as word combinations. The version used in the present study had been amended slightly to reflect Australian English, with permission from the authors and publisher, and did not affect the reliability of original instrument (Edith L. Bavin, personal communication).The MB-CDI has high internal consistency with an alpha value of .96 and high reliability for test-retest correlations (with an average test-retest gap of 1.38 months) for word production of .95 with correlations above .90 at each age. It shows good concurrent validity when compared to laboratory-based assessments, with correlations ranging between .53 and .73.

Procedure

Each child was tested either in their home over one or two sessions, with breaks given within each session as needed. Each session began with 10–20 minutes of free play and/or book reading with the child in order to relax and familiarise them with the researcher. The order of the tests was counterbalanced, with some children receiving the ToPP and the others receiving the Word Games task first. The children sat at a child-sized table and chair with the researcher sitting beside them. Parents were present during testing. A second researcher coded the tests live; however, each session was also video recorded to check ambiguous responses and recode both tests for inter-rater reliability.

The procedure for both the ToPP and the Word Games task followed the standardised instructions for the tests. Before testing began in the ToPP, the children were engaged in a warm-up session that served to familiarise them with the experimenters and the toys. Each child was then guided through a series of structured play scenarios designed to elicit different forms of pretend play. In each scenario the child was presented with specific toys relevant to that scenario and asked to play either by imitating the researcher (modelled play) or by creating their own play in response to specific requests such as '*what else can you do?'* or *'can you do something different?'* (elicited play). Following the instructions for the non-verbal version of the test, the researchers used gestures, pointing, looking, eye contact, touch, single words or short phrases to convey to the child what was expected of them. Children were coded according to their ability to model the researcher and the complexity of their symbolic play. Twelve percent (6/50 children) of the data were re-coded for reliability. Reliability was high (Cohen's kappa = .81).[1] Disagreements were adjudicated by the final author.

All children were tested on the nouns subtest of the Word Games task first, followed by the predicates subtest. The pictures were presented in sets of three, with each set appearing in the same order each time. However, the layout of each set of three was randomised for each child. Each set consisted of a picture corresponding to a comprehension

question, a production question and a distracter. For the comprehension component the child was asked to indicate the correct picture, for example, *'where is the cat?'* (nouns subtest) or *'who is singing?'* (predicates subtest). The child was given one opportunity to answer this question, and responses were recorded when the child pointed to, picked up, showed or otherwise indicated the correct picture. Next the researcher took the comprehension and distracter pictures away, leaving the picture corresponding to the production question. The researcher then attempted to elicit either a verbal and/or gestural description of the picture by asking either *'what's this?'* (nouns subtest) or *'what is he doing?'* (predicates subtest). Children were given two opportunities to answer this question. The testing session was video recorded for later coding.

The task was coded as follows. Spoken answers were classified as correct, incorrect or 'no response'. Responses were marked as correct when the child provided an accurate label for the picture. Mispronunciations were accepted. For some pictures more than one answer was accepted if both were common references to that picture or if the answer given was age appropriate. For example, 'drawing' was accepted as correct for a picture of a boy writing, since from a 2-year-old's perspective the picture could represent both activities. Responses were marked incorrect if the child answered with a word that was different from the word that was supposed to be elicited by the picture. Items were marked with no response if the child did not answer or informed the researcher that they did not know the answer.

All gestures that were produced as children interacted with the researcher were coded. Gestures produced with or without speech and before or after the spoken answer was provided were counted. Only manual gestures, body movements and movements of the head were coded. Non-manual gestures (e.g. facial expression, eye gaze and posture) were not included, and nor were conventionalised/emblematic gestures (e.g. holding up both hands with palms facing up to signify 'I don't know'), since these have not been shown to have the same relationship with language acquisition as have other gesture types (i.e. deictics and iconics). Gestures were coded according to the following two categories:

Deictic gestures. This category included showing, giving and pointing. Pointing includes extension of the index finger directed to (touching or patting) the picture, pointing with other fingers or with the palm extended. It is important to point out that, since the children were engaged in a joint activity with an experimenter, we considered all pointing in this context communicative.

Iconic gestures. This category included any concrete imagistic representation of the picture's meaning, where the child gestured the action usually performed with the object depicted (e.g. performing the action of combing hair by splaying the fingers and moving them through their hair in response to a picture of a comb) or when they gestured in imitation of the object (e.g. the child put their arm up in the air for an elephant's trunk).

Each token of a gesture was counted if it represented a unique non-verbal communication during a specific item. For instance, if a child produced both a deictic and iconic gesture on the same item then both were counted. However, if a child pointed twice to the same picture it was only counted as one deictic gesture (repeated iconic gestures were

Table 1. Descriptive statistics for the ToPP and the MB-CDI.

	Mean (SD)	Range	Skewness	Kurtosis
ToPP[a]	9.72 (3.7)	[3, 19]	.56	−.22
MB-CDI[b]	288.1 (190.1)	[16, 654]	.32	−1.07

[a]Max score = 34.
[b]Max score = 826.

also only counted as one gestural token). This was because the inclusion of repetitive gestures may have artificially inflated or obscured any association between spontaneous gesture use and play and spoken language. Finally, although they were included in our initial coding scheme, no child produced any beat gestures. This is in line with McNeill (1992), who argued that beat and metaphoric gestures emerge much later, typically between the ages of 5 and 11 years.

Inter-observer reliabilities were calculated on 10% of the data (5/50 children). Reliability was high: the two observers agreed on 95.8% (183/191) of the gestures (Cohen's kappa = .79, indicating 'substantial agreement' (Landis & Koch, 1977). All disagreements were adjudicated by the final author.

The Word Games task was administered according to the test instructions, which require the collection of comprehension and production data. However, only performance on the production component of the task was included in the subsequent analyses. This is because the Word Games task was used as a means with which to elicit spontaneous gestures from the children. In the comprehension component children must gesture to indicate their response; that is, they must point to the picture that they believe matches the experimenter's spoken utterance. As such, the task requirements are confounded with any measure of gesture that could be calculated on the comprehension component of the task. Unlike the comprehension component, the production component does not require children to gesture at all; therefore their use of gestures in this component is revealing about their tendency to gesture in general. Children's answers on the production component of the test were coded according to modality: (a) unimodal spoken production – answers were spoken only, (b) unimodal gestural productions – answers were gestured only (irrespective of gesture type), (c) bimodal production – answers were gestured and spoken (irrespective of gesture type).

Results

We first present the descriptive statistics for each measure. Table 1 presents the descriptive statistics for the MB-CDI and the ToPP.

Table 1 shows a broad distribution of scores on the ToPP and the MB-CDI. Both measures were normally distributed. We next present the descriptives for the Word Games task. Since we are only interested in their production, we only present the results from the production component of the task. The overall mean was 12.86 (*SD* = 8.6, *range*: [0, 34]).[2] Figure 1 shows the distribution of the response types across three categories: unimodal spoken responses, unimodal gestural responses and bimodal gesture + speech responses.
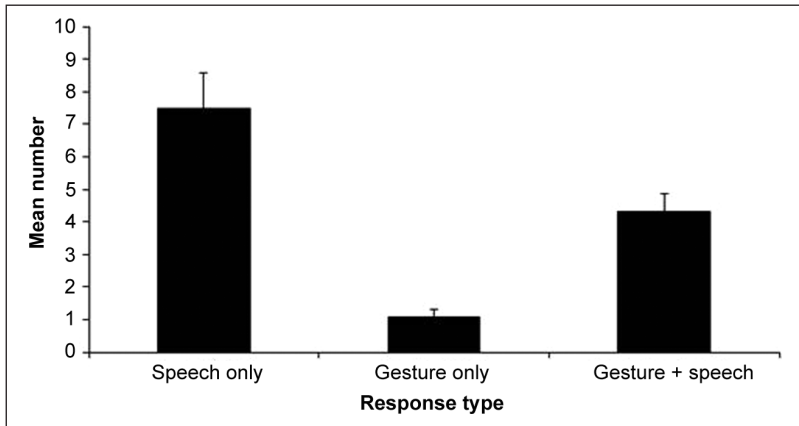
Figure 1. Distribution of response types on the production component of the Word Games task.

Figure 1 shows that, while unimodal spoken responses were most common, bimodal gesture + speech responses were also common. In contrast, unimodal gesture only responses were relatively rare. Of the responses that contained gestures ($N = 283$, $M = 5.66$ per child), 78.8% ($N = 223$) were deictic gestures, whereas 21.2% ($N = 60$) were iconic.[3] Within the category of deictic gesture, 96.4% were pointing gestures. The preponderance of deictic gestures in the context of the type of naming task we used is consistent with Stefanini et al. (2009).

We now consider the interrelationships between the children's play, gesture and spoken language. Table 2 presents the simple bivariate correlations between children's scores on the three sections of the ToPP, children's gesture, and total score on the MB-CDI. Two gesture variables were used in the subsequent analyses: (i) bimodal gesture + speech combinations (irrespective of type) and (ii) gesture use overall (i.e. bimodal gesture + speech combinations as well as unimodal gestural utterances, irrespective of gesture type).

Table 2 shows that many of the measures were interrelated. Across domains, *representational toy alone* play was strongly associated with both bimodal gesture + speech combinations and gesture use overall, whereas the other measures of play were not. Bimodal gesture + speech combinations were strongly associated with MB-CDI total score, whereas the association between total gesture use and MB-CDI total score was lower yet still significant.

We next ran two multiple linear regression analyses that aimed to test our hypotheses. The first tested the independent contributions of the play measures and performance on the MB-CDI on gesture use in gesture + speech combinations using hierarchical linear regression.[4] Age (in months) was entered in Block 1, followed by the three play measures and the MB-CDI total measure in Block 2. The results are presented in Table 3.

The model was significant at Block 1 [$F(1, 48) = 18.43$, $p < .001$], with age significantly and positively predicting bimodal gesture + speech combinations. The model was significant at Block 2 [$F(5, 44) = 9.57$, $p < .001$]; the addition of the play measures and

Table 2. Simple bivariate correlations between measures of play, gesture and language.

	Object substitution[a]	Representational toy alone[b]	Self-alone[c]	Bimodal speech + gesture	Total gesture	MB-CDI total
Age	.40**	.48***	.19	.53***	.31*	.75**
Object substitution[a]		.25#	−.11	.21	.07	.34*
Representational toy alone[b]			.33*	.58***	.47***	.29*
Self-alone[c]				.07	.05	.25#
Bimodal speech + gesture					.89***	.52***
Total gesture						.33*

[a]ToPP Section 2.
[b]ToPP Section 3.
[c]ToPP Section 4.
#p < .10, *p < .05, **p < .01, ***p < .001 (all p-values two-tailed).

Table 3. Summary of multiple regression analysis for bimodal gesture + speech combinations ($N = 50$).

Variable	B	SE	β	t	Sig. (p)
Block 1					
Constant	−10.31	3.44		−2.96	.004
Age	.58	.14	.53	4.29	< .001
Block 2					
Constant	−.13	4.03		−.03	.97
Age	.02	.20	.02	.11	.91
Object sub.	−.25	.28	−.11	−.91	.37
Rep. toy alone	1.14	.27	.55	4.31	< .001
Self-alone	−.47	.24	−.23	−1.95	.06
MB-CDI total	.01	.003	.44	2.70	.01

Note: Block 1 R^2 = .28, Block 2 R^2 = .52.

the MB-CDI total resulting in a significant increase in variance explained [$F(4, 44) = 5.59$, $p = .001$]. After the addition of the variables in Block 2 age no longer predicted gesture + speech combinations. In contrast, *representational toy alone* play and MB-CDI significantly and positively predicted gesture + speech combinations. *Representational toy alone* play uniquely explained 20.3% in gesture + speech combinations; MB-CDI-total explained 8%.

Our next regression tested the contributions of play and gesture to spoken language knowledge, as measured by MB-CDI total. That is, we aimed to predict variance in linguistic knowledge, as measured by the MB-CDI, using our measures of play and gesture use. In this analysis we removed *self-alone* play as a predictor, since it was only weakly associated with MB-CDI scores and was not associated with either gesture or age. *Self-alone* play was replaced with the children's spoken score of the Word

Table 4. Summary of multiple regression analysis for MB-CDI total (N = 50).

Variable	B	SE	β	t	Sig. (p)
Block 1 Constant	−452.92	115.66		−3.92	< .001
Age	26.66	5.03	.49	5.3	< .001
Word Games spoken	11.66	2.25	.48	5.17	< .001
Block 2 Constant	−287.4	119.45		−2.41	.02
Age	17.47	5.84	.32	2.99	.005
Word Games spoken	12.98	2.15	.53	6.04	< .001
Object sub.	.55	8.52	.005	.06	.95
Rep. toy alone	−9.09	9.04	−.01	−1.01	.32
Gesture + speech	16.38	4.38	.34	3.74	.001

Note: Block 1 R^2 = .73, Block 2 R^2 = .79.

Games task. This variable acted as a control: since the Word Games task is also a measure of vocabulary knowledge, we had to mitigate against the possibility that any association between gesture and spoken language was due to children's spoken performance on the task. A hierarchical multiple linear regression was conducted. Age and Word Games spoken production score were entered in at Block 1, followed by the two remaining play measures and gesture + speech combinations at Block 2.[5] The results are presented in Table 4.

The model was significant at Block 1 [$F(2, 47) = 61.99, p < .001$], with age and Word Games spoken production score significantly and positively predicting spoken language knowledge as measured by the MB-CDI. The model was significant at Block 2 [$F(5, 44) = 33.6, p < .001$]; the addition of gesture + speech combinations but not the play measures resulted in a significant increase in variance explained [$F(3, 44) = 4.8, p = .006$]. Gesture + speech combinations explained 6.6% of unique variance in MB-CDI total scores.

Discussion

The current study aimed to investigate the tripartite relationship between play, gesture and spoken language in infants aged 18–31 months. We hypothesised that play would be more strongly associated with gesture than with spoken language, and that gesture use would be more strongly associated with spoken language than would play. *Representational alone* play and spoken language as measured by the MB-CDI were significantly associated with gesture use, and both explained a significant amount of unique variance in gesture use over and above the influence of age. Furthermore, although all forms of play were weakly correlated with spoken language, only gesture use explained a significantly unique portion of variance in spoken language knowledge once all variables were included in a multiple regression analysis. Our hypotheses were therefore supported.

The data are broadly consistent with the strength of associations reported by Bates et al. (1979), who also observed that play was more strongly associated with gesture than

with spoken language, and that gesture was more strongly associated than play with spoken language. Thus our results are consistent with the interpretation that the relationships between the variables reflect the existence of local homologies or sets of interrelated developmental schemata between (i) (symbolic) play and gesture and (ii) gesture and spoken language. How might we narrow down the source of these interrelationships? We suggest that an adequate explanation of these phenomena requires reference to two concepts: (i) action-based schemas and (ii) joint activity.

The relationship between gesture and spoken language has been argued to emerge out of action-based schemas. Stefanini et al. (2009), who used the same gesture elicitation task used in the current study, suggested that the 2-year-old children in their sample gestured when labelling pictures because it recreated a direct action-based link between the object or action to be labelled and the verbal label, arguing that children's lexical knowledge at this age is still not fully decontextualised from their sensorimotor experience. Other empirical evidence for the emergence of gesture from action-based schemas has considered the emergence of iconic gestures, where representational action sequences are used to represent concepts the children have yet to label (e.g. Acredolo & Goodwyn, 1985; Capirci et al., 2005; Namy et al., 2000, 2008; for evidence from comprehension see Marentette & Nicoladis, 2011). It is likely that children use pre-linguistic sensorimotor conceptual information to break into language (Mandler, 1992). Children's early vocabularies are dominated by words that are high in concreteness and imageability, suggesting that early acquisition is guided by physical and perceptual salience (for computational evidence see Howell, Jankowicz, & Becker, 2005; Jansen & Watter, 2012). This proposed link between gesture and action has been argued to explain the emergence of language in ontogeny, and the evolution of the linguistic system itself (for reviews see Bates & Dick, 2002; Corballis, 2010; Gentilucci & Corballis, 2006).

Where does play fit into this explanation? It seems that at least some gestures emerge out of play-based joint activity (e.g. Capirci et al., 2005; Namy et al., 2000, 2008), providing a direct link between play-based action, gesture and spoken language. However, we suggest that the activity of play is also likely to provide the contextual frame that provides the basis for communicative development. The Cultural Learning approach to symbolic development posits that play and gesture should be related because they both reflect infants' emerging ability to engage in collective or shared intentionality, which is nurtured in joint play contexts (e.g. Rakoczy, 2006, 2008; Tomasello & Rakoczy, 2003). On this explanation play has two complementary functions: it is the context within which infants come to understand that others have intentions and desires (i.e. mind states) upon which they act, and it is also the catalyst for the insight that we live in a symbolic world (e.g. We can pretend Teddy is driving a car even if there is no car present). That is to say, play provides one major context in which the child can come to understand that human social interaction results in the construction of a shared and meaningful reality. This capacity provides the foundation for the mutual exchange of meaning in communicative contexts, which is initially dominated by gesture (Nelson, 2007). The use of gesture is later augmented but not replaced by spoken language.

The action-based and socio-cognitive explanations of these developmental associations are likely to be complementary. The action-based explanation provides a cognitive explanation of the emergence of symbolic function that is grounded in action. The

cultural learning explanation situates this cognitive process within the rich social context of human interaction, arguing that these behaviours necessarily emerge out of joint activity. The social context of development is often obscured in studies of play, gesture and language, particularly in experimental contexts which demand tight controls. However, the majority of play and communicative interaction is social. For instance, Haight and Miller (1993) showed that children's early play occurs with their primary caregiver, and later occurs with a peer or sibling. Similarly, communicative interaction, be it achieved through gesture or spoken language (most likely both), is a fundamentally interactive process.

A number of specific results deserve further discussion. In the current data only symbolic play that involved the mental projection or transformation of objects, states and actions was significantly associated with gesture use (i.e. Section 3 of the ToPP – *representational alone* play). This type of play is truly symbolic, and, by virtue of its significant association with children's gesture use, is indicative of the symbolic nature of children's gestures. It is unclear whether the relationship between gesture and play only holds for this variety of play, or whether the effect that we observed reflects the age range that we tested. Other types of play (e.g. combinatorial play, object substitution) have been reported to be related to language acquisition at different stages in development (e.g. Bates et al., 1979; McCune, 1995), and so it is possible that different types of play are also related to gesture development, but that our study did not capture these relationships because they occur outside the age range tested. This is a matter for future research.

It is also important to note that young infants' use of iconic gestures is rather infrequent in comparison to both conventional and deictic gestures, which are by far the most frequent types (e.g. Özçalişkan & Goldin-Meadow, 2005; Pizzuto & Capobianco, 2005). This was also evident in the current study, where over 75% of the children's gestures were deictics (mostly pointing). Deictic gestures play an important role in language acquisition. Pizzuto and Capobianco (2005) followed six Italian-speaking children longitudinally from 12 to 24 months. They showed that at 12 months single deictic gestures (pointing, showing, requesting) were as common as single referential word use. They also showed that the most frequent form of two-element utterances was deictic gesture + referential word combinations, which increased from 12 to 18 months, and stabilised thereafter. They interpreted their results to suggest that, for young children, naming is generally carried out across modalities, typically with deictic gesture + speech combinations. A recent meta-analysis by Colonnesi, Stams, Koster, and Noom (2010) supports this assertion. They reviewed 25 studies that investigated the relationship between pointing and spoken language acquisition, reporting significant concurrent ($r = .52$) and longitudinal ($r = .35$) relationships between the two. This relationship held only for declarative but not imperative pointing, and became stronger with age.

There is some debate as to whether children's early pointing is social or non-social (see Carpendale & Carpendale, 2010; Tomasello et al., 2007). Regardless of its origins, pointing appears to have a clear socio-cognitive basis in a majority of typically developing infants by the beginning of the second year. Liszkowski and Tomasello (2011) reported on a correlational study of pointing in 12-month-old infants. They showed that infants' spontaneous pointing was positively associated with their mother's spontaneous pointing in

joint activity. Additionally, infants who pointed with their index finger vocalised more often when pointing than did infants who pointed with an open hand, and better comprehended the function of pointing (i.e. to direct another's attentional state). These data suggest that pointing is associated with more advanced socio-cognitive knowledge; in particular, an underlying ability to engage in shared intentionality. Children's tendency to vocalise when pointing suggests that, even during pointing, their vocal and gestural systems are tightly coupled, supporting the suggestion by Capirci et al. (2005) and Stefanini et al. (2009) that the two systems might be linked via the representational properties of the motor system. Future research is needed to further investigate this issue.

The current data suggest numerous avenues for future research. First, longitudinal data are needed to strengthen the concurrent associations found in this sample. Second, cross-cultural data could potentially shed light on the topic. Child-rearing practices, attitudes towards play and interactional styles differ across cultures. Comprehensive cross-cultural studies are few and far between yet sorely needed (but see Callaghan et al., 2011). Gesture use varies with joint activity contexts within cultures (Puccini, Hassemer, Salomo, & Liszkowski, 2010), suggesting that different activities in which caregivers and infants engage support the acquisition of different components of language. Finally, Smith and Jones (2011) have recently suggested that symbolic play is related to spoken language acquisition through visual object recognition. They showed that children's recognition of sparse two-dimensional figures was more strongly associated with object substitution in play than was spoken noun vocabulary. Their data suggest that category knowledge, as measured by visual perception tasks, is important for children's symbolic play (e.g. a child must recognise a pencil as resembling a toothbrush in order to substitute one for the other in play). How this result relates to gesture is unclear, although it is likely that gesture and the recognition of sparse visual figures both rely on access to conceptual knowledge of categories.

To conclude, we have presented data that are consistent with early work by Bates et al. (1979), who reported early developmental relationships between (i) play and gesture, and (ii) gesture and spoken language development. Since this seminal study there have been numerous studies that have investigated the relationship between any two of these three variables, but to our knowledge there have been few if any investigations of the tripartite relationship between all three. The consideration of all three variables across development has the potential to provide key insights into the core social and cognitive factors that drive language acquisition. The current study provides the foundation for future research designed to further explore and narrow down the nature of the relationships we have observed.

Notes

1. All disagreements were in Section 4, where children are encouraged to play without the aid of any toy props, something that our young infants found rather difficult, and which can be ambiguous to code.
2. Max score = 44.
3. This ratio of deictic to iconic gestures is very similar to that found in naturalistic observations of monolingual (English-speaking) and bilingual children of a similar age range as the children tested in the present study (Iverson, Capirci, Volterra, & Goldin-Meadow, 2008; Mayberry & Nicoladis, 2000).
4. An identical analysis was conducted using overall gesture use as the dependent measure, and yielded similar results. *Representational toy alone* play significantly predicted gesture use ($\beta = .53$, $p = .001$), but the contribution of the MB-CDI total was marginal ($\beta = .37$, $p = .07$). No other effects were significant.
5. The results did not change when overall gesture use was included instead of gesture + speech combinations.

References

Acredolo, L. P., & Goodwyn, S. W. (1985). Symbolic gesturing in language development: A case study. *Human Development*, *28*, 40–49.

Acredolo, L. P., & Goodwyn, S. W. (1988). Symbolic gesturing in normal infants. *Child Development*, *59*, 450–466.

Bates, E. (1976). *Language and context*. New York, NY: Academic Press.

Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). Cognition and communication from nine to thirteen months: Correlational findings. In E. Bates, L. Benigni, I. Bretherton, L. Camaioni, & V. Volterra (Eds.), *The emergence of symbols: Cognition and communication in infancy* (pp. 69–140). New York, NY: Academic Press.

Bates, E., Camaioni, L., & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly*, *21*, 205–226.

Bates, E., & Dick, F. (2002). Language, gesture, and the developing brain. *Developmental Psychobiology*, *40*, 293–310.

Bavelas, J. B., & Chovil, N. (2000). Visible acts of meaning. An integrated message model of language use in face-to-face dialogue. *Journal of Language and Social Psychology*, *19*, 163–194.

Bello, A., Caselli, M. C., Pettenati, P., & Stefanini, S. (2010). *PinG – Parole in Gioco. Una prova di comprensione e produzione lessicale per la prima infanzia*. Firenze, Italy: Giunti O.S. Organizzazioni Speciali.

Bornstein, M. H., Vibbert, M., Tal, J., & O'Donnell, K. (1992). Toddler language and play in the second year: Stability, covariation and influences of parenting. *First Language*, *12*, 323–338.

Callaghan, T. C., Moll, H., Rakoczy, H., Warneken, F., Liszkowski, U., Behne, T., & Tomasello, M. (2011). Early social cognition in three cultural contexts. *Monographs of the Society for Research in Child Development*, *76*, 1–142.

Capirci, O., Contaldo, A., Caselli, M. C., & Volterra, V. (2005). From action to language through gesture: A longitudinal perspective. *Gesture*, *5*, 155–177.

Capirci, O., Iverson, J. M., Pizzuto, E., & Volterra, V. (1996). Gestures and words during the transition to two-word speech. *Journal of Child Language*, *23*, 645–673.

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.

Carpendale, J. I. M., & Carpendale, A. B. (2010). The development of pointing: From personal directedness to interpersonal direction. *Human Development*, *53*, 110–126.

Clark, H. (1996). *Using language*. Cambridge, UK: Cambridge University Press.

Colonnesi, C., Stams, G. J. J. M., Koster, I., & Noom, M. J. (2010). The relationship between pointing gesture and language: A meta-analysis. *Developmental Review*, *30*, 352–366.

Corballis, M. C. (2010). The gestural origins of language. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*, 2–7.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates communicative development inventories (CDIs): Users' guide and technical manual* (2nd ed.). San Diego, CA: Singular Press.

Gentilucci, M., & Corballis, M. C. (2006). From manual gesture to speech: A gradual transition. *Neuroscience & Biobehavioral Reviews*, *30*, 949–960.

Haight, W. L., & Miller, P. J. (1993). *Pretending at home: Early development in a sociocultural context*. Albany: State University of New York Press.

Howell, S. R., Jankowicz, D., & Becker, S. (2005). A model of grounded language acquisition: Sensorimotor features improve lexical and grammatical learning. *Journal of Memory and Language*, *53*, 258–276.

Iverson, J. M., Capirci, O., & Caselli, M. C. (1994). From communication to language in two modalities. *Cognitive Development*, *9*, 23–43.

Iverson, J. M., Capirci, O., Volterra, V., & Goldin-Meadow, S. (2008). Learning to talk in a gesture-rich world: Early communication in Italian vs. American children. *First Language*, *28*, 164–181.

Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, *16*, 367–371.

Jansen, P. A., & Watter, S. (2012). Strong systematicity through sensorimotor conceptual grounding: An unsupervised, developmental approach to connectionist sentence processing. *Connection Science*, *24*, 25–55.

Kelly, C. A., & Dale, P. S. (1989). Cognitive skills associated with the onset of multiword utterances. *Journal of Speech and Hearing Research*, *32*, 645–656.

Kendon, A. (2000). Language and gesture: Unity or duality. In D. McNeill (Ed.), *Language and gesture: Window into thought and action* (pp. 47–63). Cambridge, UK: Cambridge University Press.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge, UK: Cambridge University Press.

Kidd, E., & Holler, J. (2009). Children's use of gesture to resolve lexical ambiguity. *Developmental Science*, *12*, 903–913.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174.

Lewis, V., & Boucher, J. (1997). *Manual of the test of pretend play*. London, UK: Psychological Corporation.

Lewis, V., Boucher, J., Lupton, L., & Watson, S. (2000). Relationships between symbolic play, functional play, verbal and non-verbal ability in young children. *International Journal of Language and Communication Disorders*, *35*, 117–127.

Liszkowski, U., & Tomasello, M. (2011). Individual differences in social, cognitive, and morphological aspects of pointing. *Cognitive Development*, *26*, 16–29.

Lowe, M., & Cosetllo, A. J. (1988). *Symbolic play test* (2nd ed.). Windsor, UK: NFER-Nelson.

Mandler, J. M. (1992). How to build a baby: II. Conceptual primitives. *Psychological Review*, *99*, 587–604.

Marentette, P., & Nicoladis, E. (2011). Preschoolers' interpretations of gesture: Label or action associate? *Cognition*, *121*, 386–399.

Mayberry, R., & Nicoladis, E. (2000). Gestures reflect language development: Evidence from bilingual children. *Current Directions in Psychological Science*, *9*, 192–196.

McCune, L. (1995). A normative study of representational play at the transition to language. *Developmental Psychology*, *31*, 198–206.

McCune-Nicolich, L. (1981). Toward symbolic functioning: Structure of early pretend games and potential parallels with language. *Child Development*, *53*, 785–797.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.

Namy, L. L., Acredolo, L., & Goodwyn, S. (2000). Verbal labels and gestural routines in parental communication with young children. *Journal of Nonverbal Behavior*, *24*, 63–79.

Namy, L. L., Vallas, R., & Knight-Schwarz, J. (2008). Linking parent input and child receptivity to symbolic gestures. *Gesture*, *8*, 302–324.

Namy, L. L., & Waxman, S. R. (1998). Words and gestures: Infants' interpretations of different forms of symbolic reference. *Child Development*, *69*, 295–308.

Nelson, K. (2007). *Young minds in social worlds: Experience, meaning, and memory*. Cambridge, MA: Harvard University Press.

Nielsen, M., & Dissanayake, C. (2004). Pretend play, mirror self-recognition and imitation: A longitudinal investigation through the second year. *Infant Behavior and Development*, *27*, 342–365.

Ogura, T. (1991). A longitudinal study of the relationship between early language development and play development. *Journal of Child Language*, *18*, 273–294.

O'Reilly, A. W., Painter, K. M., & Bornstein, M. H. (1997). Relations between language and symbolic gesture development in early childhood. *Cognitive Development*, *12*, 185–197.

Özçalişkan, S., & Goldin-Meadow, S. (2005). Gesture is at the cutting edge of early language development. *Cognition*, *96*, B101–B113.

Piaget, J. (1962). *Play, dreams and imitation in childhood*. New York, NY: W. W. Norton.

Pizzuto, E., & Capobianco, M. (2005). The link and differences between deixis and symbols in children's early gestural-vocal system. *Gesture*, *5*, 179–199.

Puccini, D., Hassemer, M., Salomo, D., & Liszkowski, U. (2010). The type of shared activity shapes caregiver and infant communication. *Gesture*, *10*, 279–296.

Puccini, D., & Liszkowski, U. (2012). 15-month-old infants fast map words but not representational gestures of multimodal labels. *Frontiers in Psychology*, *3*, 101–108.

Rakoczy, H. (2006). Pretend play and the development of collective intentionality. *Cognitive Systems Research*, *7*, 113–127.

Rakoczy, H. (2008). Pretence as individual and collective intentionality. *Mind & Language*, *23*, 499–517.

Rowe, M. L., & Goldin-Meadow, S. (2009). Early gesture selectively predicts later language learning. *Developmental Science*, *12*, 182–187.

Rowe, M. L., Özçalişkan, S., & Goldin-Meadow, S. (2008). Learning words by hand: Gesture's role in predicting vocabulary development. *First Language*, *28*, 182–199.

Shore, C., O'Connell, B., & Bates, E. (1984). First sentences in language and symbolic play. *Developmental Psychology*, *20*, 872–880.

Smith, L. B., & Jones, S. S. (2011). Symbolic play connects to language through visual object recognition. *Developmental Science*, *14*, 1142–1149.

Stefanini, S., Bello, A., Caselli, M. C., Iverson, J. M., & Volterra, V. (2009). Co-speech gestures in a naming task: Developmental data. *Language and Cognitive Processes*, *24*, 168–189.

Stefanini, S., Caselli, M. C., & Volterra, V. (2007). Spoken and gestural production in a naming task by young children with Down syndrome. *Brain and Language*, *101*, 208–221.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*, 675–691.

Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development*, *78*, 705–722.

Tomasello, M., & Rakoczy, H. (2003). What makes human cognition unique? From individual to shared to collective intentionality. *Mind and Language*, *18*, 121–147.

Volterra, V., & Erting, C. (1994). *From gesture to language in hearing and deaf children*. Washington, DC: Gallaudet University Press.

Werner, H., & Kaplan, B. (1963). *Symbol formation*. New York, NY: John Wiley.