The onset and mastery of spatial language in children acquiring British Sign Language

Gary Morgan\textsuperscript{a,}* , Rosalind Herman\textsuperscript{a},
Isabelle Barriere\textsuperscript{b}, Bencie Woll\textsuperscript{c}

\textsuperscript{a} City University, London, United Kingdom
\textsuperscript{b} Yeled v Yalda Early Childhood Center, Brooklyn, NY 11218, United States
\textsuperscript{c} DCAL, UCL, London, United Kingdom

Abstract

In the course of language development children must solve arbitrary form-to-meaning mappings, in which semantic components are encoded onto linguistic labels. Because sign languages describe motion and location of entities through iconic movements and placement of the hands in space, child signers may find spatial semantics-to-language mapping easier to learn than child speakers. This hypothesis was tested in two studies: a longitudinal analysis of a native signing child’s use of British Sign Language to describe motion and location events between the ages 1–10 and 3–0, and performance of 18 native signing children between the ages of 3–0 and 4–11 on a motion and location sentence comprehension task. The results from both studies argue against a developmental advantage for sign language learners for the acquisition of motion and location forms. Early forms point towards gesture and embodied actions followed by protracted mastery of the use of signs in representational space. The understanding of relative spatial relations continues to be difficult, despite the iconicity of these forms in the language, beyond 5 years of age.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Language development; Sign language; Gesture; Classifiers

1. Introduction

The majority of children develop language with beguiling ease. Yet learning the correct mappings between meanings and words is a complex problem (Bloom, 2000; Casasola, 2005; Chiat, 2000; Bowerman & Choi, 2003), not least because of the arbitrary relationship between linguistic
form and meaning (Quine, 1960). The animal that lives with people as a pet, runs around the house, barks and chews bones gets called ‘dog’ in English but ‘perro’ in Spanish. Neither sequence of sounds has an obvious link to the concept of a dog (Smith, 2003). In contrast, signs in sign languages often have iconic qualities that display direct visual links to the concepts they represent. For example, the sign CAT in British Sign Language (BSL) makes reference to whiskers at the side of the face. Similarly, for some motion verbs in BSL the handshapes and movement patterns are, to some extent, representative of the physical shapes of the referent objects and their paths of motion. This iconicity is not present in the phonology of the words ‘go down’ or ‘bajar’ in Spanish.

Motion and location events are made up of several semantic elements including path, movement, manner, figure and ground. When children learn how to talk about movement and location they need to learn the specific rules for which formal elements are important in their language (words, parts of speech, grammatical morphemes, construction types) and how they map onto concepts. English-speaking children develop spatial language between ages 2 and 6 years, although mastery may take several years (Johnston, 1984; Kuczaj & Maratsos, 1975; Sowden & Blades, 1996). The same is true for other spoken languages, including German (Grimm, 1975) and the Mexican language Upper Necaxa Totonac (Varela, 2006). English locatives are acquired in a predictable order. First come topological meanings that do not require measurement or perspective (‘in’, ‘on’ and ‘under’); then proximity notions (‘next to’ ‘between’ etc.); and, finally, projective and three-dimensional Euclidean spatial notions are gradually acquired (‘in front of’ and ‘behind’). Within and across spoken languages, there is protracted and consistent order of acquisition of spatial locatives (Bowerman, 1996).

Our aims in this article are twofold: first to describe deaf children’s onset of spontaneous spatial language productions, and second to investigate the eventual mastery of spatial semantics in BSL. Our motivation stems from the iconicity of person and object motion and location descriptions in BSL and what this potential modality difference between speech and sign will mean for the onset, pattern, and rate of development of spatial language.

2. Embodied cognition and gestural communication

Many toddlers show a strong preference for gestures over verbal communication in their spontaneous interactions (Capirci, Iversen, Pizzuto, & Volterra, 1996; Iversen & Goldin-Meadow, 2005; Volterra, Caselli, Capirci, & Pizzuto, 2005). When children start communicating about the location of entities or objects moving, they also use gestures, sometimes several weeks before they say words with the same meanings (Özçalışkan & Goldin-Meadow, 2005). Children learning sign language might also be able to use gesture to talk about motion and location before they have mastered the form-meaning mappings common in the signed language they are being exposed to. One theory of word learning extending from the early Piagetian framework is the embodied cognition approach (Barsalou, 2003; Lakoff, 1987; Piaget & Inhelder, 1956). Embodiment is the dynamic interactive coupling between brain, body and environment. In this framework children develop cognitive representations for movement and location through the building up of embodied sensory-motor/perceptual features related to their own and other objects’ movements and locations (Howell, Jankowicz, & Becker, 2005). A child who wants to express the action of throwing a ball or the movement of an object in space may gesture a throwing action or spread her arms to represent an aeroplane flying. Children’s first words for motion, for example, often emerge from expressions of their own movements (Johnston & Slobin, 1979).
3. Expression of motion and location forms in sign language

When signers talk about space there is an apparent iconicity between what they do with their hands and what the referents they are talking about do in real space. An English sentence such as ‘the pen is on the table’ contains three words that encode the identity of a figure, its location and the identity of a ground. Similarly, the sentence ‘the car drove under the bridge’ encodes the semantic components of figure, ground, location and manner of movement (Talmy, 1985, 2003). BSL and other sign languages have conventions regarding how these semantic components get mapped onto linguistic forms. For example, when describing the figure, there are a set of handshapes that can be used to represent classes of referents with similar forms and meanings. Sign language researchers term these handshapes ‘classifiers’ (see Emmorey, 2003, for more details). In Fig. 1, a signer describes the location of three objects on a table: a cup, a pen and a bunch of keys, each using classifiers. The BSL convention is for the ground referent to be mentioned first and so the sign TABLE is signed in space in front of the signer by moving two flat hands apart at waist height to create a representation of a surface. As each object is mentioned, the noun is articulated first, followed immediately by a corresponding classifier handshape located in the space in front of the signer. The signer uses the following signs: CUP CLASSIFIER-curved hand, PEN CLASSIFIER-extended index finger and KEY CLASSIFIER-spread and bent fingers. Just the classifiers appear in the pictures in Fig. 1.

The form and orientation of the hands may be associated with referents having similar shapes or belonging to a semantic class of objects (e.g., long thin objects). For example, sentences referring to the location of pens, people, poles, upright paint brushes or Big Ben would require the index finger classifier handshape in BSL. Adult signers use many types of asymmetrical two-handed constructions to mark the existence of two objects in space. For example, a figure-ground relationship is articulated through two classifier handshapes used together. The dominant or moving hand typically maps out the figure, while the non-dominant hand represents the ground component. Describing a pen in a cup, a signer would use one hand to represent the ground referent through a curved hand classifier, and the second hand would show how the figure is located inside the cup through an index finger positioned within the confines of the curved hand. The spatial meanings for objects being in figure-ground relations - such as: ‘behind’, ‘under’, ‘in-front’, ‘bottom-left’, ‘inside-right’, ‘top-left’ - are mapped onto the two hands in this way. In sign languages, but not in spoken languages, the position of the hands is akin to the relative locations of the objects in real space.

Spatial meanings therefore appear to be iconic in BSL. However, there are conventions for how each part of a spatial event is mapped out. For example, the signer articulates these handshapes in sign space in the same positions as his viewpoint—with the curved handshape on his left side, the extended index handshape in the middle and the handshape with bent fingers on his right.

![Fig. 1. Sentence expressing relative locations of three objects on a table.](image-url)
When interpreting such sentences, the viewer is required to mentally reverse the spatial array as it is presented from the signer’s perspective and uses three-dimensional spatial relations.

In the preceding example, the utterance CUP-LEFT is actually on the right side of space from the viewer’s perspective. Relational meanings are therefore related to the perspective of the person signing in the same fashion as right-left and deictic expressions in all spoken languages are determined by the speaker (Emmorey & Tversky, 2002).

In the example, shown in Fig. 2, the signer articulates a figure’s movement in relation to a ground referent using classifiers. The signer, describing a scene in which a car passed under a bridge, chooses a flat handshape with fingers together to represent the figure (vehicle classifier) and a flat handshape bent at the knuckles to represent the ground (the bridge), then moves the ‘vehicle’ hand under the ‘bridge’ hand. Each classifier is preceded by a noun for CAR and BRIDGE. Just the classifiers are shown in Fig. 2.

This overview of BSL illustrates that sign languages use language conventions to map out even basic motion and location situations and these are significantly different from the ways that hearing people use gesture. Supalla (1990) and Slobin and Hoiting (1994) discuss movement and manner combinations in serial verb constructions, where a single event is split between two verbs, with one handshape and movement describing the manner of an entity’s movement, followed by a separate path description. The language adopts conventions for how each part of the motion event is mapped onto different articulators in a temporal order.

In a study of the expression of motion events by both child and adult signers of Nicaraguan Sign Language (NSL) and Nicaraguan hearing adult’s use of gestures for the same events, Senghas, Kita and Özyürek (2004) describe NSL signers producing spatial events by splitting the event into separate meaning units, with path and manner in different parts of the sentence. However, when hearing people were tested on the same motion events, rather than separating out the different meaning components, they used holistic gesture forms that conflated motion, manner and path together.

4. The developmental time course for spatial language in signed and spoken languages

Sign languages are learned by deaf and hearing children of deaf parents in ways very similar to those by which spoken languages are acquired by hearing children. First signs appear just before

Fig. 2. Signed sentence ‘the car goes under the bridge’.
12 months; a vocabulary spurt typically occurs at 18 months; two-sign combinations appear at 2–0; the 500-sign stage is reached by 36 months; grammar emerges between 2–0 and 3–0; and discourse functions are acquired in the years leading up to school age (Chamberlain, Morford & Mayberry, 2000; Morgan & Woll, 2002; Schick, Marshark & Woolfe, 2007). Previous research on pronominal and verb sign, suggests that the iconicity of signs does not help deaf children learn language (Meier, 1987; Petitto, 1987).

Slobin et al. (2003) report early use of handshapes and path descriptions in children learning ASL and Sign Language of the Netherlands (SLN). In spoken language research, path expressions emerge very early, even in the one- and two-word speech of children. This is so regardless of whether the Path is expressed as a preposition (as in English) or as a verb (as is more common in Korean) (Bloom, 1973; Choi & Bowerman, 1991). Choi and Bowerman (1991) reported that 14–21-month olds who are learning English produce ‘out’, ‘up’ and ‘down’ to encode their own Paths and ‘on’, ‘in’, and ‘off’ for those of objects. For example, ‘in’ is used to describe movement of self into a shopping trolley or ‘down’ to comment on a doll falling from a sofa.

In sign language examples from Slobin et al. (2003), a deaf child aged 2–8 with non-native SLN input from his hearing mother moves a fist with thumb and pinkie extended in a downward arc to express the notion ‘the plane flies down’. Another Dutch child at 2–6 produced two curved spread fingers handshapes and moved them in an upward, slow, zigzag path to show a ‘balloon drifting away’. A third even younger child, aged 2–1 and learning ASL, copied the mother in producing a two handed construction where the non-dominant hand, acts as a ground (representing a chair) with a relaxed spread fingers handshape and the dominant hand with the index and middle finger touching and extended, was placed on top the non-dominant hand to encode the figure-ground meaning ‘the doll stood on a toy chair’. It is not possible from these small numbers of examples to argue that the children are using productive knowledge of the SLN or ASL classifier system. Evidence that a linguistic form is productive in a child’s language, e.g., the past tense ‘ed’ in English, relies on finding multiple uses by the child of that form across different contexts, e.g., ‘walked’, ‘talked’ and even ‘eated’, rather than isolated examples (Pizzuto & Caselli, 1992). In addition because gesture and sign cannot be clearly separable from each other, as is the case with gesture and speech, the use of criteria for mastery of sign is crucial (Goldin-Meadow, Mylander, & Butcher 1995).

A number of previous studies have reported that sign learners take several years to develop classifiers (De Beuzeville, 2004; Kantor, 1980; Newport, 1990; Newport & Meier, 1985; Schick, 1990; Tang, Sze, & Lam, 2007). In one study of ASL, Newport and Meier (1985) argue that children have difficulty integrating the handshapes necessary to encode the correct figure and ground components while expressing the path or manner through the motion verb. Parts of the event are available to the children but mapping out the whole event correctly is difficult. Before 5 years of age, children map out some figure and simple paths through motion and location verbs. Children aged 5 years and over sometimes break down motion and location forms by sequencing different parts of the event linearly, as opposed to simultaneously, which occurs in the adult language. In one example from Newport and Meier (1985), a child described the CURVED-UPWARD movement of a bird by splitting apart the spatial verb into ARC plus UPWARD. De Beuzeville (2004) reported a similar error in a 5-year-old Australian Sign Language (Auslan) learner who described an event in which a plane spiralled as it flew; the child used the correct thumb and pinkie finger extended from the fist to categorise the figure, but moved the hand in a straight flat line, then stopped the movement, pivoted the hand and finally continued with a straight flat movement. Similar ‘errors’ are reported for children acquiring Hong Kong Sign Language (Tang et al., 2007).
Additionally, the handshape used to encode the figure or ground also causes problems. Supalla (1982) found that children after 5 years of age produced the correct handshape for moving figures 84–95% of the time, but sometimes used a ‘general classifier’ (a flat hand) instead of a specific one and often omitted the handshape necessary to encode the ground part of the utterance on the secondary hand. Engberg-Pedersen (2003) and Tang et al. (2007) both report that children as old as 6 or 7 years of age frequently omit the handshape representing the ground in spatial descriptions.

These studies reviewed have all looked at fairly old children (in language acquisition terms) and their performance on elicited language tests. These sorts of tasks may involve additional demands from memory or attention. Nevertheless, elicitation is a core methodology in grammar acquisition research (Berko, 1958; Bishop, 2003).

Very few studies have used sentence comprehension to probe when mastery of spatial language occurs in deaf children acquiring signed languages. This methodology is useful because comprehension is less demanding on the child than language production. Martin and Sera (2006) reported on the performance of a group of 11 ASL learning deaf children aged 4–9 years on a comprehension task carried out through sign-picture matching. In Martin and Sera’s (2006) study, the ASL terms ‘away’, ‘right’ and ‘left’ were the most difficult locative items, with children scoring 35–38% correct where chance was 25%. The other terms tested produced the following correct scores for comprehension: ‘towards’ (47%); ‘behind’ (58%); ‘in front’ (60%); ‘below’ (87%) and ‘above’ (91%). The study was largely based on static locations and did not test comprehension of motion events; however, the results suggest that knowledge of location in ASL is continuing to develop in children between the ages of 4–9 years.

By 2 years of age, children acquiring English generally have learned to use prepositions for encoding topological arrangement of objects, e.g., ‘on’, ‘above’ or ‘below’ (Clark, 1972). Later, projective relations are expressed. Children acquiring English, Italian, Serbo-Croatian and Turkish do not produce ‘front/back’ (e.g., ‘the ball is in front of the tree’) until about 5 years of age. The use of ‘left’ and ‘right’ to specify the location of one object with respect to another using three-dimensional Euclidean principles appears still later, at about 11 or 12 years (Choi & Bowerman, 1991; Johnston & Slobin, 1979; Sinha, Thorseng, Hayashi, & Plunkett, 1994).

How children exposed to sign language, as a maternal language, learn to use and understand motion and location forms is intrinsically interesting for language development researchers as it sets up a question regarding the brain’s plasticity for using systematic input when the default modality (speech) shifts to the visual-manual channel. The apparent visual closeness of gesture to sign (what we have been up to now referring to as iconicity) should lead to radical differences in the onset, pattern and mastery of spatial language between modalities.

Two separate studies were carried out to address this question. In a case study of native sign language acquisition between 1 and 3 years, we chart the onset of spatial language, as the child moves from gesture to conventionalised uses of BSL. In the second study we investigate older children’s comprehension of more complex spatial language between 3 and 5 years of age.

5. Method

5.1. Study 1

5.1.1. Subject

The subject is a Deaf boy referred to by the pseudonym Mark, acquiring BSL from his native signing parents and three older siblings (all native signers). There are no available measures of BSL
ability for children of this age, but he displays no developmental impairments and native signers confirmed his signing was typical for a native signer of his age. He was filmed in naturalistic interaction in the home from the ages of 1–10 to 3–0.

5.1.2. Data collection and coding

Deaf and hearing investigators fluent in BSL filmed Mark in 2–3 h sessions at least once a month. The Mark corpus consists of 37 h of spontaneous signing with 3174 child utterances transcribed in the Berkeley Transcription System (BTS), which is compatible with the CHILDES software (Slobin et al., 2001). Independent inter-coder reliability on 10% of the transcription was consistently over 90%. Any disagreement between coders was discussed to a consensus. If agreement was not possible, the item was discounted from the analysis. BTS is ideal for analysing motion and location forms as it specifically codes handshape, path/direction of motion, whether two objects are encoded and, if so, the nature of their spatial relationship (e.g., ‘into’ or ‘beside’).

It is often difficult to decide in signing children younger than 3 years of age whether a communicative manual action is a gesture or a sign. Trained Deaf and hearing researchers fluent in BSL followed standard criteria for including utterances as signed in our analysis. The sign must have been directed toward another person, used spontaneously by the child and not be the direct manipulation of an object or person in the child’s environment (see Abrahamsen, Lamb, Brown-Williams, & McCarthy, 1991; Caselli & Volterra, 1990; Casey, 2003 for more details). Additionally, we compared the child’s signing with the fluent input he received. All utterances were then categorised into two groups, gestures and classifiers. Gestures were motion and location descriptions that could not be characterised as signing based on the above criteria and were made up of the following types:

1. Whole body pantomime depiction. The child himself represents a figure in movement or in a specific location and does not use handshape classifiers to map out the figure.
2. Directional traces. Index finger traces a path without information about the figure through a handshape classifier.
3. Real object manipulation. Hands move onto a real world surface to represent figures, locations and grounds (e.g., V hand moves onto a real bicycle to express ‘person rides bicycle’).

These different types of gestures successfully express different semantic aspects of motion and location events but do not use the representational sign space in front of the signer. The whole body depiction expresses manner information, the index finger trace refers to path and the real object/sign combinations express figure and path but rely on real-world objects to encode the ground information. If the child spontaneously attempted to use classifiers in sign space to represent a figure and its semantic class, as well as a figure’s movement, path, manner or location in sign space, we coded these as classifier utterances. They were made up of the following types:

1. Adult-like utterances.
2. Figure motion or location expressed but ground omitted.
3. Incorrect, or omission of, handshape used for the figure but with a movement and/or ground expressed.
4. Incorrect movement for path but with a figure and/or ground.
5. Two-handed forms for two figures or figure plus ground constructions. In these two-figure utterances we counted both handshapes as separate entities, e.g., ‘two cars cross’ where two flat hands were used to describe the movement of two vehicles.
5.1.3. Results

We observed 17 gestures and 43 classifier utterances in the corpus. Their usage is shown in Fig. 3. Early in the observation period, both groups were used for motion and location descriptions, but from 2 to 6 onwards, gesture forms disappeared and were replaced by classifier utterances.

First we describe the range of meanings found in gesture utterances. Before age 2–0, the whole body depiction was used to describe movements such as lifting the arms for ‘jumping’ and moving the hands forward to describe ‘falling’. These forms encode only the figure and path semantic components. It is not physically possible to encode both figure and ground semantic components simultaneously with the whole body. In order to encode the ground information without recourse to real physical space, some use of a handshape classifier or a second hand interacting as ground is required. Therefore, the child’s attempts to map out figure and motion or location information manually were successful but limited. Using the body as a direct representation of the behaviour of another entity in space is one means of grounding meaning in embodied actions.

Between 2–0 and 2–6, Mark mapped out more event information about ground, path or manner using either finger tracing, real-world objects or the physical ground itself. In several utterances, quite elaborate manners of movement and paths were expressed through tracing of an index finger, e.g., POURING, ZIGZAGGING, PIROUETTING, OVERTAKING and CROSSING-OVER. Each of these motion and location descriptions was preceded by a sign for the nominal CAR, PLANE, MAN, etc., but the child did not combine a handshape classifier for figure with the movement of the hand.

Within the 2–0–2–6 period, Mark exploited gestures with real-world objects. For example, he moved a real toy car in the representational sign space in front of him to depict a figure moving over a bumpy path. Conversely, he moved a flat palm handshape classifier (vehicle) along a table surface or the floor to depict a particular path the figure took. This type of symbolic play using gesture and objects to stand for other entities has been reported for hearing children of this age (Volterra & Erting, 1994). These second types of gesture utterances are more abstract than the whole body depictions as they combine parts of BSL motion and location forms (handshapes and conventionalised movements), but with real-world anchors. Again, gestures that interact with grounded experience (objects, surfaces, textures, etc.) may provide signing children with their first language-to-concept mapping opportunities.

Next, the 43 classifiers were coded for meaning components in BTS and any errors were noted (e.g., in selection of handshape for figure—flat hand for person or something falls down and the child produces an upward movement). The following handshapes were used to encode figures in

Fig. 3. The appearance of classifiers and gestures in Mark’s spontaneous signing.
these 43 examples: extended index finger; spread hand; flat hand; 2-finger hand; bent 2-finger hand; pinkie and thumb hand; fist and an unmarked form coded as ‘relaxed’ open hand. See Fig. 4 for pictures. Mark expressed the following movement and location meanings:

- linear path (forward, down, up);
- circular path (circle, pirouette);
- manner of movement (jump/hop, fly, walk, fall, bob);
- locations and changes of locations (move onto, move off, be under, be behind); postures (stand, sit, be upside down);
- 2-object constructions (meet, move-side-by-side, cross).

The appearance of these components with different figures is shown in Fig. 5. Errors are asterisked and the repetition of the same token in that session is marked with the code (2).

Isolated adult-like forms were observed in the earliest filmed sessions (at 1–10, 1–11, and 2–1). Some of these appear to be quite elaborate encodings of motion and location events (with figure, path and manner semantic components), e.g., *AIRPLANE-MOVE-FORWARD (1–10);
The first two of these examples contain handshape selection errors, with Mark choosing the extended index finger handshape and a spread fingers handshape to represent the plane and helicopter respectively, rather than the pinkie and thumb handshape as in adult input.

Across the classifiers there were eight errors (18.6% of all errors) in selection of handshape, compared with just two errors for verbs of movement and location (4.6%). An example of a path error was CAR-MOVE-SIDWAYS (1–11) when the target was MOVE-FORWARD. Mark’s early and almost errorless use of path and location descriptions is interesting, especially coupled with the apparent semantic complexity of these utterances.

As the child grew older, more diverse motion and location forms were used, but there was a preference for the flat handshape and a limited number of movement and location meanings were expressed. Even at the end of the data-collection period, errors in handshape selection continued. For example, at 3–0, Mark used the relaxed open handshape erroneously on two occasions to describe a figure’s motion (a ball and a car). These errors in handshape selection for classifiers are not phonological problems, as the same handshapes were being used appropriately in lexical signs. Thus, Mark produced a handshape error when describing a vehicle’s movement, while at the same age this flat handshape was error-free in lexical signs such as BOOK, HAPPY or TO-LIKE.

To begin to address the productivity question – does the child have systematic knowledge of how forms are used as part of a system, rather than just providing isolated examples? – four standard criteria for attributing degrees of mastery of grammar were used (Pizzuto & Caselli, 1992; Goldin-Meadow et al., 1995):

1. Single emergence: no clear evidence of analysis of a handshape, motion or location component.
2. Weak evidence of productivity: use of a handshape, movement or location component with single items even on multiple occasions.
3. Strong evidence of productivity: use of a handshape, movement or location component with more than one item.
4. Adult-like use.

Among the utterances in Fig. 5, there is only one form – FLY – that was used with different handshapes (three) up to age 2–6. Thus, there is only weak evidence of productivity of motion and location components before this age. Between the ages of 2–6 and 3–0, FORWARD appeared with four different handshapes and UP with three different handshapes, providing stronger evidence of productivity. We can conclude, therefore, that although Mark used motion forms before 2–6, they did not become systematic until between 2–6 and 3–0.

As a second measure of productivity, the systematic use of the same handshape with different motion and location forms was judged absent before 2–6. The early uses of classifier handshapes were isolated to single motion and location meanings. After 2–6, the flat hand was used with 3 motion forms: FORWARD, FLY and UP. This use constitutes weak evidence of productivity. Strikingly, it was not until after 2–9 that more handshapes began to be combined with different motion and location components. These were the flat hand, appearing with four different motion/locations, the bent-finger and two-finger handshape with four motion/locations each, and the pinkie and thumb handshape with two different motion/locations.

To follow up the development of motion and location forms in BSL, we next carried out a sentence comprehension test with a larger group of older children.
5.2. Study 2

5.2.1. Participants

Two groups of children acquiring BSL as a native language through interaction with their Deaf parents were tested. The younger group consisted of nine children (six girls and three boys) aged 36–48 months (mean 42 months). Two of the younger children were hearing children of Deaf parents. The older group consisted of nine children (six girls and three boys) aged 49–59 months (mean 55 months). Two of the older children were hearing children of Deaf parents. All the children had age-appropriate BSL and non-verbal IQ as measured by a BSL assessment (Herman, Holmes, & Woll, 1999) and subtests of the Snijders-Oomen non-verbal intelligence test (Snijders, Tellegen, & Laros, 1989).

5.2.2. Procedure

After explaining the procedure to the children, 15 sentences were signed in BSL on a video by a fluent Deaf signer using child-appropriate register. All items describe a figure in a location or path. The test sentences are listed in Table 1.

Participants then chose a corresponding picture from a choice of four alternatives presented in a booklet. Targets were accompanied by semantic or grammatical distracters. Picture item number 15 is shown in Fig. 6.

5.2.3. Results

Performance across the two age groups is shown in Fig. 7. Children’s performance in the 3–0 to 3–11 group was 33%, not significantly above chance (25%), compared with 52% by the older children. In the 4–0 to 4–11 group, children’s scores are significantly higher than those of the younger children ($p = 0.002$). However, scores are still fairly low for sentences that encoded projective and Euclidean spatial relations, indicating that comprehension of BSL motion and location sentences is far from complete at 5 years of age.

---

1 The Deaf parents were not themselves native signers, however all were fluent signers and used BSL as their preferred language. It is conventional in the literature to refer to children of Deaf parents as native signers. Hearing children of Deaf parents exposed to BSL from birth onwards are also considered native signers albeit bilingual children. For more details see Petitto et al. (2001).
<table>
<thead>
<tr>
<th>Item number</th>
<th>Spatial relation tested</th>
<th>BSL sentence</th>
<th>English gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Plural</td>
<td>CAR OBJECT-ROW ROW ROW</td>
<td>Rows of cars</td>
</tr>
<tr>
<td>2</td>
<td>Relative location: topological</td>
<td>BOOK OBJECT-ON</td>
<td>The book is on (the bed)</td>
</tr>
<tr>
<td>3</td>
<td>Relative location: topological</td>
<td>TEDDY OBJECT-LOCATED</td>
<td>One teddy</td>
</tr>
<tr>
<td>4</td>
<td>Relative location: topological</td>
<td>BALL TABLE OBJECT-ON</td>
<td>The bull is on (the table)</td>
</tr>
<tr>
<td>5</td>
<td>Movement/path</td>
<td>OBJECT-TWO-PEOPLE-MEET</td>
<td>Two people meet</td>
</tr>
<tr>
<td>6</td>
<td>Relative location: topological</td>
<td>DOG OBJECT-IN</td>
<td>The dog is in (the box)</td>
</tr>
<tr>
<td>7</td>
<td>Movement/path</td>
<td>OBJECT-PERSON-COME-DOWN-ESCALATOR</td>
<td>The person is coming down the escalator</td>
</tr>
<tr>
<td>8</td>
<td>Plural</td>
<td>OBJECT-FEW-CUPS</td>
<td>A few cups</td>
</tr>
<tr>
<td>9</td>
<td>Relative location: projective</td>
<td>CAR OBJECT-BEHIND</td>
<td>The car is behind (the house)</td>
</tr>
<tr>
<td>10</td>
<td>Relative location: projective</td>
<td>BOX BED OBJECT-UNDER</td>
<td>The box is under (the bed)</td>
</tr>
<tr>
<td>11</td>
<td>Relative location: projective</td>
<td>OBJECT-IN-QUEUE</td>
<td>A queue</td>
</tr>
<tr>
<td>12</td>
<td>Relative location: Euclidean</td>
<td>DOG OBJECT-IN-FRONT</td>
<td>The dog is in front (of the box)</td>
</tr>
<tr>
<td>13</td>
<td>Relative location: Euclidean</td>
<td>OBJECT-CAR-ROW-BOTTOM-LEFT</td>
<td>The row of cars is in the bottom left (of the picture)</td>
</tr>
<tr>
<td>14</td>
<td>Relative location: Euclidean</td>
<td>DOG OBJECT-INSIDE-RIGHT</td>
<td>The dog is lying inside to the right side (of the box)</td>
</tr>
<tr>
<td>15</td>
<td>Relative location: Euclidean</td>
<td>HOUSE OBJECT-TOP-RIGHT</td>
<td>The house is in the top right part of a cross-roads scene</td>
</tr>
</tbody>
</table>
Fig. 7. Mean comprehension scores on 15 items by age group.

The test sentences focused on different domains of meaning: movement and path descriptions (item 5 and 7); relative locations (items 2, 4, 6, 9, 10, 12, 13, 14 and 15); and Pluralisation through spatial means (items 1, 3, 8 and 11).

We present results only for Path and Locations. The younger group – Path 39% (S.D. = 42) and Location: 29% (S.D. = 14) – performed less well than the older group – Path 61% (S.D. = 42) and Location 53% (S.D. = 21) – in both domains. An independent-groups \( t \)-test showed a non-significant difference between groups for Path and a significant difference for Location (\( p = 0.011 \)); see Fig. 8.

Finally, we broke down scores by age group on just the eight location items. Scores for these items are presented in Fig. 9. The older group outscored the younger one on all sentences. The
items that required some reversal in perspective were most difficult for all children, especially right-left relations. All children scored below chance on right-left distinctions. Herman et al. (1999) reported that children by age 11–12 years, as well as, adult signers scored at ceiling on these items.

6. General discussion

Results from both studies reported here suggest that children exposed to sign language learn the correct meaning-to-form mappings in their language very gradually in spite of the available iconicity. In answer to our main question concerning the impact of modality on language acquisition, we find the onset, pattern and mastery of spatial semantics in a visual-manual language is very similar to acquisition studies of different spoken languages (Brown, 2001; Casasola, 2005; Johnston & Slobin, 1979; Sowden & Blades, 1996; Varela, 2006). Our findings add a signed language to the literature on cross-linguistic comparisons of spatial semantics in language acquisition. They also reinforce the need to analyse more closely the role of gesture in hearing children’s development of spatial language. Many meaning components were expressed by the child in our case study through gesture before he developed conventional signs. The same may be true of hearing children’s first use of gesture.

6.1. Embodiment and uses of gesture in early sign language acquisition

The results of Study 1 suggest embodied understanding of movement and location concepts can be mapped onto gesture significantly before conventionalised signs have been learned (Evans, Alibali, & McNeill, 2001; Howell et al., 2005). Because of the closeness of natural gestural communication to sign language in this domain, the visual-manual modality lends itself to early gestural means of communicating spatial events. The acquisition of spatial language beyond these early gestures requires a distancing from embodied descriptions to conventionalised uses of representational space.

A significant skill for encoding spatial semantics efficiently in BSL is mastery of the coordination of manual articulators within an external area in front of the signer. Path expressions emerge early, even in the one- and two-word speech of children (Bloom, 1973; Choi & Bowerman, 1991). Mark’s early signs encoded simple movements of single figures with linear paths, such as UP and FORWARD. These descriptions are expressed through whole body depictions, finger traces and actions onto objects. Movement of the whole body acting as a one-to-one map between a person’s movement and its expression is perhaps less abstract than a conventionalised move-
ment of a handshape in representational sign space. There were very few location meanings or explicit descriptions of ground across the data. Reasons for this may stem from the linguistic complexity of these forms in the grammar of BSL, the cognitive demands of using two hands to express the figure-ground information simultaneously or the mapping of more complex spatial meanings involving perspective and three-dimensional reference onto the manual articulators in a representational signs space. For encoding projective and Euclidean concepts, the correct choice of handshape, position in space and addressee viewpoint needs to be considered (Slobin et al., 2003).

Some of Mark’s gestures, where the handshape acts as a figure but interacts with real-world space (e.g., a two finger hand moving onto a real bicycle in the child’s environment, to express a person riding) may be an intermediate stage between gesture and more linguistic use of signs in a representational space. The figure and ground elements of motion and location conventionalised forms in BSL require manipulation of two handshapes. Children learning sign may interact with real-world objects, as substitutes for the figures (e.g., real cars) or grounds (e.g., real bicycles). More spontaneous data on sign language acquiring children of this age is required to substantiate this suggestion.

6.2. Mastering conventions

Language provides discrete, categorical and combinatorial symbols to express meanings previously expressed holistically through gesture (Volterra & Erting, 1994). Children learning sign have to build up categories of form-meaning pairs. For example, ten different types of zigzag event may be described initially with close adherence to how each of the objects zigzagged in the real world, through motion gestures. As language develops, children begin to form a category for similar but not identical movements that can be described with the same generalised BSL motion verb, ‘object-zigzags’. The figure component in this utterance can be varied, by using different classifier handshapes to describe a person, small animal or vehicle.

Natural gesture can be used to express parts of events involving figures located and moving in space. Is exposure to signed language a prerequisite for the mastery of spatial semantics? Deaf children not exposed to sign language have been reported to create gesture-based communication systems, referred to as homesign (Goldin-Meadow, 2003; Zheng & Goldin-Meadow, 2002). Morford (2002) examined narratives of two adolescent homesigners in order to investigate whether homesign shares characteristics with ASL in the expression of motion events. Morford asked whether the homesigners would combine the elements of figure, ground, path and manner in single signs. It was found that their use of homesign did not resemble ASL in this respect. In particular, the homesigners combined fewer conceptual elements in their signs, and one of the two homesigners rarely encoded path at all.

This is difficult to understand as we observed path in the very early sign-like communication of Mark, at 2 years of age, and path sentences were clearly understood by 3-year-olds in Study 2. The difference may be due to methods for collecting spontaneous versus elicited data. The types of event descriptions required in the Morford study were fairly complex, involving Frog story narratives, where each part of the story has several elements to be described. Homesigners might have done better on spontaneous spatial event descriptions or in less demanding language comprehension tasks. Path descriptions required in these Frog story scenarios might be crucially tied into exposure to language models. The importance of language exposure for later skill in encoding motion, location, manner and path components onto the classifier system is also substantiated in studies of deaf children learning ASL from non-native input. Motion and location
constructions can develop to full mastery in young deaf children, even if the input is non-optimal (Ross & Newport, 1996).

Morford concluded that the development of discrete, categorical and combinatorial symbols is not an inevitable outcome of either the visual-spatial modality or the iconicity of meaning to form links (Morford, 2002). Morford’s study is an important one, as it focuses on the effects of isolation from accessible adult linguistic models, on subsequent language development. Adolescent homesigners are able to exploit gestures, but without a language model these gestures do not extend to conventionalised language. The crucial contact with an adult language model provides the resources to talk about and understand motion and location events with categories and combinations of form-meaning pairs in a grammar. Slobin (1996) writes that the child determines appropriate mappings between form and meaning on the basis of patterns through the exposure to specific languages.

In the sentence comprehension study reported in Study 2, the simple path descriptions were more easily understood than relative locations by the 3-year olds. The distinction between horizontal and vertical axis is observed in spoken language acquisition (Clark, 1972) and also in sign language (Martin & Sera, 2006). Performance in the current study, in both age groups, on sentences with right-left distinctions was poorer than top-bottom. The 3-year-olds scored below chance (25%) on ‘behind’, ‘under’, ‘in-front’, ‘bottom-left’, ‘inside-right’ and ‘top-left’. All of these meanings are encoded by two hands simultaneously describing a figure-ground relationship from the signer’s perspective. They require that the addressee map the signs onto a conceptual representation space, which has to be reversed to match the pictures the children were asked to choose. Some of these sentences were still not understood by several of the 4-year-olds.

Therefore, once deaf children are beyond talking about figure and motion meanings in topological scenes, they have to master the conventions for expressing projective and Euclidean meanings. Recall that in BSL, figures have to be first mentioned through nouns followed immediately by appropriate classifiers positioned in sign space. All of the sentences in Study 2, which encoded relative locations, required our subjects to identify both figure and ground entities, establish the spatial relationship expressed between the two hands in relation to their own viewer’s perspective and, finally, match this meaning information to a set of corresponding pictures.

Newport and Meier (1985) first highlighted that in ASL production tasks, children younger than 5 years of age have difficulty integrating the handshapes necessary to encode the correct figure-ground components and often separate out these two semantic elements. Newport and Meier (1985) explained these errors as reflecting the child’s difficulty with the morphological complexity of these constructions.

In the present study, the comprehension of two-handed, relative location meanings is very similar to what has been reported in previous studies of spoken language development. Children learn the conventions for specifying the location of one object with respect to another, using projective and Euclidean principles at around 11 or 12 years (Choi & Bowerman, 1991; Johnston & Slobin, 1979; Sinha et al., 1994). We can add that in comprehension tasks some cognitive reversal of perspective also contributes to the protracted development of two-handed constructions that encode relative locations. Piaget and Inhelder (1956) first argued that three-dimensional spatial relations were a late development because they required an understanding of the integrated horizontal and vertical coordinates, used to observe and represent the physical world.

Although classifiers located in signing space are a close visual representation of the concepts they encode, children still have to be able to map these concepts onto language using conventional linguistic forms. Deaf children acquiring sign languages as natural first languages do not exploit
this iconicity and enjoy a developmental advantage compared with hearing children learning spoken language spatial forms. As reported for other linguistic domains (Meier, 1987; Petitto, 1987) the mapping problem is not avoided even in the domain of space. Children learning a language crafted in the visual-spatial modality are confronted with the same complex concept-to-language mapping problem as their hearing peers.

Acknowledgements

We thank Dan Slobin, Adam Schembri and Laura Lakusta and three anonymous reviewers for comments and suggestions, Penny Roy and Sian Hurren for assistance with data analysis and coding. This work was supported by a grant from the ESRC.

References


