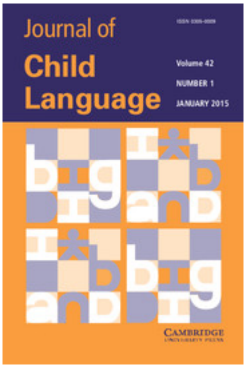



# On the way to language: event segmentation in homesign and gesture


ASLI ÖZYÜREK, REYHAN FURMAN and SUSAN GOLDIN-MEADOW



Journal of Child Language / Volume 42 / Issue 01 / January 2015, pp 64 - 94
DOI: 10.1017/S0305000913000512, Published online: 20 March 2014

*J. Child Lang.* **42** (2015), 64–94. 
doi:10.1017/S0305000913000512

# On the way to language: event segmentation in homesign and gesture*

ASLI ÖZYÜREK

*Radboud University Nijmegen and Max Planck Institute for Psycholinguistics, the Netherlands*

REYHAN FURMAN

*University of Alberta, Edmonton, Canada*

AND

SUSAN GOLDIN-MEADOW

*University of Chicago, USA*

(*Received* 26 *June* 2012 – *Revised* 16 *May* 2013 – *Accepted* 20 *October* 2013 – *First published online 20 March 2014*)

ABSTRACT

Languages typically express semantic components of motion events such as manner (roll) and path (down) in separate lexical items. We explore how these combinatorial possibilities of language arise by focusing on (i) gestures produced by deaf children who lack access to input from a conventional language (homesign); (ii) gestures produced by hearing adults and children while speaking; and (iii) gestures used by hearing adults without speech when asked to do so in elicited descriptions of motion events with simultaneous manner and path. Homesigners tended to conflate manner and path in one gesture, but also used a mixed form, adding a manner and/or path gesture to the conflated form sequentially. Hearing speakers, with or without speech, used the conflated form, gestured manner, or path, but rarely used the

[*] This research was supported by an ERC Starting Grant given to the first author and R01DC00491 from NICHD to the last author. We would like to thank Koc University in Istanbul, Turkey, for logistic support during the project and all the deaf and hearing participants in their cooperation and for welcoming us into their homes. We would also like to thank Professor Shanley Allen and Professor Sotaro Kita for sharing the stimulus materials, the creation of which was supported by an NSF Grant awarded to them and to the first author. Finally, we are grateful to Carolyn Mylander, Tilbe Goksun, Engin Arik, Beyza Sumer, and Burcu Sancar for helping in data collection and coding. Address for correspondence: Asli Özyürek, Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525JT, Nijmegen, the Netherlands. e-mail: asliozu@mpi.nl

mixed form. Mixed form may serve as an intermediate structure on the way to the discrete and sequenced forms found in natural languages.

## INTRODUCTION

Events of motion contain many aspects that can occur simultaneously. Take, for example, a person running downstairs. The runner is pumping his legs and moving downward at the same time. One of the key features of human language is that it can take a holistic event, break it into segments, and express the segments sequentially, allowing speakers to focus on each of the decomposed semantic elements of the event in turn. For example, in English, we describe the running event as *run down*, representing the manner of the movement (*run*) separately from its trajectory (*down*). But dividing holistic events into segments and then combining those segments into organized strings is not the only way we can represent events. We use maps, pictures, pantomimes, etc., which represent events more holistically; for example, a mime who is depicting an act produces movements that correspond, part-for-part, to the act, thus evoking the act as a whole. Segmented (but not holistic) representations have the potential to give rise to new combinations, one of the basic design features of language (Hockett, 1960) and a characteristic of all spoken languages (Talmy, 1985), as well as established (Supalla, 1990) and newly emerging (Senghas, Kita & Özyürek, 2004) sign languages. Moreover, segmented representations allow speakers and signers to focus on a single piece of an event (e.g. to highlight manner and not path) in a way that holistic representations do not (e.g. miming movement along a path inevitably brings with it a depiction of manner).

We ask in this paper whether segmenting and sequencing the pieces of a motion event is such a central and robust feature of human language and communication that it can be reinvented by a child who does not have access to a conventional language. We tackle this question by comparing the gestures produced by deaf children not exposed to accessible input from a conventional language (homesigners) to gestures produced by hearing speakers (adults, children, and their own hearing mothers) in the same community.

### *Componentialization in hearing and deaf children*

Children exposed to conventional languages learn to componentialize elements of a motion event early in development, whether they are learning a spoken language (Allen *et al.*, 2007; Choi & Bowerman, 1991; Özçalışkan & Slobin, 1999) or a sign language (Supalla, 1982). However, not all children are exposed to models of language. Deaf children whose hearing losses are so severe that they cannot acquire spoken language, and whose hearing parents have not exposed them to sign language, lack an accessible model for language. Nevertheless, these children communicate using gestures, called homesigns (Goldin-Meadow, 2003).

Homesigns are characterized by many, although not all, of the properties of natural language, including a stable lexicon (Goldin-Meadow, Butcher, Mylander & Dodge, 1994), structure at word (Goldin-Meadow, Mylander & Butcher 1995; Goldin-Meadow, Mylander & Franklin, 2007) and sentence (Feldman, Goldin-Meadow & Gleitman, 1978; Goldin-Meadow & Feldman, 1977) levels, sentence-level negation and question modulators (Franklin, Giannakidou & Goldin-Meadow, 2011), nominal constituents (Hunsicker & Goldin-Meadow, 2012), grammatical categories such as subject (Coppola & Newport, 2005), and a demarcated distinction between noun and verb categories (Goldin-Meadow *et al.*, 1994). The question we ask here is whether homesigners introduce a language-like segmentation strategy (e.g. a rolling gesture produced in place, followed by a gesture moving across space) into gesture sentences that express both Manner and Path of spontaneous motion.

The gestures that homesigners produce have, in fact, been found to display one type of segmentation and combination. For example, in order to describe putting down a round penny, a homesigner child first held up a 'round' handshape (thumb and index forming a circle, 'penny'), followed by a flat palm moved downward ('down'), thus producing two segmented gestures strung together ('penny–down') rather than a single holistic gesture that combined both semantic elements (i.e. moving the circle-shaped hand down, 'penny + down'; Goldin-Meadow, 2003; Goldin-Meadow *et al.*, 1995). Even hearing adults, when asked to use their hands without speech to describe an event, will produce segmented gestures, each representing a different semantic element. For example, when asked to describe with their hands a simple event in which a circle moves diagonally across the screen, hearing speakers behave like homesigners – they produce two separate gestures, one representing the circle ('penny') and one presenting the diagonal downward movement down ('down') (Gershkoff-Stowe & Goldin-Meadow, 2002; Goldin-Meadow, McNeill & Singleton, 1996). Segmenting semantic elements (in this case, figure and path) out of a motion event thus appears to be a basic aspect of cognition, easily incorporated into communication (see Goldin-Meadow, So, Özyürek & Mylander, 2008, for evidence that hearing adults can segment other semantic elements, e.g. the patient and the endpoint, out of motion events when using their hands to describe the event).

However, segmenting a figure from its path might be different from segmenting two movements that take place simultaneously (e.g. separating the act of running down the street into the PATH along which the runner moves and the MANNER of movement that propels the runner along the path). As mentioned earlier, despite the simultaneity of path and manner in the actual event, established languages rarely conflate manner and path into the same lexical item and instead encode manner and path in two separate lexical items, using different lexicalization patterns depending on the typology of the language (Talmy, 1985). For example, in English (a satellite-framed language), manner is expressed in the verb and path in a satellite, as in *The child runs* (manner) *down* (path) *the street*. In contrast, in Turkish (a verb-framed language), path is expressed in the main verb and manner in a subordinate verb as in *Çocuk koşarak tepeden aşağı indi* – 'child as running (manner) descended the hill (path)'.[1] The sign languages in which path and manner have thus far been studied, American Sign Language (ASL) and the Sign Language of the Netherlands (SLN, also known as Nederlandse Gebarentaal, NGT), also convey manner and path in separate lexical items in verbal predicates. For example, Slobin and Hoiting (1994) have proposed that sign languages use serial-verb constructions (manner-path) and are best characterized as complex verb-framed languages (although manner and path can be combined within a single sign in some classifier constructions in sign languages; Supalla, 1990). Action segmentation of this sort has even been observed in newly emerging sign languages (e.g. Nicaraguan Sign Language; Senghas *et al.*, 2004). Do we see action segmentation and sequencing in homesign?

Previous research has shown that American and Chinese child homesigners are able to segment manner and path into separate gestures when communicating about crossing space events – they produce gestures conveying path alone and gestures conveying manner alone, in addition to producing manner + path gestures (Zheng & Goldin-Meadow, 2002). However, the previous work did not ask whether manner and path gestures were combined within a single gesture sentence and, if so, how those sentences were structured. Nor did the study examine the gestures produced by hearing individuals in the community, gestures that might have served as input to the homesign system.

## PRESENT STUDY

To determine whether young children who are creating their communication systems without benefit of a community of language users introduce action segmentation and combination into those systems, we asked Turkish

1 Note that Turkish is different from other verb-framed languages (e.g. Spanish) in that manner is not expressed as a gerund, but as a subordinate verb linked to the main verb with the connective morpheme *-arak* (in *koşarak*). The awkward translation offered in the text is meant to capture this subordinate clause arrangement.

homesigners, seven in Study 1 and five in Study 2, to gesture about a series of events designed to elicit manner and path descriptions.

Homesigners are not exposed to a conventional sign language. However, they do see the gestures that their hearing parents and other hearing people around them produce as they talk (cf. Iverson, Capirci, Longobardi & Caselli, 1999; Özçalışkan & Goldin-Meadow, 2005; Shatz, 1982). Thus, we compared the gestures produced by the homesigners to gestures that the children's hearing mothers and eighteen other hearing adults in Study 1 produced in response to the motion events to determine whether the gestures hearing speakers produce provide a model for the deaf children's homesigns.

Hearing children are also exposed to the gestures for motion events produced by adult speakers of their language. The difference, however, is that the hearing children experience other people's gestures, and produce their own gestures, in the context of speech. Thus, we examined the gestures that Turkish hearing children, fourteen in Study 1 and five in Study 2, produced in response to the motion events to determine whether experiencing and producing gesture in the context of speech matters compared to using gestures only.

Finally, we explore whether the patterns found in the deaf children's gestures, if not copied from hearing speakers' gestures, might be a response simply to the fact that the manual modality is the deaf child's sole means of communication. If so, requiring hearing adults to describe motion events using gesture without speech might result in gestures that resemble the deaf children's. As mentioned earlier, we know that when hearing adults are called upon to use the manual modality as their sole means of communication, they can segment semantic elements such as figure and path into separate gestures and combine those gestures into structured strings (Goldin-Meadow *et al.*, 1996, 2008). To determine whether segmentation and combination of manner and path will also arise when hearing speakers rely solely on the manual modality, we asked the eighteen Turkish hearing adults who initially described the events in speech in Study 1 to describe them a second time in Study 3, this time using only their hands.

In sum, we investigate whether homesigns – gestural systems that develop in a deaf child without conventional language input – contain the roots of action segmentation and sequencing. We compare three conditions under which action segmentation and combination have the potential to arise in the manual modality: (i) when gestures have been a child's only means of communication throughout development (deaf homesigners); (ii) when gestures are produced along with speech (hearing adults, hearing children, hearing mothers of the deaf homesigners gesturing while talking); and (iii) when gestures are recruited on-the-spot to replace speech (hearing adults

TABLE 1. *Ages of the seven deaf homesigners at each data collection session in Study 1 (in years;months)*

| Name | I | II | III | IV | V | VI |
|---|---|---|---|---|---|---|
| Rana | 3;2 | 3;4 | 3;5 | 3;7 | 3;8 | 3;10 |
| Kaan | 3;4 | 3;7 | 3;8 | 3;9 | 3;11 | 4;0 |
| Sina | 4;4 | 4;6 | 4;8 | 4;9 | 4;10 | 5;0 |
| Nur | 4;5 | 4;6 | 4;7 | 4;9 | 4;11 | 5;0 |
| Irem | 4;8 | 4;11 | 5;1 | 5;3 | 5;4 | 5;6 |
| Emre | 4;4 | 4;6 | 4;9 | 4;10 | 5;0 | 5;5 |
| Ela | 4;9 | 4;11 | 5;2 | 5;3 | 5;5 | 5;7 |

gesturing without talking). We focus on descriptions of spontaneous motion events (i.e. actors moving across space on their own). In these events, manner (e.g. rolling) takes place throughout the crossing-space event (e.g. moving down). Events of this sort can be represented in the manual modality either holistically (rotating the hand while moving it down), or componentially (rotating the hand in place, followed or preceded by moving the hand down), and thus provide fertile ground for exploring the conditions that give rise to action segmentation and combination.

## STUDY 1

### METHOD

#### *Participants*

Seven Turkish deaf children, ranging in age from 3;2 (years; months) to 4;9 ($M = 4;2$), participated in a longitudinal study and were videotaped at home every one to three months (Table 1). The children were congenitally deaf, with bilateral hearing losses (70–90 dB), and no other reported cognitive or physical disabilities (in Turkey, there are few opportunities for deaf children to be given normed cognitive evaluations; however, during our year-long observations, we did not notice any major cognitive or social deficiencies in the children in our study – they all performed our tasks without difficulty). The children's hearing parents had chosen to educate them using oral methods. None of the children had cochlear implants, but all wore hearing aids, although they did not use them regularly and, in addition, had very little (if any) speech therapy. Although able to produce an occasional Turkish word, the children did not combine words into sentences. Moreover, none had been exposed to conventional sign language or had contact with another deaf child or adult. The deaf children had not attended preschool of any sort during the observational period and spent their days at home with their mothers (at the time of our observations, deaf children did not begin school until age seven in Istanbul; the first preschool for deaf children was established after this study was conducted). The deaf children's hearing mothers also participated in the study.

In addition, fourteen Turkish hearing children, ranging in age from 3;0 to 6;10 ($M = 4;9$) and drawn from families of the same socioeconomic status as the deaf children were videotaped at home, and eighteen Turkish adults, undergraduate students in Istanbul, were videotaped on campus of a Turkish university (Koç University). All hearing participants were native Turkish speakers.

### *Procedure*

Hearing participants (adults, children, mothers of deaf children) were told that they would see a series of animated vignettes on a laptop, and were asked to tell the experimenter what happened after each vignette; the speakers gestured spontaneously while talking and it was these co-speech gestures that we compared to the homesigners' gestures. To elicit responses from the deaf homesigners, after each vignette, the experimenter produced a two-handed flip gesture along with a quizzical look and pointed at the screen.

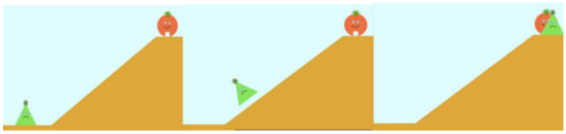
All participants were shown six spontaneous motion events highlighting manner and path (Özyürek, Kita & Allen, 2001), along with thirty-six other action events (Goldin-Meadow *et al.*, 2008), in random order. During the retelling, a still picture of the initial scene of the event, which included all objects in the event, was placed in front of the participants as a memory aid. The children, and occasionally the adults (particularly when asked to gesture without speech), pointed at the picture as a way to refer to an object in the event or traced a trajectory on the picture to refer to the path or the manner.

The deaf children were part of a longitudinal study and thus were shown the events six times at sessions taking place over the course of several months (see Table 1). The hearing children and the deaf children's hearing mothers described the events once. Four of the mothers told the vignettes to their own children; three told them to the experimenter. The mothers performed the task after the sixth session; thus all of the deaf children described the vignettes before their mothers did their descriptions.

The hearing adults described the vignettes twice, first using speech and whatever gestures they spontaneously produced, second using gesture and no speech, always in the same order. The gestures that the hearing adults produced without speech will be described in Study 3.

### *Materials*

We focused on the six animations (each 6–15 seconds) designed to highlight simultaneous manner and path of spontaneous events: ROLL + ASCEND, ROLL + DESCEND, ROTATE + ASCEND, ROTATE + DESCEND, JUMP + ASCEND,



| | | |
|---|---|---|
| Entry event: Triangle Man slides horizontally in from the left | Target event: Triangle Man jumps as it ascends the hill | Closing event: Triangle Man slides in front of the Tomato Man at top of the hill |

Fig. 1. (Colour online) Selected scenes from the JUMP + ASCEND motion event.

JUMP + GO AROUND. For example, ROTATE + ASCEND involved an animated object turning on its horizontal axis as it ascended vertically in the air. Each clip involved a round red smiling character and a triangular-shaped green frowning character moving within a simple landscape. All the vignettes were designed so that the target event – which was always in the middle of the sequences of events – had a distinct beginning and a distinct end point (Figure 1); these events were used in Özyürek *et al.*, 2005, 2008), Kita *et al.* (2007), and Allen *et al.* (2007) to investigate expressions of spontaneous motion events with simultaneous manner and path in hearing speakers. The path of the moving figure in the target event always followed a different direction from the path followed in the entry and closing events (e.g. from a horizontal path in the entry event in Figure 1, to a diagonal movement up the hill in the target event, followed by another horizontal path in the closing event. The manner also occurred only in the target event. This design made it easy to identify the boundaries of the target event in speech and in gesture during coding.

In creating the stimuli, we made sure that the target motion event was indeed a spontaneous (i.e. not caused) event. In two of the events, ROLL + DOWN and ROLL + ASCEND, the figure in the target event is given a bump during the entry event, but when the figure changes direction in the target event, it is clear that it is moving under its own steam. In the four other events, there is no bump during the entry event.

*Coding*

As in previous papers (Allen *et al.*, 2007; Kita *et al.*, 2007; Özyürek *et al.*, 2005, 2008), we included only the gestures that displayed the direction of the figure's path during the target event. The stroke (meaningful phase) of the gesture (Kendon, 1980; McNeill, 1992) was used to segment gestures and determine their meaning. To determine onset and offset of gesture strokes, we considered changes in the parameters of shape, placement of the hand, trajectory of motion, and tension of the hands (for more on gesture phases and how to recognize and code them, see Kita *et al.*, 1998).

Following Kita *et al.* (2007) and Özyürek *et al.* (2005, 2008), gestures were divided into three categories:

1. Path gestures depicted the trajectory that the moving object took (e.g. ascending movement of the hand representing moving upward).
2. Manner gestures depicted the manner by which the object moved as it changed its location (e.g. repetitive circular movement representing rolling).[2]
3. Manner + Path gestures simultaneously depicted both manner and path within the gesture's stroke (e.g. hand moves repetitively in a circle as it ascends representing rolling upward). Single points to objects were not included in the analysis. However, points that traced either the trajectory of a path or the manner of movement were included.

We also coded gesture strings. Our goal was to analyze gestures in the same way for all participants. We therefore needed to consider gesture without regard to speech. We divided gestures into strings using motoric criteria. Following Goldin-Meadow and Mylander (1984), string breaks were coded when participants paused or relaxed their hands. On average, pauses lasting longer than 1·5 s. constituted a break between gesture strings. A string could contain one or more gesture strokes, e.g. a circular movement of the hand (stroke 1, Manner) followed, without pause, by a downward movement of the hand (stroke 2, Path) was considered two gestures within a single string.

We categorized gesture strings into five types:

1. Path alone (no gestures referring to manner).
2. Manner alone (no gestures referring to path).
3. Conflated (both manner and path were produced within a single gesture stroke, i.e. Manner + Path, with no other gestures referring to manner or path in the string).
4. Sequenced (at least one Manner gesture conjoined with at least one Path gesture and no Manner + Path gestures in the string).
5. Mixed (a Manner + Path gesture combined with a Manner or Path gesture, i.e. two gestures). When a combination included Manner + Path gesture combined with a Manner and a Path gesture (i.e. three gestures), we also coded it as a Mixed form (although this happened rarely).

2 Occasionally participants used their entire bodies to enact the manner of the event. The homesigners produced a total of thirteen of these full body enactments throughout the study (2·2 per session, approximately 0·3 per child), the hearing children produced three (approximately 0·2 per child), and the silent hearing adults produced three (approximately 0·2 per adult). All of these gestures conveyed manner of motion without path and were therefore counted as Manner gestures.

Reliability was calculated on 30% of the participants' event descriptions in each group by two independent coders. Agreement between coders was 94% for categorizing both gestures and gesture strings (Cohen's Kappa score was 0·87 for both).

All of the statistical analyses reported below are omnibus ANOVAs followed by post-hoc tests exploring pairwise comparisons when needed.

## RESULTS

### *Homesigners across sessions and ages*

Homesigners are unique in that they are not able to make use of the spoken language that surrounds them, nor are they exposed to a conventional sign language. The homesigners were participating in a longitudinal study and were shown the vignettes at each of the six sessions at which they were observed (see Table 1). In order to compare their data to the data collected on the hearing participants, who responded to the vignettes only once, we collapsed the homesigners' scores across the six sessions, using the mean across sessions for each homesigner. The decision to use a single score for each homesigner was motivated by the fact that we did not find differences in the string types homesigners produced across sessions. For each type of gesture string, we performed a repeated measures ANOVA on the number of strings produced by each child with session (1 through 6) as the within-subject factor. Furthermore, to determine whether age of the child contributed to each effect, the age of each child at session 1 was included as a covariate in each ANOVA analysis since the children began the study at different ages (see Table 1). The analyses revealed no main effect of session for any of the gesture string types: Path Only strings ($M$= 2·36, $SE$=2·37; $F(5,25)=0·26$, $p=·80$, partial $\eta^2=·05$); Manner Only strings ($M$=0·61, $SE$=0·98; $F(5,20)=0·39$, $p=·67$, partial $\eta^2=·09$); Conflated strings ($M$=2·52, $SE$=2·15; $F(5,25)=0·19$, $p=·80$, partial $\eta^2=·04$); Mixed strings ($M$=1·21, $SE$=1·34; $F(5,25)=0·35$, $p=·73$, partial $\eta^2=·07$); Sequenced strings ($M$=0·26, $SE$=0·43; $F(5,25)=2·58$, $p=·11$, partial $\eta^2=·34$). No significant interaction effects were found between mean number of gesture strings per session and age for any of the gesture string types: Path Only strings ($F(5,25)=0·26$, $p=·78$, partial $\eta^2=·05$); Manner Only strings ($F(5,20)=0·50$, $p=·61$, partial $\eta^2=·11$); Conflated strings ($F(5,25)$= 0·30, $p=·71$, partial $\eta^2=·06$); Mixed strings ($F(5,25)=0·36$, $p=·72$, partial $\eta^2=·05$); Sequenced strings ($F(5,25)=2·65$, $p=·10$, partial $\eta^2=·34$).

### *Homesign compared to co-speech gesture*

We looked first at the total number of manner and path gestures participants produced and found no significant differences across groups: ($M$=9·9,

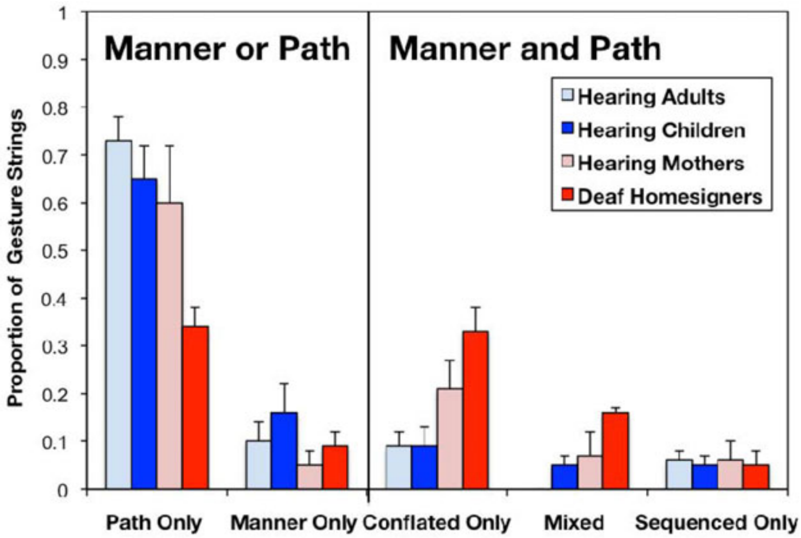



Fig. 2. (Colour online) The mean proportion of gesture strings produced by participants in each group in Study 1 to describe the set of motion events: hearing adults, hearing children, hearing mothers of deaf children, all of whom produced their gestures while talking, and deaf children who used homesign rather than speech to communicate. Gesture strings contain one or more gestures and are classified according to type (Path only, Manner only, Conflated, Mixed, Sequenced). Error bars reflect standard errors.

$SE = 1{\cdot}1$) per participant across the six vignettes for homesigners; ($M = 6{\cdot}2$, $SE = 0{\cdot}6$) for hearing adults; ($M = 6{\cdot}3$, $SE = 1{\cdot}1$) for hearing children; ($M = 7{\cdot}7$, $SE = 2{\cdot}0$) for hearing mothers; ($F(3,42) = 2{\cdot}00$, $p = {\cdot}13$, partial $\eta^2 = {\cdot}14$), one-way ANOVA with group (homesigners, hearing adults, hearing children, hearing mothers) as the independent factor, and total number of manner and path gestures as the dependent factor. We then looked at the total number of gesture strings participants produced and also found no significant differences across groups: ($M = 7{\cdot}0$, $SE = 0{\cdot}9$) gesture strings per participant across the six vignettes for the homesigners; ($M = 5{\cdot}0$, $SE = 0{\cdot}5$) for the hearing adults; ($M = 5{\cdot}4$, $SE = 0{\cdot}8$) for the hearing children; ($M = 6{\cdot}4$, $SE = 1{\cdot}4$) for the hearing mothers; ($F(3,42) = 1{\cdot}15$, $p = {\cdot}34$, partial $\eta^2 = {\cdot}08$) one-way ANOVA with group (homesigners, hearing adults, hearing children, hearing mothers) as the between-subjects independent factor, and number of gesture strings as the dependent factor.

We turned next to the TYPES of gesture strings the participants produced. Figure 2 presents the number of gesture strings of each type produced by a participant, taken as a proportion of all gesture strings that the participant produced, and averaged across all of the participants within each of the four groups. In other words, the proportions in this figure (and all subsequent figures) were calculated by individual and then averaged to create a mean proportion per group.

We focused first on the gesture strings containing 1-event component strings, that is strings containing either a Path or a Manner gesture (the two sets of bars on the left). A 4×2 ANOVA with GROUP (homesigners, hearing adults, hearing children, hearing mothers) and STRING TYPE (Manner only, Path only) as independent factors, and proportion of 1-event component strings as the dependent factor, revealed an effect of group ($F(3,84)=3{\cdot}2$, $p=\cdot03$, partial $\eta^2=\cdot1$); an effect of string type ($F(1,84)=90{\cdot}43$, $p<\cdot000$, partial $\eta^2=\cdot53$); and no interaction ($F(3,84)=2{\cdot}45$, $p=\cdot07$, partial $\eta^2=\cdot08$). LSD pairwise post-hoc comparisons revealed that homesigners used 1-event component strings significantly less often ($M=0{\cdot}43$, $SE=0{\cdot}04$) than hearing adults ($M=0{\cdot}84$, $SE=0{\cdot}05$) ($p=\cdot005$) and hearing children ($M=0{\cdot}80$, $SE=0{\cdot}07$) ($p=\cdot012$), but not significantly less often than their mothers ($M=0{\cdot}65$, $SE=0{\cdot}12$) ($p=\cdot199$). All groups used more Path only strings than Manner only strings.

We then examined 2-component gesture strings in which both manner and path were conveyed. We again calculated the number of gesture strings of each type produced by a participant, taken as a proportion of all gesture strings that the participant produced, and averaged across all participants within each of the four groups. A 4×3 ANOVA with GROUP (homesigners, hearing adults, hearing children, hearing mothers) and STRING TYPE (Conflated, Mixed, Sequenced) as independent factors, and proportion of 2-component strings as the dependent factor, revealed an effect of group ($F(3,126)=8{\cdot}9$, $p<\cdot000$, partial $\eta^2=\cdot18$) and string type ($F(6,126)=16{\cdot}2$, $p<\cdot000$, partial $\eta^2=\cdot21$), and an interaction between string type and group ($F(2,126)=3{\cdot}07$, $p=\cdot008$, partial $\eta^2=\cdot13$). We describe these effects for each of the string types in the next paragraphs.

LSD pairwise comparisons revealed that CONFLATED strings were used significantly more often by homesigners ($M=0{\cdot}34$, $SE=0{\cdot}05$) than by hearing adults ($M=0{\cdot}09$, $SE=0{\cdot}03$), ($p<\cdot000$) and hearing children ($M=0{\cdot}09$, $SE=0{\cdot}04$) ($p=\cdot001$). But homesigners did not differ significantly from their mothers ($M=0{\cdot}22$, $SE=0{\cdot}06$) ($p=\cdot13$). The hearing mothers produced approximately the same number of Conflated gestures whether or not they addressed their deaf child: the four hearing mothers who described the vignettes to their deaf child produced, on average, 1·8 Conflated gesture strings; the remaining three who described the vignettes to the experimenter produced 2·0. No other differences were found among the hearing groups.

Turning next to MIXED strings, LSD pairwise comparisons revealed that homesigners ($M=0{\cdot}17$, $SE=0{\cdot}02$) used significantly more Mixed forms than all three hearing groups: hearing children ($M=0{\cdot}05$, $SE=0{\cdot}02$) ($p=\cdot002$);

hearing mothers ($M = 0{\cdot}07$, $SE = 0{\cdot}05$) ($p = {\cdot}02$); and hearing adults, who did not use any mixed forms. Mothers also used Mixed forms significantly more often than the other hearing adults ($p = {\cdot}05$); however, only two hearing mothers produced Mixed combinations, one who described the vignettes to her deaf child and one who described them to the experimenter. In contrast, all seven deaf children produced at least one instance of a Mixed form. The participants created Mixed strings by combining Conflated gestures equally often with Manner gestures (19 in total) or with Path gestures (17 in total), and less often with both Manner and Path gestures (2 in total, both produced by homesigners).

No differences were found among the groups with respect to SEQUENCED strings (homesigners, $M = 0{\cdot}06$, $SE = 0{\cdot}03$; hearing adults, $M = 0{\cdot}07$, $SE = 0{\cdot}02$; hearing children, $M = 0{\cdot}05$, $SE = 0{\cdot}02$; and hearing mothers, $M = 0{\cdot}06$, $SE = 0{\cdot}04$). There was only one instance in which a Path gesture was sandwiched between two Manner gestures, and one in which a Manner gesture was sandwiched between two Path gestures, both produced by homesigners.

### DISCUSSION

We have found that, when describing motion events, Turkish homesigners often mention both manner and path within a single gesture string, and do so significantly more often than hearing speakers do in the gestures they produce along with their speech to describe the same events. When homesigners mention both manner and path, they use two different forms: the Conflated form in which manner and path are combined within a single gesture (manner + path), and the Mixed form in which conflated gestures are combined with a segmented gesture for manner or path (e.g. manner + path – manner). Note that the Conflated form represents the motion event holistically and iconically since both the manner and the path take place simultaneously in the actual event. In contrast, the Mixed form segments out either the manner or the path and is thus a step away from iconicity. We also found that the homesigners produced the Mixed form significantly more often than hearing adults, hearing children, and their own hearing mothers, suggesting that this segmentation strategy is not directly copied from the gestural input that the children see. These patterns were presents as early as age three years in the homesigners.

## STUDY 2

One difficulty with Study 1 is that the homesigners' data came from six observation sessions, whereas all of the hearing participants were observed only once. The fact that we found no differences across the deaf children's six sessions suggests that the relatively large numbers of Mixed gesture strings that the homesigners produced were not attributable to their being observed a number of times. However, to verify these findings, we asked, in Study 2, an additional five Turkish homesigners and five Turkish hearing children to describe the same motion events only once, and examined their gestures.

In addition to replicating the patterns in Study 1, Study 2 had one other methodological goal. As described in Study 1, during the retellings of the vignettes, a still picture of the initial scene of the event, which included all objects in the event, was placed in front of the participants as a memory aid. All seven of the homesigners and eleven of the fourteen (·78) hearing children in Study 1 had referred to the pictures, and produced ·70 and ·51 of their manner and path gestures, respectively, on the pictures. In contrast, only four of the eighteen (·22) hearing adults and four of the seven (·57) hearing mothers in Study 1 referred to the pictures, producing ·09 and ·25 of their manner and path gestures, respectively, on the pictures. To determine whether the pictures had influenced the gestures that the children produced in Study 1, we modified the procedure in Study 2 and presented the vignettes without the pictures, thus making it likely that the children would produce their gestures in neutral space (i.e. at chest level) rather than on the pictures, as did the adults in Study 1.

## METHODS

### *Participants*

Five Turkish deaf children, ranging in age from 4;7 to 7;2 ($M=6;3$), were videotaped once in their homes; none of the children had participated in Study 1. As in Study 1, all of the homesigners were congenitally deaf, with bilateral hearing losses (70–90 dB), did not have cochlear implants, used hearing aids, and had no other reported cognitive or physical disabilities. None had been exposed to a conventional sign language or had contact with another deaf child or adult. None of the children had attended preschool of any sort, and all spent their days at home with their mothers.

In addition, five Turkish hearing children, ranging in age from 3;9 to 6;8 ($M=5;3$), and drawn from families of the same socioeconomic status as the deaf children, were videotaped once at home. All of the hearing children were native Turkish speakers.

### *Procedure, materials, coding*

The procedure, materials, and coding for Study 2 were identical to Study 1, with the exception that the children were not given pictures to act as a memory aid during their retellings of the vignettes.

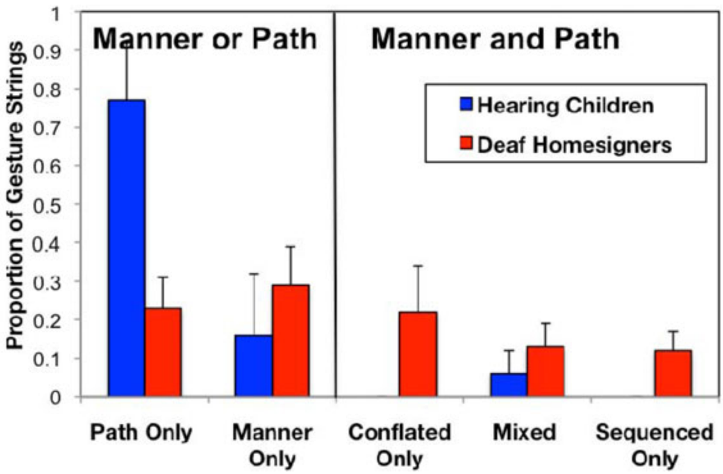



Fig. 3. (Colour online) The mean proportions of gesture strings produced by new groups of hearing children and deaf homesigners in Study 2 asked to describe the set of motion events without referring to static pictures of the objects in the events. Gesture strings contain one or more gestures and are classified according to type (Path only, Manner only, Conflated, Mixed, Sequenced). Error bars reflect standard errors.

RESULTS

One of the hearing children produced no gestures and was therefore removed from the statistical comparisons. As in Study 1, we found that the deaf and hearing children did not differ significantly in the total number of manner and path gestures they produced: 10·2 ($SE$=2·52) per participant for homesigners, 6·0 ($SE$=2·12) for hearing children ($F(1,7)=1{\cdot}52$, $p$=·26, partial $\eta^2=\cdot18$), one-way ANOVA with hearing status as the independent factor and total number of manner and path gestures as the dependent factor. In addition, the deaf and hearing children also did not differ significantly in the total number of gesture strings they produced: ($M$=7·0, $SE$=1·18) gesture strings per participant for the deaf children; ($M$=4·75, $SE$=1·38) for the hearing children ($F(1,8)=1{\cdot}55$, $p$=·25, partial $\eta^2=\cdot18$), one-way ANOVA with hearing status (deaf, hearing) as the between-subjects independent factor, and number of gesture strings as the dependent factor.

Figure 3 presents the number of gesture strings of each type produced by a participant, taken as a proportion of all gesture strings that the participant produced, and averaged across the participants within each of the two groups. As in Study 1, we focused first on 1-component strings, containing either a Path or a Manner gesture. A two-way ANOVA with hearing status (deaf, hearing) and string type (Manner only, Path only) as independent factors, and mean proportion of 1-component strings as the dependent factor, revealed no main effect of hearing status ($F(1,14)=2{\cdot}61$, $p$=·13, partial $\eta^2 = \cdot 16$), a marginal effect of string type ($F(1,14) = 4{\cdot}67$, $p = \cdot 049$, partial $\eta^2 = \cdot 25$), and a significant interaction ($F(1,14) = 6{\cdot}96$, $p = \cdot 02$. partial $\eta^2 = \cdot 33$). Further one-way ANOVA tests conducted to explore this interaction revealed that deaf children used fewer Path only strings than hearing children ($F(1,7) = 10{\cdot}05$, $p = \cdot 02$, partial $\eta^2 = \cdot 59$). There were no differences in Manner only strings between groups ($F(1,7) = 0{\cdot}48$, $p = \cdot 51$, partial $\eta^2 = \cdot 06$). Deaf children used Path only and Manner only strings equally often ($F(1,8) = 0{\cdot}21$, $p = \cdot 66$, partial $\eta^2 = \cdot 03$). Hearing children, in contrast, produced more Path only than Manner only strings ($F(1,6) = 6{\cdot}84$, $p = \cdot 04$, partial $\eta^2 = \cdot 53$).

Turning to the 2-component strings in which both manner and path were conveyed, we again calculated the number of gesture strings of each type produced by a participant, taken as a proportion of all gesture strings that the participant produced, and averaged across all participants within each of the two groups. The deaf children used all three types of 2-component strings: Conflated ($M = 0{\cdot}22$, $SE = 0{\cdot}12$), Mixed ($M = 0{\cdot}13$, $SE = 0{\cdot}07$), and Sequenced ($M = 12$, $SE = 0{\cdot}05$). The hearing children produced only one type: Mixed ($M = 0{\cdot}06$, $SE = 0{\cdot}06$). A $2 \times 3$ ANOVA with GROUP (homesigners, hearing children) and STRING TYPE (Conflated, Mixed, Sequenced) as independent factors, and proportion of 2-component strings as the dependent factor, revealed an effect of group ($F(1,21) = 5{\cdot}3$, $p = \cdot 03$, partial $\eta^2 = \cdot 20$), but no effect for string type ($F(2,21) = 0{\cdot}27$, $p = \cdot 77$, partial $\eta^2 = \cdot 03$), and no interaction ($F(2,21) = 0{\cdot}57$, $p = \cdot 58$, partial $\eta^2 = \cdot 05$). Thus, deaf children used significantly more Conflated, Mixed, and Sequenced string types than hearing children.

### DISCUSSION

Study 2 paralleled the findings of Study 1. When describing motion events, the Turkish homesigners once again conveyed both manner and path within a single string, using Mixed, Conflated, and Sequenced forms to do so. Moreover, they produced fewer 1-component strings and more 2-component strings (of all types) than hearing children did in the gestures they produced along with their spoken descriptions of the same events. In addition, because we did not provide children in Study 2 with still pictures to use as memory aids, we can be certain that the patterns found in Study 1 (and replicated in Study 2) were not influenced by having pictures present during the retellings.

## STUDY 3

Why do deaf homesigners use Conflated and Mixed forms so often to convey manner and path? As mentioned earlier, the homesigners use the manual modality as their sole means of communication. One possibility, then, is that Conflated and Mixed forms arise whenever communication is done with the hands alone. Alternatively, these forms may arise only in gestures that have been used for communication for many years and have transformed into a semi-structured system (as in the homesigners). Study 3 explores this possibility by observing the gesture strings that are produced when hearing speakers recruit the manual modality on-the-spot as their sole means of communication. We analyzed the gestures that the eighteen hearing adults produced when asked to describe the motion events without speech, comparing them first to the gestures that the same adults produced when describing the events with speech (analyzed in Study 1), and then to the gestures that the seven homesigners in Study 1 produced when describing the events.

### METHODS

#### *Participants*

The participants were the eighteen hearing adults and the seven deaf homesigners who participated in Study 1. We used the homesigners in Study 1 (as opposed to Study 2) for this comparison because we found that twelve of the eighteen (67%) hearing adults referred to the pictures when asked to describe the events without speech, producing ·43 of their manner and path gestures on the pictures. As noted earlier, homesigners in Study 1 also produced many manner and path gestures on the pictures and, in this sense, are more comparable to the hearing adults in the no-gesture condition than the homesigners in Study 2 who did not have the pictures available.

#### *Procedure, materials, coding*

The procedure, materials, and coding for Study 3 were identical to Study 1, with the exception that, in addition to describing the vignettes with speech (and spontaneous gesture), the hearing adults were asked to describe the vignettes a second time using only their hands and not their mouths. The order of retellings was always the same: first with speech, then without speech. We followed this order because putting the silent gesture condition first might have encouraged the participants to focus on gesture and, as a result, alter their subsequent co-speech gestures. The still pictures of the initial scene of each event were available to the participants when they described the events with and without speech, as was the case for all groups in Study 1.

### RESULTS

#### *Gesture without speech compared to co-speech gesture*

The hearing adults produced, on average, 6·9 ($SE$=0·2) total manner and path gestures across the vignettes when asked to gesture without speaking,

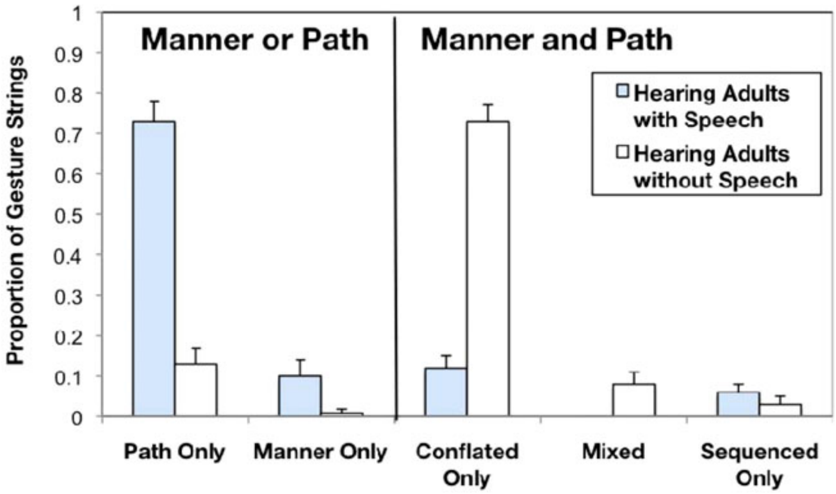



Fig. 4. (Colour online) The mean proportions of gesture strings produced by the hearing adults in Study 1 asked to describe the set of motion events using only their hands (right-hand bars), compared to the same hearing adults asked to describe the events using speech (left-hand bars; these data also appear in Figure 2); this comparison is described in Study 3. Gesture strings are classified according to type (Path only, Manner only, Conflated, Mixed, Sequenced). Error bars reflect standard errors.

a number that did not differ from the total number of manner and path gestures they produced while speaking ($F(1,17) = 1{\cdot}42$, $p = {\cdot}25$, partial $\eta^2 = {\cdot}08$), repeated measures ANOVA with group as the within-subjects factor and total number of manner and path gestures as the dependent factor. In contrast, the silent gesturers produced, on average ($M = 6{\cdot}0$, $SE = 0{\cdot}1$) gesture strings across the six vignettes, which was significantly more gesture strings than they produced while speaking ($F(1,17) = 4{\cdot}63$, $p = {\cdot}046$, partial $\eta^2 = {\cdot}21$), repeated measures ANOVA with group as the within-subjects factor and total number of gesture strings as the dependent factor.

Figure 4 displays the number of gesture strings of each type produced by a participant, taken as a proportion of all gesture strings that the participant produced, and averaged across all of the participants within each of the two conditions: hearing adults when producing gestures with speech vs. without speech. We first focused on 1-component strings and conducted a $2 \times 2$ repeated measures ANOVA on the proportion of 1-event component strings, with CONDITION (gesture with speech, gesture without speech) as the within-subjects factor and STRING TYPE (Manner only, Path only) as the between-subjects factor. We found main effects of condition ($F(1,34) = 82{\cdot}06$, $p < {\cdot}000$, partial $\eta^2 = {\cdot}71$) and string type ($F(1,34) = 77{\cdot}28$, $p < {\cdot}000$, partial $\eta^2 = {\cdot}69$), and an interaction between the two ($F(1,34) = 42{\cdot}25$,

$p < \cdot000$, partial $\eta^2 = \cdot55$). Adults used fewer 1-component strings in the gesture without speech condition than in the gesture with speech condition. The interaction reflected the fact that the difference in the proportion of Path only gesture strings adults produced with speech ($M = 0\cdot73$, $SE = 0\cdot06$) vs. without speech ($M = 0\cdot14$, SE $= 0\cdot04$) was greater ($F(1,17) = 92\cdot56$, $p < \cdot000$, partial $\eta^2 = \cdot85$) than the difference in the proportion of Manner only gesture strings they produced with speech ($M = 0\cdot11$, $SE = 0\cdot04$) vs. without speech ($M = 0\cdot009$, $SE = 0\cdot01$; $F(1,17) = 4\cdot73$, $p = \cdot04$, partial $\eta^2 = \cdot22$).

We conducted a similar analysis on the 2-component gesture strings, using a $3 \times 2$ ANOVA with CONDITION (gesture with speech, gesture without speech) as the within-subjects factor and STRING TYPE (Conflated, Mixed, Sequenced) as the between-subjects factor, and proportion of 2-component strings as the dependent factor. We found main effects of condition ($F(1,51) = 102\cdot1$, $p < \cdot000$, partial $\eta^2 = \cdot67$) and string type ($F(2,51) = 124\cdot99$, $p < \cdot000$, partial $\eta^2 = \cdot83$), and an interaction between the two ($F(2,51) = 84\cdot21$, $p < \cdot000$, partial $\eta^2 = \cdot77$).

Adults used more 2-component strings in the gesture without speech condition than in the gesture with speech condition. Further post-hoc analyses (repeated measures) revealed that the adults produced significantly more Conflated gesture strings ($F(1,17) = 157\cdot27$, $p < \cdot000$, partial $\eta^2 = \cdot90$) in the gesture without speech condition ($M = 0\cdot74$, $SE = 0\cdot04$) than in the gesture with speech condition ($M = 0\cdot07$, $SE = 0\cdot03$). No difference was found for the Sequenced strings (without speech, $M = 0\cdot03$, $SE = 0\cdot02$; with speech, $M = 0\cdot07$, $SE = 0\cdot02$) ($F(1,17) = 0\cdot85$, $p = \cdot37$, partial $\eta^2 = \cdot05$). No Mixed forms were produced in the gesture with speech condition, and very few in the gesture without speech condition ($M = 0\cdot08$, $SE = 0\cdot03$).

When required to use only their hands to communicate, the hearing adults put all of the necessary information into their hands, producing gesture strings containing both manner and path. Moreover, they tended to combine manner and path within a single gesture, producing significantly more Conflated strings when they gestured without speech than when they gestured with it.

### *Gesture without speech compared to homesign*

Next we compared the gestures that the hearing adults produced without speech to the gestures produced by the seven homesigners who participated in Study 1. Figure 5 displays the silent hearing adults' gesture strings in relation to the homesigners' gesture strings. We first focused on the proportion of 1-component strings the two groups produced. A $2 \times 2$ ANOVA with GROUP (homesigners, silent gesturers) and STRING TYPE (Path only, Manner only) as independent factors, and proportion of 1-component gesture strings as the dependent factor, revealed main effects of group

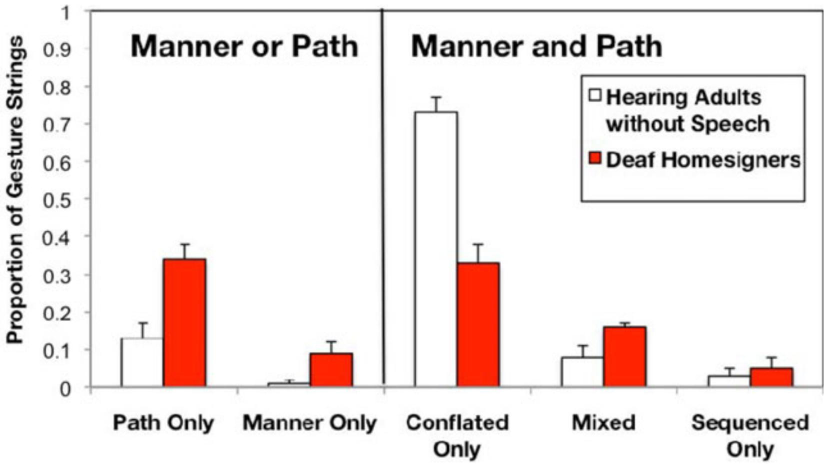



Fig. 5. (Colour online) The mean proportions of gesture strings produced by the hearing adults in Study 1 when asked to describe the set of motion events using only their hands (left-hand bars; these data also appear in Figure 4), compared to the deaf homesigners in Study 1 who relied exclusively on gesture to describe the events (right-hand bars; these data also appear in Figure 2); this comparison is described in Study 3. Gesture strings are classified according to type (Path only, Manner only, Conflated, Mixed, Sequenced). Error bars reflect standard errors.

($F(1,46)=14{\cdot}41$, $p<{\cdot}000$, partial $\eta^2={\cdot}24$) and string type ($F(1,46)=25{\cdot}23$, $p<{\cdot}000$, partial $\eta^2={\cdot}35$), but no interaction between the two ($F(1,46)=2{\cdot}56$, $p={\cdot}12$, partial $\eta^2={\cdot}05$). The homesigners conveyed more 1-component events ($M=0{\cdot}43$, $SE=0{\cdot}04$) than the silent gesturers ($M=0{\cdot}15$, $SE=0{\cdot}02$) and thus did not appear to be as sensitive to the importance of conveying BOTH manner and path information as the silent gesturers. Both groups produced more Path only than Manner only 1-component strings.

A similar analysis was conducted on the 2-component gesture strings. We conducted a 3×2 ANOVA, with GROUP (homesigners, silent gesturers) and STRING TYPE (Conflated, Mixed and Sequenced) as independent factors, and the proportion of 2-component strings as the dependent factor. We found main effects of group ($F(1,69)=8{\cdot}69$, $p={\cdot}004$, partial $\eta^2={\cdot}11$) and string type ($F(2,69)=86{\cdot}59$, $p<{\cdot}000$, partial $\eta^2={\cdot}72$), and an interaction between the two ($F(2,69)=21{\cdot}77$, $p<{\cdot}000$, partial $\eta^2={\cdot}39$). Post-hoc tests revealed that the silent gesturers used Conflated gesture strings significantly more often than homesigners ($F(1,23)=27{\cdot}13$, $p<{\cdot}000$, partial $\eta^2={\cdot}54$). In contrast, homesigners used Mixed gesture strings marginally more often than the silent gesturers ($F(1,23)=3{\cdot}6$, $p={\cdot}06$, partial $\eta^2={\cdot}14$). There were no differences between the groups in Sequenced strings ($F(1,23)=0{\cdot}36$, $p={\cdot}55$, partial $\eta^2={\cdot}02$).

## DISCUSSION

We have found that the gestures Turkish homesigners produce to describe spontaneous motion events look different from a hearing person's gestures, even when those gestures are produced without speech. When told to use gesture and not speech to describe motion events, Turkish hearing adults mentioned both manner and path within a single gesture string, and did so even more often than the deaf homesigners. Importantly, however, the hearing adults did NOT use the same forms to convey these pieces of information as the deaf homesigners – the hearing adults exclusively conflated manner and path into one gesture; the deaf children produced these Conflated forms, but also produced forms in which manner and/or path was segmented out into a separate gesture produced along with the Conflated gesture (i.e. the Mixed form). Thus, although the need to convey everything in the manual modality is likely to have encouraged both homesigners and silent gesturers to express manner and path within a single gesture string, this need did not dictate the form of the gestures: Silent gesturers relied exclusively on the Conflated form, whereas homesigners used the Mixed form as well as the Conflated form.

## GENERAL DISCUSSION

Turkish homesigners, who are developing their communication systems without the benefit of a conventional language model, nevertheless express two basic elements of motion events – manner and path – in their gestures. These observations corroborate those of Zheng and Goldin-Meadow (2002), who found that Chinese and American homesigners produced manner and path gestures when describing motion events in naturalistic interactions. However, our findings take the phenomenon several steps further by analyzing how these two elements are combined within a gesture string and comparing them to the gestures produced by hearing children and adults in the same cultural community.

### *Emphasis on both manner and path*

We found, first, that in approximately half of their gesture strings (see the right-most bars in Figures 2 and 3), homesigners mention BOTH manner and path within a single gesture string; moreover, they display this pattern as early as age three, with no further developmental change through age five. This finding is itself interesting given that many studies of preschool-aged hearing children describing motion events in speech (e.g. Allen *et al.*, 2007; Özyürek *et al.*, 2008; Papafragou & Selimis, 2010) have found that children of this age tend to mention only one component of the motion event (typically the path, but see Bunger, Trueswell & Papafragou, 2012, and Papafragou, Massey & Gleitman, 2006, for the effect of typology of the language), rather than mentioning both manner and path. We speculate that the homesigners' inclusion of both manner and path at these early stages may reflect the influence of modality. Recent work by Sümer, Zwitserlood, Perniss, and Özyürek (2013) has shown that deaf children learning Turkish Sign Language from their deaf parents frequently include both manner and path in their sign sentences beginning at age four, and do so significantly more often than age-matched hearing peers who are learning Turkish. It may be easier to convey both manner and path in the manual modality, which supports an iconic mapping between form and meaning.

Second, we found that when homesigners mention manner and path within a single gesture sentence, they often conflate the two components into one gesture (manner + path). Note that the conflated form portrays the motion event holistically. But homesigners go beyond holistic representation when they combine conflated gestures with a segmented gesture for manner or path—that is, when they produce the mixed form. Homesigners produce this form significantly more often than hearing adults, hearing children, and their own hearing mothers. Even when hearing adults are called upon to use only their hands to communicate, they rarely produce the mixed form, preferring instead to use the conflated form. Importantly, in another study of four Turkish homesigners (all of whom participated in Study 1) interacting with their hearing mothers in unscripted play sessions at home, we also found evidence of the mixed form; and, just as importantly, we found that the mothers did not produce the mixed form during these interactions (Goldin-Meadow, Namboodiripad, Mylander, Özyürek & Sancar, in press). We consider the implications of these findings on mixed and conflated gesture forms in the next sections.

### *The mixed form*

The mixed form, which combines a conflated form with at least one segmented form, is interesting in large part because it represents a step towards segmentation and combination. Segmentation may not be difficult for Turkish homesigners to introduce into their action gestures simply because they routinely see hearing Turkish speakers produce decomposed gestures along with their descriptions of motion events (see Figure 1, and Kita & Özyürek, 2003; Özyürek *et al.*, 2005, 2008). These segmented gestures could have served as a model for the homesigners' path alone and manner alone gestures (and even for their few sequenced gesture strings), but segmented gestures are not a good model for the homesigners' mixed form, which involves combining a segmented form with a conflated form. If the homesigners were merely taking their own conflated form and combining it with the most frequent segmented form they see (i.e. path), they should produce more mixed forms with path in the segmented slot (i.e. manner + path – path) than with manner in the segmented slot (manner + path – manner), and they do not.

But perhaps homesigners do not need to see segmentation and combination in order to use it; segmentation and combination offer a number of communicative benefits and thus may be easy for children to discover on their own, possibly facilitated by interaction sequences with their communicative partners. First, segmenting out one component of a simultaneous event allows the language user to combine that component with other elements, thus leading to new combinatorial possibilities not imaginable with the conflated form alone. Second, segmenting out one component allows the language user to focus on one aspect of the event in a topic–focus construction. Importantly, this type of focusing is not possible with conflated or single-component constructions. As a result, the pressure to highlight certain components of an event might have led homesigners to produce mixed forms, and to do so more often than hearing speakers. Hearing speakers can use speech to accomplish this type of focusing and thus need not rely on co-speech gesture for this function. Interestingly, however, even when asked to use gesture on its own without speech, hearing speakers still do not use the mixed form as often as homesigners do – silent gesturers primarily produce conflated forms.

The mixed form represents a small step towards segmentation and combination within action gestures and it is interesting that the silent gesturers do not take it, particularly since they do exhibit other linguistic properties in the gestures they create on the spot (e.g. Goldin-Meadow *et al.*, 2008). This finding suggests that some of the properties found in the homesigners' gestures may require time to develop. Note, however, that the onset of action segmentation and combination must have taken place prior to the onset of our study since the mixed form was present in the homesigners' earliest sessions and did not change in frequency over the six sessions (a fact that also suggests the form was not an adaptive response to the task demands per se). Generating a communication system over a period of years thus appears to be a process that is distinct from inventing gestures on the spot – although there are, of course, many differences between homesigners and silent gesturers (e.g. age, cognitive maturity), making time span only of many potential factors that could account for the difference between the groups.

Turning to a longer time span (development over generations rather than over childhood), we note that our findings cohere well with Senghas *et al.* (2004), who studied changes in how action is segmented in the newly evolving Nicaraguan Sign Language (NSL). Nicaraguan Sign Language was born thirty years ago when deaf children were brought together for the first time in an educational setting but with no sign language instruction.

Every year, new students entered the school and the peers developed a common, rule-governed sign language (Kegl, Senghas & Coppola, 1999; Senghas & Coppola, 2001). Senghas *et al.* (2004) explored action segmentation across the first three cohorts of NSL, and found that each new cohort introduced more manner and path segmentation and sequencing (i.e. our sequenced forms) than the previous cohort. Interestingly, the co-speech gestures of the hearing community within which these individuals live displayed no segmentation; the gesturers conflate manner and path into a single gesture.

Our homesigners live in a very different cultural context, and are much younger, than the Nicaraguan gesturers and the NSL signers described by Senghas *et al.* (2004), and we found that they developed a form not previously identified in either the Nicaraguan gesturers or the Nicaraguan signers. Prompted by the discovery of the mixed gesture string in our data, Senghas, Özyürek, and Goldin-Meadow (2010, 2013) reanalyzed the original Nicaraguan data reported in Senghas *et al.* (2004) to determine whether and how often the mixed form was used. They discovered that all three cohorts of NSL signers, as well as the Spanish-speaking gesturers, produced the mixed form. However, the mixed form was the dominant response only for the first cohort of NSL signers, the transitional group between the speakers (whose gestures were not linguistically structured) and the second and third cohorts (whose signs were taking on more and more linguistic properties; cf. Senghas & Coppola, 2001). However, the fact that the first cohort of NSL signers, who are older and therefore have spent more time in the emerging language community than the second and third cohorts, use segmentation and combination LESS often than any other group of signers (Senghas *et al.*, 2004) suggests that time on task does not fully account for the rise of segmentation and combination. The age of the language creator/learner may play an important role as well in order for full segmentation and combinatorial possibilities to emerge (see Senghas, 2003).

We speculate on the basis of our Turkish homesign data that the original Nicaraguan homesigners who came together and created NSL were already producing some mixed sentences (containing conflated and segmented forms). The mixed gesture strings became the dominant form in the first cohort, setting the stage for the sequenced strings (containing only segmented forms) that have come to dominate the signs of the second and third cohorts. The fact that our homesigners produced the mixed form suggests that children who have not had contact with either an accessible conventional language model or other deaf individuals (i.e. children fashioning a communication system without a linguistic community) can introduce action segmentation and combination into their gestures. These children have thus taken a step towards segmentation not found in gesturers. However, the fact that our homesigners produced very few of the sequenced manner-path gesture sentences found primarily in the second or third cohorts of NSL suggests that they have not yet achieved the fully segmented and sequenced forms found in signers. Taken together, the findings suggest that the mixed form may constitute an early step in the emergence of manual communication systems – one that retains an element of iconicity and holistic representation (i.e. Conflated strings) while at the same time allowing the signer to single out and focus on a piece of the event. We might therefore expect to find the mixed form in homesigns developed around the globe, perhaps even in individuals who use homesign into adulthood.

### *The conflated form*

The other dominant form in the homesigners was the conflated manner + path gesture used on its own. The homesigners used this form significantly more often than the hearing adults and hearing children did in their co-speech gestures. The hearing adults and children preferred instead to produce individual path gestures, as we would expect given the syntactic structures of Turkish, their spoken language (Kita & Özyürek, 2003; Özyürek *et al.*, 2005, 2008). Given that the gestures speakers produce along with their descriptions of motion events tend to parallel the speech they accompany (Özyürek *et al.*, 2005, 2008), it is not surprising that the Turkish hearing speakers in Study 1 produced path-only speech descriptions in more than half of their responses. The fact that Turkish speakers used the conflated form in the silent condition, but relied primarily on decomposed gestures (path only and manner only, with a preference for the first) during speech also provides further evidence for the claim that gestures are shaped not directly by the imagery of the event but by the habitual linguistic packaging of event components (Kita *et al.*, 2007).

The homesigners neither understood nor produced spoken Turkish. Their gestures were thus not constrained by Turkish and were free to assume whatever form the child chose. Their choice of the conflated form might have been motivated by iconicity, as this form represents the actual event more closely than the segmented forms. Another possibility is that the homesigners learned the conflated form from their hearing mothers. Recall that the hearing mothers used the conflated form more often than the other hearing adults and almost as often as their deaf children, possibly to be able to communicate with their child in the most iconic way possible. Moreover, when Turkish hearing adults are asked to communicate using only their hands, they increase the number of conflated gestures they produce. If the hearing mothers had addressed their children using gesture without speech, they might have produced an even greater number of conflated gestures. In this regard, it is important to note that the mothers of Turkish deaf homesigners rarely produce gestures without speech when addressing their children in spontaneous interactions (Flaherty & Goldin-Meadow, 2010). Thus, although it is possible that the deaf children learned the conflated form from their hearing mothers, we cannot rule out the possibility that the hearing mothers produced their conflated gestures in response to their children's conflated gestures.

*Gesture with speech and without it in the input to children*

Our findings suggest that, aside from seeing their hearing mothers produce a slightly larger number of conflated gestures than hearing Turkish speakers typically produce, the Turkish deaf children in our study are likely to have been exposed to the same types of manner and path gestures as Turkish hearing children in their community (see also Goldin-Meadow & Saltzman, 2000, who found few frequency differences between the gestures hearing mothers produce with their deaf vs. hearing children in China and the US). The difference between Turkish hearing and deaf children is that hearing children interpret the gestures they see in the context of speech – most severe to profoundly deaf children are unable to make effective use of the speech that surrounds them, even when provided with a hearing aid. This difference seems to affect how children use their gestural input. Hearing children integrate their gestures with the speech they hear, producing gestures comparable to those produced by hearing adults. In contrast, deaf children transform the gestures they see into a homesign system characterized by language-like structure; they use, for example, the mixed form found only rarely in any of the hearing speakers' gestures.

Our findings are the first to explicitly compare how deaf and hearing children respond to the gestures they see and, as such, they make it clear that there is no one 'child' gesture pattern. The findings also underscore the fact that gesture is part of an integrated gesture–speech system for hearing children, but must serve all of the functions of language for the deaf children and, as a result, needs transformation.

*Gesture with speech and without it in the output*

We found that the deaf children's gestures look different from a hearing person's gestures even when those gestures are produced without speech. When hearing adults are told to use only gesture to describe motion events, they (like the deaf children) find it essential to mention both manner and path within a single gesture sentence, presumably in response to the need to convey all of the relevant information in the manual modality. Importantly, however, this pressure does not dictate the form of the resulting gesture string – deaf homesigners, in addition to using the conflated form on its own, often add decomposed segments to the conflated gesture, thus creating the mixed form; hearing adults prefer to use the conflated form on its own.

One additional point is worth highlighting in this regard. Segmenting and sequencing the action components of an event appears to be less robust in communication than segmenting and sequencing an entity and the event in which it is involved. Hearing adults asked to communicate using only their hands routinely produce segmented and sequenced gestures representing the figure and path of an event (e.g. circle followed by path; Goldin-Meadow *et al.*, 1996, 2008), as do homesigners. In contrast, neither group produces many segmented and sequenced gestures representing the manner and path of an event within a single gesture string (e.g. roll gesture, followed by down gesture). However, by producing a sizeable number of mixed forms (e.g. roll + down, followed by roll or down), the homesigners have taken a step towards action segmentation and combination that silent hearing adults do not take (see also Özçalişkan & Goldin-Meadow, 2013).

## CONCLUSION

In sum, we have found that homesigners, who do not have access to a language model that they can process, introduce action segmentation and combination into their gestural communication systems even though the manual modality lends itself to holistic representation (e.g. rolling down a hill is a single act that is easily represented using a single gesture incorporating both manner and path). Homesigners do conflate manner and path in the same gesture, but they also combine those conflated gestures with segmented gestures for manner and/or path (the mixed gesture string), thus taking the first step toward a segmented representational form. In contrast, hearing speakers in the same community rarely combine conflated gestures with segmented gestures into a mixed string and, in fact, produce large numbers of conflated gestures only when they are forced to use gesture to communicate. Thus the mixed form may be indexing an intermediate stage in the development of manual language systems, one that bridges the transition from conflated forms that have no segmentation to sequenced forms that are fully segmented.

The segmentation patterns we observe in the homesigners' mixed form are consistent with patterns found in deaf children learning sign languages from their deaf parents. Deaf children have been found to display a preference for linear sequencing even in situations where adult signers use simultaneous constructions. For example, Meier (1987) found that children learning American Sign Language initially break complex verb expressions down into sequential morphemes, despite the fact that adult ASL signers produce these verb elements within a single simultaneous movement (see also Supalla, 1982; Newport, 1981). Our homesigners display similar tendencies even though they are not exposed to a conventional language model.

Our results are also in line with recent experimental and simulation studies of language emergence. The conflated representations of manner and path in homesign systems reveal an initial bias for iconic and holistic representation, corroborating claims about iconicity as the base out of which linguistic structures might have emerged (Garrod, 2007; Gasser, 2004; Theisen, Oberlander & Kirby, 2010) and as a feature that can still be found in modern-day languages, signed (Perniss, Thompson & Vigliocco, 2010) and spoken (Shintel, Nusbaum & Okrent, 2006). At the same time, our results underscore the fact that homesigning children are able to pull away from iconicity (even if not totally), suggesting that children may be predisposed to prefer communication systems characterized by segmentation and combination. However, the fact that homesigners do not display the sequencing found in later cohorts of Nicaraguan Sign Language makes it clear that children cannot do it all, and that other forces (e.g. having a community within which the language is socially shared; transmitting the language from one generation to the next; Christiansen & Kirby, 2003; Fay, Garrod, Roberts & Swoboda, 2010; Goldin-Meadow, 2010; Senghas *et al.*, 2010, 2013) must have collaborated to make human language what it is.

## REFERENCES


Allen, S., Özyürek, A., Kita, S., Brown, A., Furman, R., Ishizuka, T. & Fujii, M. (2007). Language-specific and universal influences in children's packaging of manner and path: A comparison of English, Japanese, and Turkish. *Cognition* **102**, 16–48.

Bunger, A., Trueswell, J. C. & Papafragou, A. (2012). The relation between event apprehension and utterance formulation in children: Evidence from linguistic omissions. *Cognition* **122**, 135–49.

Choi, S. & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition* **41**, 83–121.

Christiansen, M. & Kirby, S. (2003). *Language evolution*. Oxford: Oxford University Press.

Coppola, M. & Newport, E. L. (2005). Grammatical 'Subjects' in home sign: Abstract linguistic structure in adult primary gesture systems without linguistic input. *Proceedings of the National Academy of Sciences* **102**, 19249–53.

Fay, N., Garrod, S., Roberts, L. & Swoboda, N. (2010). The interactive evolution of human communication systems. *Cognitive Science* **34**(3), 351–86.

Feldman, H., Goldin-Meadow, S. & Gleitman, L. (1978). Beyond Herodotus: The creation of language by linguistically deprived deaf children. In A. Lock (ed.), *Action, symbol, and gesture: The emergence of language*, 351–414. New York: Academic Press.

Flaherty, M. & Goldin-Meadow, S. (2010). Does input matter? Gesture and homesign in Nicaragua, China, Turkey, and the USA. In A. D. M. Smith, M. Schouwstra, B. de Boer & K. Smith (eds), *Proceedings of the Eighth Evolution of Language Conference*, 403–4. Singapore: World Scientific Publishing Co.

Franklin, A., Giannakidou, A. & Goldin-Meadow, S. (2011). Negation and structure building in a home sign system. *Cognition* **118**(3), 398–416.

Garrod, S. (2007). Foundations of representation: Where might graphical symbol systems come from? *Cognitive Science* **31**(6), 961–87.
Gasser, M. (2004). Origins of arbitrariness in language. In K. Forbes, D. Gentner & T. Regier (eds), *Proceedings of the Annual Conference of the Cognitive Science Society*, 26, 434–40. Mahwah, NJ: Lawrence Earlbaum.
Gershkoff-Stowe, L. & Goldin-Meadow, S. (2002). Is there a natural order for expressing semantic relations? *Cognitive Psychology* **45**(3), 375–412.
Goldin-Meadow, S. (2003). *The resilience of language: What gesture creation in deaf children can tell us about how all children learn language*. New York: Psychology Press.
Goldin-Meadow, S. (2010). Widening the lens on language learning: Language creation in deaf children and adults in Nicaragua. *Human Development* **53**, 235–312.
Goldin-Meadow, S., Butcher, C., Mylander, C. & Dodge, M. (1994). Nouns and verbs in a self-styled gesture system: What's in a name? *Cognitive Psychology* **27**, 259–319.
Goldin-Meadow, S. & Feldman, H. (1977). The development of language-like communication without a language model. *Science* **197**, 401–3.
Goldin-Meadow, S., McNeill, D. & Singleton, J. (1996). Silence is liberating: Removing the handcuffs on grammatical expression in the manual modality. *Psychological Review* **103**, 34–55.
Goldin-Meadow, S. & Mylander, C. (1984). Gestural communication in deaf children: The effects and non-effects of parental input on early language development. *Monographs of the Society for Research in Child Development* **49**, 1–121.
Goldin-Meadow, S., Mylander, C. & Butcher, C. (1995). The resilience of combinatorial structure at the word level: Morphology in self-styled gesture systems. *Cognition* **56**, 195–262.
Goldin-Meadow, S., Mylander, C. & Franklin, A. (2007). How children make language out of gesture: Morphological structure in gesture systems developed by American and Chinese deaf children. *Cognitive Psychology* **55**, 87–135.
Goldin-Meadow, S., Namboodiripad, S., Mylander, C., Özyürek, A. & Sancar, B. (in press). How children create language out of gesture: The effect of co-speech gesture on homesign in American and Turkish deaf children. *Journal of Cognition and Development*.
Goldin-Meadow, S., So, W.-C., Özyürek, A. & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences* **105**, 9163–8.
Goldin-Meadow, S. & Saltzman, J. (2000). The cultural bounds of maternal accommodation: How Chinese and American mothers communicate with deaf and hearing children. *Psychological Science* **11**, 311–8.
Hockett, C. F. (1960). Origins of Speech. *Scientific American* **203**, 88–96.
Hunsicker, D. & Goldin-Meadow, S. (2012). Hierarchical structure in a self-created communication system: Building nominal constituents in homesign. *Language* **88**(4), 732–63.
Iverson, J. M., Capirci, O., Longobardi, E. & Caselli, M. C. (1999). Gesturing in mother–child interactions. *Cognitive Development* **14**, 57–75.
Kegl, J., Senghas, A. & Coppola, M. (1999). Creation through contact: Sign language emergence and sign language change in Nicaragua. In M. DeGraff (ed.), *Language creation and language change: Creolization, diachrony, and development*, 179–237). Cambridge, MA: MIT Press.
Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (ed.), *Relationship of verbal and nonverbal communication*, 207–28. The Hague: Mouton.
Kita, S., van der Hulst, H. & van Gijn, I. (1998). Movement phases in signs and co-speech gestures, and their transcription human coders. In I. Wachsmuth & M. Fröhlich (eds), *Gesture and sign language in human–computer interaction: lecture notes in artificial intelligence*, 23–35. Berlin: Springer.
Kita, S. & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language* **48**, 16–32.

Kita, S., Özyürek, A., Allen, S., Brown, A., Furman, R. & Ishizuka, T. (2007). Relations between syntactic encoding and co-speech gestures: Implications for a model of speech and gesture production. *Journal of Language and Cognitive Processes* **22**, 1212–36.

McNeill, D. (1992). *Hand and mind*. Chicago: University of Chicago Press.

Meier, R. P. (1987). Elicited imitation of verb agreement in American Sign Language: iconically or morphologically determined? *Journal of Memory and Language* **26**, 362–76.

Newport, E. (1981). Constraints on structure: Evidence from American Sign Language and language learning. In W. A. Collins (ed.), *Aspects of the development of competence: Minnesota Symposia on Child Psychology*, 93–124. Hillsdale, NJ: Lawrence Erlbaum Associates.

Özçalişkan, S. & Goldin-Meadow, S. (2005). Gesture is at the cutting edge of language development. *Cognition* **96**, 101–13.

Özçalişkan, S. & Goldin-Meadow, S. (2013). How speaking shapes the native language of gesture in describing motion. Paper presented in the special session on space and directionality in language, Berkeley Linguistics Society, Berkeley, CA.

Özçalişkan, S. & Slobin, D. I. (1999). Learning how to search for the frog: Expression of manner of motion in English, Spanish, and Turkish. In A. Greenhill, H. Littlefield & C. Tano (eds), *Proceedings of the 23rd Annual Boston University Conference on Language Development*, 541–52. Boston, MA: Cascadilla Press.

Özyürek, A., Kita, S. & Allen, S. (2001). Tomato Man movies: Stimulus kit designed to elicit manner, path and causal constructions in motion events with regard to speech and gestures [Videotape]. Nijmegen, the Netherlands: Max Planck Institute for Psycholinguistics, Language and Cognition Group.

Özyürek, A., Kita, S., Allen, S., Brown, A., Furman, R. & Ishizuka, T. (2008). Development of cross-linguistic variation in speech and gesture: Motion events in English and Turkish. *Developmental Psychology* **44**, 1040–54.

Özyürek, A., Kita, S., Allen, S., Furman, R. & Brown, A. (2005). How does linguistic framing of events influence co-speech gestures? Insights from cross-linguistic variations and similarities. *Gesture* **5**, 215–37.

Papafragou, A., Massey, C. & Gleitman, L. (2006). When English proposes what Greek presupposes: The cross-linguistic encoding of motion events. *Cognition* **98**, B75–87.

Papafragou, A. & Selimis, S. (2010). Lexical and structural biases in the acquisition of motion verbs. *Language Learning and Development* **6**, 87–115.

Perniss, P., Thompson, R. & Vigliocco, G. (2010). The role of iconicity in signed and spoken languages. *Frontiers in Language Sciences* **1**, E227.

Sandler, W., Meir, I., Padden, C. & Aronoff, M. (2005). The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences* **102**, 2661–5.

Senghas, A. (2003). Intergenerational influence and ontogenetic development in the emergence of spatial grammar in Nicaraguan Sign Language. *Cognitive Development* **18**, 511–31.

Senghas, A. & Coppola, M. (2001). Children creating language: how Nicaraguan Sign Language acquired a spatial grammar. *Psychological Science* **12**, 323–8.

Senghas, A., Kita, S. & Özyürek, A. (2004). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science* **305**, 1779–82.

Senghas, A., Özyürek, A. & Goldin-Meadow, S. (2010). Evolution of segmentation and sequencing: Evidence from homesign and Nicaraguan Sign Language. In A. D. M. Smith, M. Schouwstra, B. de Boer & K. Smith (eds), *Proceedings of the Eighth Evolution of Language Conference*, 279–89. Singapore: World Scientific Publishing Co.

Senghas, A., Özyürek, A. & Goldin-Meadow, S. (2013). Homesign as a way-station between co-speech gesture and sign language: The evolution of segmenting and sequencing. In R. Botha & M. Everaert (eds), *The evolutionary emergence of language: Evidence and inference*, 62–77. Oxford: Oxford University Press.

Shatz, M. (1982). On mechanisms of language acquisition: Can features of the communicative environment account for development? In E. Wanner & L. Gelitman (eds), *Language acquisition: The state of the art*, 102–27. New York: Cambridge University Press.

Shintel, H., Nusbaum, H. C. & Okrent, A. (2006). Analog acoustic expression in speech communication. *Journal of Memory and Language* **55**(2), 167–77.

Slobin, D. & Hoiting, N. (1994). Reference to movement in spoken and signed languages: Typological considerations. *Proceedings of the 20th Berkeley Linguistics Society*, 487–505. Online: <http://elanguage.net/journals/bls/index>

Sümer, B., Zwitserlood, I., Perniss, P. & Özyürek, A. (2013). Revisiting modality effects in children's acquisition of spatial language: Insights from Turkish and Turkish Sign Language. Paper presented at Theoretical Issues in Sign Language Research Conference, University College London.

Supalla, T. (1982). *Structure and acquisition of verbs of motion and location in American Sign Language* (Unpublished PhD dissertation). University of California, San Diego.

Supalla, T. (1990). Serial verbs of motion in ASL. In S. Fischer & P. Suple (eds), *Theoretical issues in sign language research, Vol. 1: Linguistics*, 127–52. Chicago: University of Chicago Press.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (ed.), *Language typology and syntactic description, Vol. III: Grammatical categories and the lexicon*, 57–149. Cambridge: Cambridge University Press.

Theisen, C., Oberlander, J. & Kirby, S. (2010). Systematicity and arbitrariness in novel communication systems. *Interaction Studies* **11**(1), 14–32.

Zheng, M. & Goldin-Meadow, S. (2002). Thought before language: How deaf and hearing children express motion events across cultures. *Cognition* **85**, 145–75.