

ARTICLE

What the development of gesture with and without speech can tell us about the effect of language on thought

Şeyda Özçalışkan¹, Ché Lucero^{2,3} and Susan Goldin-Meadow⁴

¹Department of Psychology, Georgia State University, Atlanta, GA, USA; ²Department of Psychology, Cornell University, Ithaca, NY, USA; ³Department of Psychology, SPARK Neuro, Inc, New York, NY, USA; ⁴Department of Psychology, University of Chicago, Chicago, IL, USA

Corresponding author: Şeyda Özçalışkan; Email: seyda@gsu.edu

(Received 17 January 2023; Revised 25 April 2023; Accepted 08 June 2023)

Abstract

Adults display cross-linguistic variability in their speech in how they package and order semantic elements of a motion event. These differences can also be found in speakers' *co-speech gestures* (gesturing with speech), but not in their *silent gestures* (gesturing without speech). Here, we examine when in development children show the differences between co-speech gesture and silent gesture found in adults. We studied speech and gestures produced by 100 children learning English or Turkish ($n = 50/\text{language}$) – equally divided into 5 age-groups: 3–4, 5–6, 7–8, 9–10, and 11–12 years. Children were asked to describe three-dimensional spatial event scenes (e.g., a figure crawling across carpet) first with speech and then without speech using their hands. We focused on physical motion events that elicit, in adults, cross-linguistic differences in co-speech gesture and cross-linguistic similarities in silent gesture. We found the adult pattern even in the youngest children: (1) Language shaped co-speech gesture beginning at age 3 years, showing an early effect of language on thinking for speaking (as measured by gestures that occur during the speech act). (2) Language did *not* affect silent gesture at any age, highlighting early limits on the effects language has on thinking and revealing a language of gesture that shows similarities across languages.

Keywords: motion events; co-speech gesture; silent gesture; child language; crosslinguistic differences

Introduction

Languages display variability in how they express experiential domains. These cross-linguistic differences not only influence how speakers *talk* about these domains but also how they *think* about the domains (e.g., Choi & Bowerman, 1991; Evans & Levinson, 2009). This view, called the strong Sapir–Whorf hypothesis (Sapir, 1961; Whorf, 1956), posits an extended effect of language on cognition, an effect that is present not only when speaking but also when not speaking (Lucy, 1992a, 1992b). A



weaker version of the Sapir–Whorf hypothesis holds that language has a more limited and transient effect on cognition. This view is reflected in Slobin’s (1996, 2004) thinking-for-speaking account, which proposes that language influences cognition during online production of speech, but not beyond speech production.

In this article, we use the gestures people produce to describe an event to explore the impact of language on thinking. We know from previous work that cross-linguistic differences in how an event is described in speech can also be found in the gestures that accompany speech (co-speech gesture, Kita & Özyürek, 2003; Özçalışkan et al., 2016a). These findings suggest an effect of language on thinking-for-speaking that goes beyond the words in a communicative act. Here, we explore *gesture with speech* in children aged 3 to 12 years to determine when language first has an effect on thinking for speaking during the act of speaking but measured by gesture.

We also know from previous work that cross-linguistic differences found in co-speech gesture do *not* appear when people are asked to describe the same event in gesture without speech (silent gesture, Özçalışkan et al., 2016b, 2018). These findings point to limits, even within communication, on the strong form of the Sapir–Whorf hypothesis. We therefore observe *gesture without speech* in the same children to determine whether limits on how language affects thinking appear during childhood and, if so, when.

We investigated co-speech gesture and silent gesture in descriptions of motion events. A motion event consists of four key elements (Talmy, 1985, 2000): a *figure* that moves (e.g., woman or boy), a *ground* anchoring the figure’s movement (e.g., house or bridge), a *path* marking the direction of the figure’s movement (e.g., into or across), and a *manner* specifying the pattern of the figure’s motion (e.g., run or crawl).

Speakers of different languages largely follow a binary cross-linguistic split in how they *package* manner and path elements of a motion event: satellite-framed languages (e.g., German, English, and Polish) and verb-framed languages (e.g., Spanish, Turkish, and Korean; Cardini, 2010; Choi & Lantolf, 2008; Chui, 2009, 2012; Gennari et al., 2002; Ibarretxe-Antuñano, 2004; Lewandowski & Özçalışkan, 2021; Naigles et al., 1998; Özçalışkan & Slobin, 1999, 2003; Tütüncü et al., 2023). Speakers of satellite-framed languages, such as English, prefer a *conflated* packaging strategy, placing manner information in the verb and path information in a satellite to the verb (preposition or particle) within the same clause (e.g., *girl RUNS INTO* [manner path] *the house*). In contrast, speakers of verb-framed languages, such as Turkish, rely on a *separated* packaging strategy, typically placing path information in the verb and manner information in an additional subordinate clause (e.g., *kız eve GİRER* [path] *KOŞARAK* [manner] ‘girl house-to ENTER [path] RUNNING [manner]’; Allen et al., 2007; Özçalışkan & Slobin, 1999; Slobin, 2004). Adult Turkish speakers also often convey only path, omitting manner from their descriptions of motion in speech (e.g., Özçalışkan, 2009, 2016).

Speakers of Turkish and English also follow a two-way split in their *ordering* of semantic elements in descriptions of motion events in speech. Adult speakers of English use a Figure–MOTION–Ground order, which locates motion in the middle position in the clause (e.g., *she* [figure] *RUNS* [motion] *into house* [ground] – consistent with the canonical subject–verb–object (SVO) order of the language. Adult speakers of Turkish, on the other hand, use a Figure–Ground–MOTION order, locating the key motion element at the final position in the clause (e.g., *eve GİRER* ‘she [figure] house-to [ground] ENTER [motion]’) – consistent with the canonical subject–object–verb (SOV) order of Turkish (Özçalışkan et al., 2016b, 2018).

Effects of language on thinking-for-speaking that go beyond words: Co-speech gesture

Adult speakers of Turkish and English display the same cross-linguistic patterns in their co-speech gestures. English speakers express manner and path components of motion simultaneously within a single gesture (e.g., rotating the hand as it moves down, manner+path), and Turkish speakers express each component in separate gestures (e.g., rotating the hand and then moving the hand down, manner–path; Kita & Özyürek, 2003; Özçalışkan et al., 2016a, 2018). Similarly, English speakers prefer to place the motion gesture in the middle of a gesture string; Turkish speakers prefer to place the motion gesture at the end of the gesture string (Goldin-Meadow et al., 2008; Özçalışkan et al., 2016b, 2018; Tütüncü et al., 2023; see Methods section, for further details and examples of packaging and ordering of motion elements). These co-speech gesture patterns indicate an effect of language on thinking that goes beyond the words in a communicative act.

When do these language-specific patterns emerge in speech and gesture? Child learners of English or Turkish begin to follow language-specific patterns of motion expression in *speech* at a relatively young age (Allen et al., 2007; Hickmann et al., 2009; Hohenstein, 2005; Özçalışkan & Slobin, 1999). Beginning at age 3–4 years, English learners use conflated descriptions (e.g., she *runs into* house), while Turkish learners rely on separated descriptions that typically convey only path information (e.g., Eve *girdi* ‘She entered the house’; Allen et al., 2007; Özçalışkan, 2009; Özçalışkan & Slobin, 1999). Children learning English or Turkish also show early sensitivity to the canonical ordering of semantic elements in their speech production (Ekmekeçi, 1986; Radford, 1990; Slobin & Bever, 1982) – Figure–MOTION–Ground in English (*the girl ran towards the fence*) versus Figure–Ground–MOTION in Turkish (*kız çite doğru koştu* ‘she fence-TO towards RAN’).

The developmental findings for *co-speech gesture* are less conclusive. Children increase their production of representational iconic co-speech gestures, depicting features of objects (e.g., holding cupped hands in air to form a ball shape) or actions on objects (e.g., moving an empty palm forward as if throwing a ball), at age 3 or 4 years (McNeill, 1992; Özçalışkan & Goldin-Meadow, 2011). However, little is known about cross-linguistic patterns in children’s early iconic gestures. In addition to being sparse, the literature on children’s co-speech gesture production focuses exclusively on packaging of motion, with largely inconclusive findings: Some studies find language-specific gesture patterns in packaging around age 3–6 years (e.g., Özçalışkan, 2007); others find a more extended timeline for language-specific gestures (e.g., Özyürek et al., 2008). However, work on *comprehension* of co-speech gestures shows that children aged 3–4 have greater difficulty understanding co-speech gestures that do *not* follow language-specific patterns than co-speech gestures that do follow language-specific patterns (Glaser et al., 2018), suggesting early attunement to language-specific patterns in co-speech gesture. There is no existing cross-linguistic work that examines developmental changes in learning language-specific patterns of ordering in the description of motion events.

Limits on the effects of language on thinking, even during communication: Silent gesture

Interestingly, the language-specific patterns found in the gestures adults produce when speaking (co-speech gesture) are *not* found when adults describe the same

events in gesture without speech (silent gesture, Özçalışkan et al., 2016a, 2016b, 2018). This finding indicates limits on the effects that language has on thinking, even within a communicative act. When is this limit first seen?

We know little about the development of silent gesture. The few studies that have been conducted focus on the structure of these gestures in speakers of a particular language (e.g., in 6- and 8-year-old German speakers, Bohn et al., 2019; in 4- and 12-year-old British English speakers, Clay et al., 2014). But these studies do not compare silent gesture to co-speech gesture in the same children nor do they compare silent gesture in child speakers of different languages. Nonetheless, the studies provide evidence for language-like structures – conventionality, abstraction, segmentation – in children’s early silent gestures. These patterns are largely independent of the grammatical structure of the child’s native spoken language. No work has yet examined similarities and differences in silent gesture in child speakers of different languages.

We address these gaps and inconclusive findings in our study by observing speech, co-speech gesture, and silent gesture in child speakers of two structurally different languages (Turkish, English) over a broad age span (3 to 12 years). We ask two questions: (1) When do language-specific patterns in children’s co-speech gestures appear in development? Based on the currently inconclusive literature on co-speech gesture, we expect that children will show language-specific adult-like patterns in co-speech gesture either later than (>ages 3–4) or at the same time as (ages 3–4) they show language-specific speech. If so, we will have evidence that language has an early effect on thinking that goes beyond words during communication (2). When do children first display cross-linguistic similarities in their silent gestures? Given the scarcity of work on silent gesture, we expect that children might or might not show the cross-linguistic similarities that adults exhibit in silent gesture at an early age (3–4). If the former holds true, we will have evidence that limits on language’s effect on thinking, even during communication, appear early in development.

Overall, our study, provides the first comprehensive analysis of developmental changes in co-speech and silent gesture in two structurally different languages, using a new corpus. It focuses on both emergence of language-specific patterns (or their lack) in packaging and ordering of semantic elements in the expression of motion events – a domain whose expression shows systematic variability and patterned regularities in adult speech and gesture.

Methods

Sample

Participants were 100 children, learning either English ($n = 50$) or Turkish ($n = 50$) as their native language, each equally divided into 5 age-groups: 3–4 ($M_{age} = 4;2$ [$SD = 0;5$]), 5–6 ($M_{age} = 5;8$ [$SD = 0;7$]), 7–8 ($M_{age} = 7;11$ [$SD = 0;7$]), 9–10 ($M_{age} = 10;1$ [$SD = 0;8$]), and 11–12 ($M_{age} = 11;11$ [$SD = 0;7$]) years with roughly comparable numbers of boys and girls in each group; see Table 1 for sample characteristics by language and age. The choice of this age range was based on earlier work, which showed that language-specific patterns in descriptions of motion arise in children’s speech at age 3–4 years (Özçalışkan & Slobin, 1999) and in their gestures between ages 3 and 9 (Özçalışkan et al., 2014; Özyürek et al. 2008). Earlier work (Özçalışkan, 2009) suggested that 10 children in each group would give 84% power for the detection of

Table 1. Child age and gender by language and age group (years; months)

	3–4	5–6	7–8	9–10	11–12
<i>English</i>					
Mean age (SD)	4;1 (0;4) 5F, 5 M	5;6 (0;2) 5F, 5 M	7;9 (0;6) 5F, 5 M	9;10 (0;8) 5F, 5 M	12;0 (0;5) 7F, 3 M
<i>Turkish</i>					
Mean age (SD)	4;2 (0;6) 7F, 3 M	5;10 (0;9) 5F, 5 M	8;1 (0;8) 6F, 4 M	10;5 (0;6) 5F, 5 M	11;10 (0;9) 5F, 5 M

Abbreviations: F, female; M, male; SD, standard deviation.

significant effects for p values $<.05$ and an effect size of $\eta^2 = 0.08$. The data from children speaking English and Turkish were gathered in the United States and Turkey, respectively, as part of a bigger research project that focused on patterns of gesturing in blind individuals.¹ Participants' families received monetary compensation. Data from 6 additional participants were excluded ($n = 3/\text{language}$) due to either speech production difficulties (e.g., stuttering) or failure to complete the experiment. The study was carried out in accordance with the Code of Ethics for the protection of human research participants. The protocol was approved by an American research university institutional review board and informed consent was obtained from the participants' families prior to their children's participation in the study.

Procedure

Data collection

Children were asked to describe 8 three-dimensional scenes one at a time. Each scene showed motion in one of three path types (to, from, over) in relation to different landmarks (house, carpet, hurdle) with various manner types (e.g., run, jump, crawl). Each scene was glued onto a small board; the scene included a landmark and three stationary yet varying poses of the same doll, depicting a motion event involving both manner and path. Children were first introduced to the figure – with the name Oya in Turkish and Eve in English – and told that she would do different kinds of activities in different scenes involving various objects. Children were also explicitly informed that the figure would be repeated 3 times in each scene, but as part of one continuous movement. The motion scenes were presented with counterbalanced order in two blocks with 4 items (see Table 2).

Children provided descriptions for the scenes in two conditions: with speech while naturally moving their hands (co-speech gesture condition: 'tell me what is happening in this scene using both your words and your hands'); and using their hands without any speech (silent gesture condition: 'tell me what is happening in this scene but only using your hands without speaking'). The descriptions for all the scenes were first elicited in speech (and co-speech gesture) followed by silent gesture. We did not counterbalance the two conditions to eliminate any effect that silent gesture might have had on the

¹The 3-dimensional motion scenes allowed for elicitation of motion descriptions in gesture and speech in a format accessible to both sighted and blind individuals as part of the larger project. The patterns of speech and gesture production were, however, in line with earlier work with adult speakers that relied on dynamic motion scenes (e.g., Özçalışkan, 2016; Tütüncü et al., 2023), suggesting that the participants in our study treated these events as if they were dynamic scenes.

Table 2. Stimulus motion events

SET 1	SET 2
<i>Run into house</i> <i>Crawl over carpet</i> <i>Run out of house</i> <i>Walk towards crib</i>	<i>Climb into treehouse</i> <i>Jump over hurdle</i> <i>Crawl out house</i> <i>Run away from motorcycle</i>

naturalness of children’s co-speech gesture. Each child did two practice trials prior to describing the scenes in each condition; these familiarization trials were not included in any of the analyses. The co-speech and silent gesture conditions were separated from each other by two other unrelated tasks – one on metaphors and one on narratives – thus eliminating any possible immediate effect of responses in the co-speech condition on responses in the silent gesture condition. Children were not allowed to touch the scenes with their hands (see Fig. 1 for data collection set-up). We explicitly asked children to use their hands along with their words in the co-speech gesture condition to elicit comparable amounts of gesture production across ages and languages.

Transcription and coding

Children’s speech responses for each language in the co-speech gesture condition were transcribed by native speakers of that language; they were then parsed into sentence units based on earlier work (Özçalışkan, 2016; Özçalışkan et al., 2016b, 2018). A sentence unit was defined as consisting of a verb, along with the arguments and subordinate clauses associated with it (e.g., *She is running into house*; *Eve koşuyor* ‘house-to running’; and *Eve giriyor koşarak* ‘House-to enters running’). All gesture responses children produced in the two conditions (co-speech, silent) were also coded. We defined gesture as movements of the hand or body that characterized movements or features of the scenes for communicative purposes. We further coded each sentence unit for (1) packaging and (2) ordering of semantic elements.

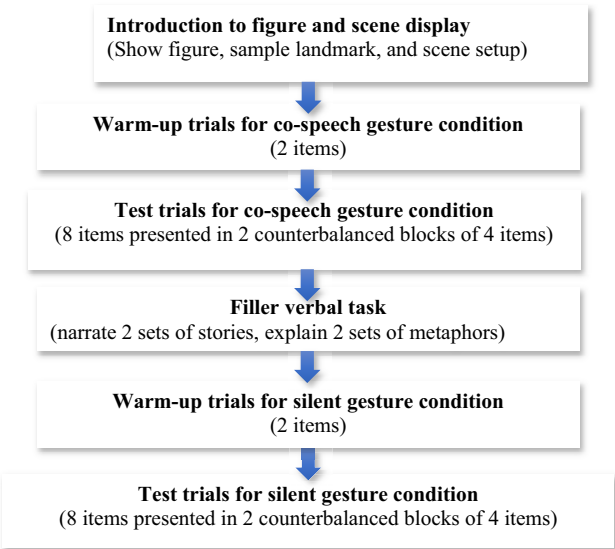


Figure 1. Flowchart for the data collection procedure

Packaging

Speech and gestures in each sentence unit were classified as either conflated or separated (see Fig. 2). Conflated sentence units included responses in which both motion components (manner+path) were expressed in a single clause or gesture. Separated sentence units included responses that expressed only manner (e.g., *she runs*, *koşar* ‘runs’), only path (e.g., *she enters the house*, *ev-e girer* ‘house-to enter’), or manner and path, expressed in separate gestures or separate clauses (e.g., *eve girer koşarak* ‘house-to enters running’) – a speech response produced once in English, but relatively often in Turkish (54 instances).

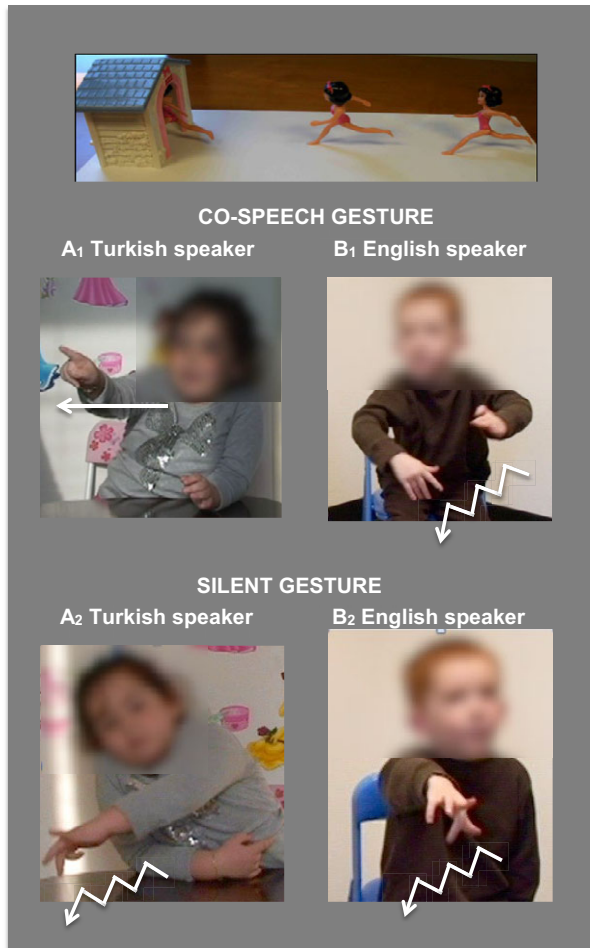


Figure 2. The three-dimensional stimulus scene depicting a girl's running motion toward house (top panel) and its description in gesture by 3- to 4-year-old children learning Turkish or English. In co-speech gesture, child learners of English combine manner and path into a single gesture (rotating both palms rapidly while moving them forward to convey running forward; B1); child learners of Turkish express only path without manner (tracing a line with right index finger left to right conveying rightward motion; A1). In silent gesture, child speakers in each language combine manner and path into a single gesture by walking middle and index fingers left to right (A2) or walking right hand forward away from speaker (B2). The jagged arrows indicate motion with both manner and path; the straight arrows indicate motion with only path.

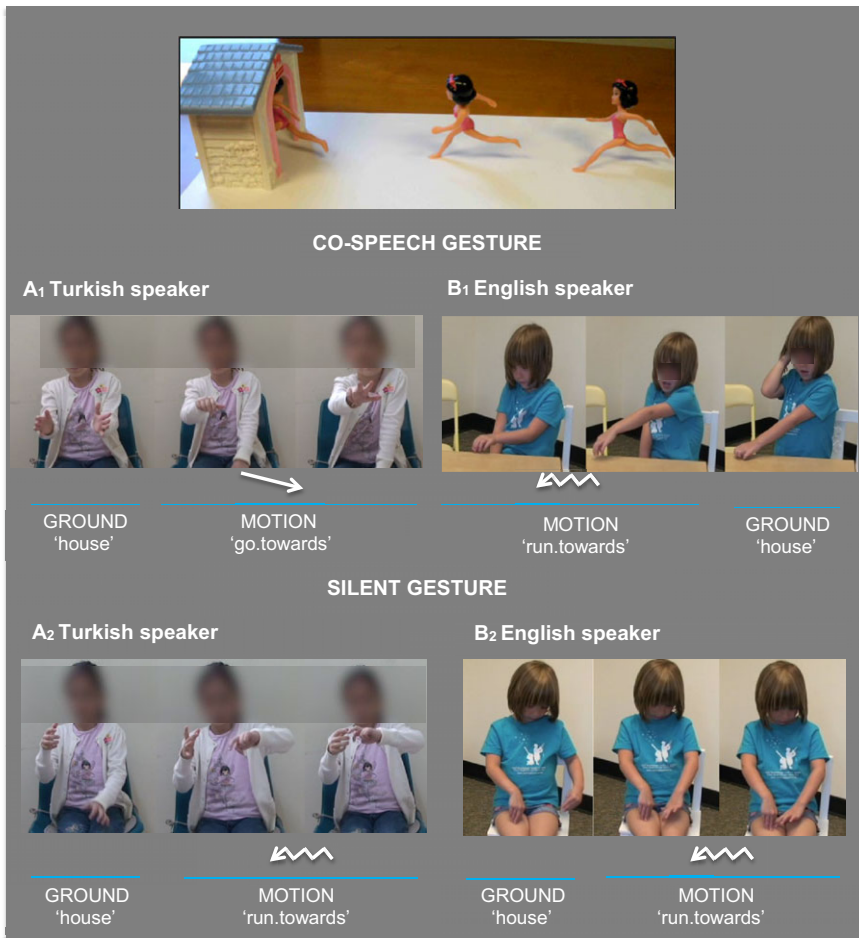


Figure 3. The three-dimensional stimulus scene depicting a girl's running motion toward a house (top panel) and its description in gesture by 5- to 6-year-old children learning Turkish (A panels) or English (B panels). In co-speech gesture, children learning English who produced ground and motion gestures along with their speech expressed motion (run_towards) before ground (house); children learning Turkish gestured ground (house) before motion (move_towards). In silent gesture, child speakers of each language expressed ground before motion (run_towards). The jagged arrows indicate motion with both manner and path; the straight arrows indicate motion with only path.

Ordering

Speech and gesture strings in each sentence unit were classified as following either *Figure–MOTION–Ground* order or *Figure–Ground–MOTION* order (see Fig. 3). Assigning spoken responses to one of these two orders was based on the location of the primary motion component, namely, the main verb that typically expressed path (Turkish) or manner (English). When expressed, secondary motion elements (e.g., prepositions, particles, and adjectives) conveying path in English and manner in Turkish were always associated with the main verb. Assigning gesture strings to either one of the orders was based on the location of the motion gesture that frequently expressed manner+path in English and only-manner, only-path, or sequential manner–path gesture in Turkish. Manner–path sequential gestures were always

contiguous (i.e., followed each other without any other gesture in between) but were infrequent in both languages (English: 7 instances, Turkish: 15 instances). When combining gestures into strings, children were likely to combine a gesture for motion with a gesture for landmark, leaving out a gesture for the figure (Fig. 3A₁,B₁). Children typically expressed the ground element with a sideways or downward facing palm (e.g., right palms in Fig. 3A₂,B₂).

Children sometimes showed a mixed pattern: They used a separated and a conflated gesture together in a single sentence unit. These instances were infrequent; $M = 1.8\%$ of sentence units (range = $0.7\% - 3.5\%$) across languages and conditions. We omitted all the 'mixed' instances from our analysis for *packaging*, as we could not classify them as belonging to either packaging category. Most of the mixed sentence units in each language were also omitted from the analysis of ordering because they expressed either only motion or a mixed ordering pattern. However, we included responses that had mixed packaging but followed consistent ordering (English: 5 instances, Turkish: 10 instances) in the order analysis – constituting 38% of the small number of responses in the mixed category in both languages.

Children predominantly represented different motion elements with each hand (placing flat left palm on the left side of the body to represent house, placing the right index finger on the right side of the body as if figure, moving fingers of the right hand left to right to convey motion) in both co-speech gesture (67%) and silent gesture (77%). The use of both hands – with or without the accompanying bodily enactment – to represent a single motion element (e.g., rapidly circling both arms simultaneously to convey running, crawling forward on carpet on all fours to convey crawling forward, and placing downward facing palms in the shape of an inverse V as if house) was relatively less frequent, accounting for 33% of the co-speech and 23% of the silent gestures. The majority of the gestures were produced in the air (~80%) in both co-speech and silent gesture, but some (~20%) were also produced either on the body (e.g., hopping fingers on the upper leg to convey jumping) or on the table in front of the participant (e.g., placing the cupped left hand on the table as if carpet and crawling fingers of the right hand over the left hand as if crawling over landmark).

We assessed reliability with independent coders who were native speakers of each language. They coded a randomly selected 10% of the data for each age and condition in each language. Inter-coder agreement was high: 97% for identification of gestures, 100% for description of gesture type, and 97% and 99% for categorization of motion elements for gesture and speech.

Analysis

The count data were analyzed using Bayesian mixed-effects Poisson generalized linear models implemented in the *stan_glmer()* function of the *rstanarm* package (Goodrich et al., 2020). The models provide fully-Bayesian inference with Markov Chain Monte Carlo (MCMC) sampling of posterior distributions to produce parameter estimates. The 'mixed effects' approach allows us to better estimate effects of interest by modeling and controlling for the idiosyncratic contributions of nuisance variables to the outcome (e.g., one stimulus scene eliciting more gestures than another stimulus scene). The Poisson linking function in the Bayesian mixed-effects models works as in any typical generalized linear model: it links a discrete dependent variable with a set of independent variables.

In our analysis, we specified our mixed-effects structures according to the 'keep it maximal' principle (Barr et al., 2013); that is, by including random subject and scene

slopes for the fixed-effect term of interest where the design allowed for their estimation, as well as random intercepts for subject and scene. We adopted the ‘weakly informative’ default priors provided by *stan_glmer()*, which show normal distribution that is centered at 0, with standard deviations of 2.5. We increased the *adapt_delta*, number of chains, iterations per chain, and warmup iterations where it was necessary to get stable parameter estimates with no divergent transitions. All models we reported converged well with *Rhat* values that fell within .005 of 1.0. All parameter estimates had associated effective sample sizes of at least 1000.

We submitted the fitted models to *bayesfactor models()* or *describe_posterior()* functions from the *bayestestR* package (Makowski et al., 2019) to support inferences from the modeling results. For each model, we report values describing the 90% credible interval (90% CI) based on the highest density interval (HDI) of the posterior distribution. The highest density interval corresponds to a range of the posterior within which all points inside the bounds of the interval have a higher probability density than the points outside the bounds of the interval. The interpretation of a 90% (HDI) CI meant that there is a 90% chance that the true parameter would fall within the CI range. The point estimate of the parameter (b) value that we provide with each 90% CI is the median of that highest-density interval of the posterior. We also report a Bayes factor with each result, which is a ratio that quantifies how strongly the data support the null hypothesis, as opposed to the alternative hypothesis, namely, values under 1.0 in favor of the null and values above 1.0 in favor of the alternative hypothesis. For example, a Bayes factor of 12.0 would mean that the alternative hypothesis is 12 times more likely than the null hypothesis (given the priors and the data), whereas a Bayes factor of 0.5 would mean the null hypothesis is twice as likely as the alternative hypothesis. For a given effect, the characterization of the strength of support (e.g., ‘the data provided strong/moderate/anecdotal evidence that...’) corresponded to the interpretation of the associated Bayes factor, using language adopted from Lee and Wagenmakers (2014). Anonymized data summaries and coding manuals can be found at the link: https://osf.io/fse6r/?view_only=a90669bcf98b4dcd8e4e24b3aef1d32b.

Results

Packaging motion elements

Speech

Children learning English or Turkish differed in the way they packaged motion components in speech (language x packaging interaction; $b = 1.93$, 90% CI = [1.31, 2.52], $BF > 100$; Fig. 4A). Children learning English preferred conflated to separated packaging ($b = -0.82$, 90% CI = [-1.09, -0.53], $BF > 100$), expressing manner and path in the same clause (e.g., *she runs into the house*). Conversely, children learning Turkish preferred separated to conflated packaging ($b = 1.36$, 90% CI = [0.71, 2.03], $BF = 8.4$), describing similar scenes by expressing path by itself (*Ev-e giriyor* ‘house-to entering’), manner by itself (*koşuyor* ‘running’), or path in the main clause and manner in the subordinate clause (*eve girer koşarak* ‘house-to enter running’). The language-specific patterns in speech were evident by 3–4 years ($b = 2.61$, 90% CI = [1.80, 3.54], $BF > 100$) and remained relatively stable over developmental time (BFs > 50). The 11–12 age-group showed the same pattern, but this cross-linguistic difference was not reliable ($b = .99$, 90% CI = [0.36, 1.60], $BF = 1.27$).

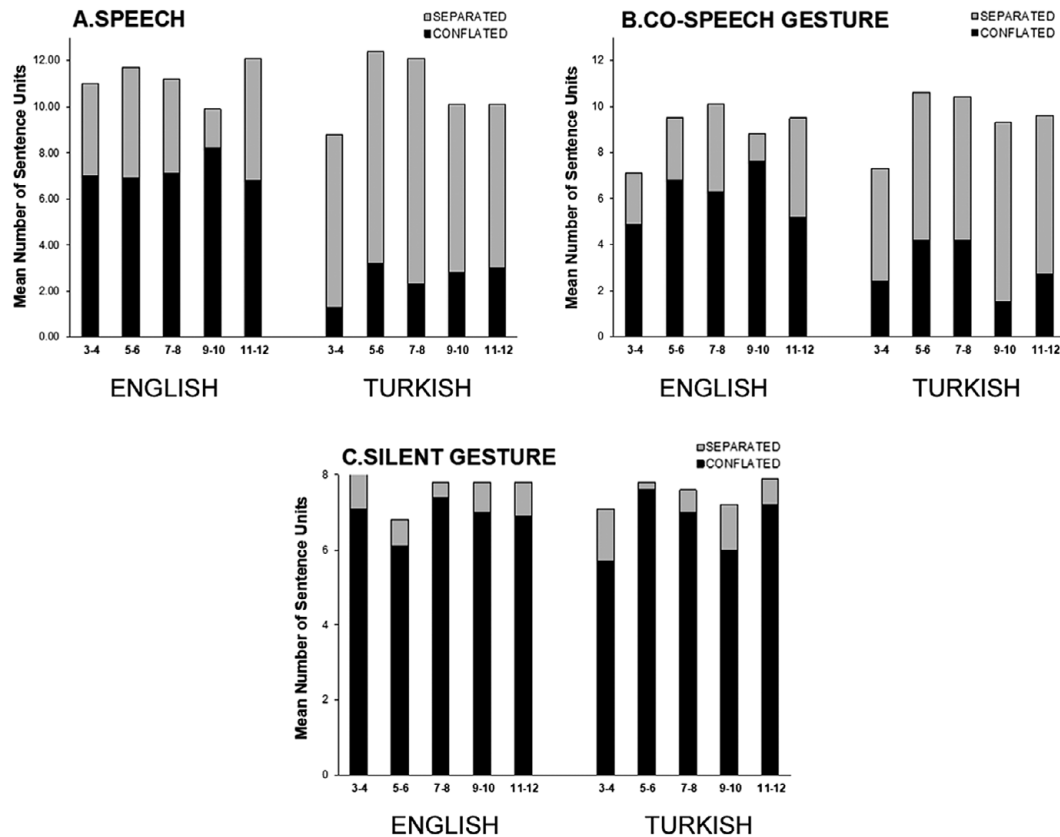


Figure 4. The mean number of sentence-units with separated or conflated packaging that children produced in speech (A), in co-speech gesture (B), and in silent gesture (C). Child native speakers of the two languages show cross-linguistic differences in speech and co-speech gesture, and cross-linguistic similarities in silent gesture, by age 3-4 years (the max possible number of sentence-units was 8 for the silent gesture condition).

Co-speech gesture

Co-speech gesture showed the same pattern of cross-linguistic differences as speech (language $\times$ packaging interaction; $b = 1.54$, 90% CI = [1.28, 1.79], $BF > 100$, Fig. 4B). Children learning English used a greater number of gestures with conflated than with separated packaging ($b = -1.47$, 90% CI = [-1.96, -0.97], $BF > 100$), expressing manner and path in one gesture (e.g., running fingers [manner] as the hand moved along a forward trajectory [path]). Turkish speakers, in contrast, opted for more separated than conflated responses ($b = 1.05$, 90% CI = [0.53, 1.51], $BF = 12.1$), producing a gesture for either path (e.g., moving finger forward) or manner (e.g., running fingers in place), or using two sequential gestures (one expressing path and one expressing manner) in the same sentence unit (e.g., running fingers in place, followed by moving the hand forward). The language-specific patterns in co-speech gesture were evident by 3–4 years ($b = 1.53$, 90% CI = [0.92, 2.15], $BF = 49.2$) and remained unchanged over developmental time, $BFs = 2.93$ –100. We thus observed language-specific patterns in the packaging of motion components across speech and co-speech gesture starting at age 3–4 years.

Silent gesture

We next asked whether child English and Turkish speakers, when communicating without speech, displayed the same patterns that adult speakers of the two languages used in their silent gestures – that is, packaging patterns without any cross-linguistic differences (Fig. 4C). Child speakers of Turkish and English preferred conflated to separated responses in their silent gesture (English; $b = -3.72$, 90% CI = [-4.73, -2.8], $BF > 100$, Turkish; $b = -2.99$, 90% CI = [-3.69, -2.32], $BF > 100$); there was no cross-linguistic difference across groups in the strength of the preference for conflated packaging ($b = 0.133$, 90% CI = [-2.65, 0.53], $BF = 0.028$). Moreover, the preference for conflated responses emerged at an early age: 3- to 4-year-old children produced more conflated than separated responses in both the American ($b = -2.79$, 90% CI = [-4.30, -1.55], $BF > 100$) and Turkish ($b = -2.97$, 90% CI = [-3.71, -2.33], $BF > 100$) groups – a pattern that held within each age-group (all $BFs < 1$), and with moderate to strong evidence of no difference in the conflated packaging preference between language groups (language–packaging interaction; $b = 0.67$, 90% CI = [-0.137, 1.45], $BF = 0.13$; see Table A.1 in Appendix for means and standard errors for each packaging type by age and language in speech, co-speech gesture, and silent gesture).

*Ordering semantic elements**Speech*

Children learning English or Turkish differed in their ordering of semantic elements in speech (language $\times$ order interaction; $b = -12.78$, 90% CI = [-16.55, -9.74], $BF > 100$; Fig. 5A). They showed greater production of Figure–MOTION–Ground than Figure–Ground–MOTION order in English ($b = 7.34$, 90% CI = [5.00, 9.85], $BF > 100$), and greater production of (Figure)–Ground–MOTION than Figure–MOTION–Ground order in Turkish ($b = -4.53$, 90% CI = [-5.32, -3.78], $BF > 100$); parentheses around the Figure indicate that it was optional and not always produced. The language-specific ordering patterns in speech were evident at age 3–4 years ($b = -11.00$, 90% CI = [-15.28, -7.23], $BF > 100$) and remained unchanged over developmental time (all $BFs > 100$).

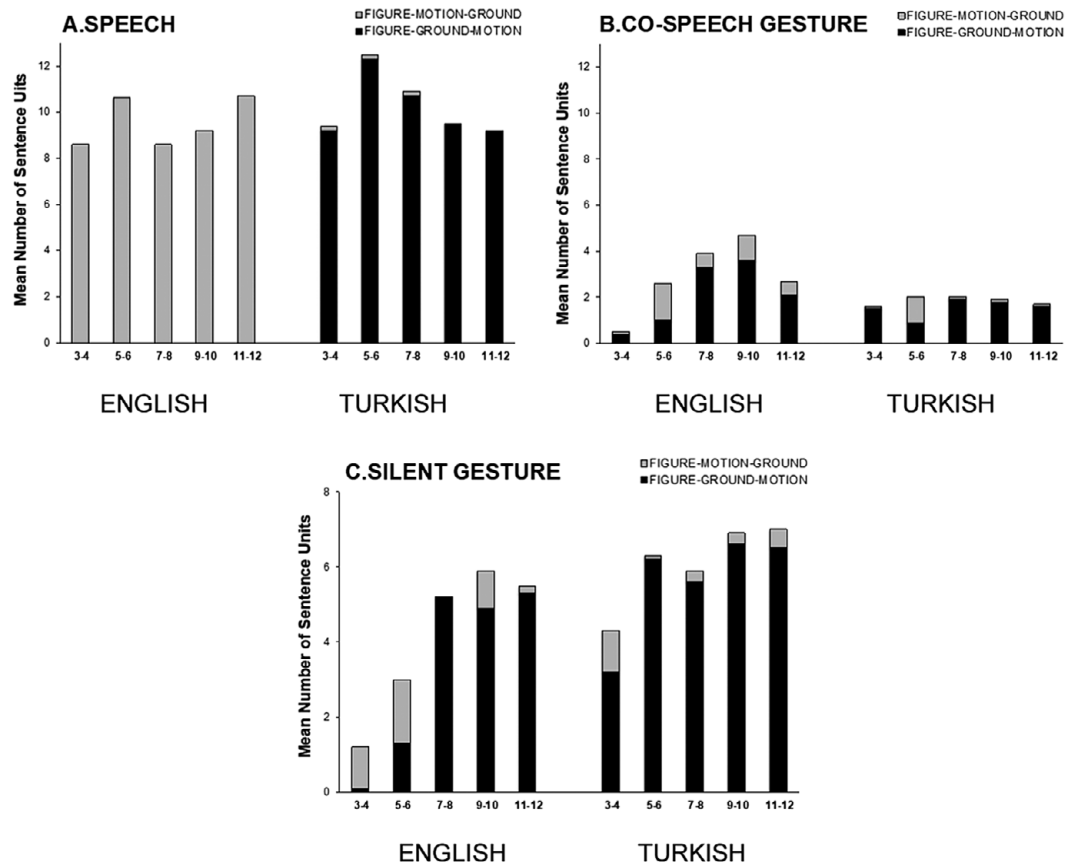


Figure 5. The mean number of sentence-units with (Figure)-Ground-MOTION or (Figure)-MOTION-Ground orders children produced in speech (A), co-speech gesture (B), and silent gesture (C). Child learners of Turkish and English showed cross-linguistic differences in speech and cross-linguistic similarities in silent gesture. These patterns were evident in speech at age 3-4, in silent gesture at age 5-6 for Turkish, and 7-8 for English. Children produced relatively few co-speech gesture responses that contained gestures for both the ground and the motion; as a result, the majority of the co-speech gestures were not analyzed for order. Children produced one string per scene in the silent gesture condition, making 8 the max possible number of sentence units for the silent gesture condition.

Co-speech gesture

Children speaking English or Turkish rarely concatenated gestures in sequences before age 5–6 years and, in fact, did not produce many concatenated gestures at any age, making it difficult to observe robust patterns. Indeed, the ordering of the children's co-speech gestures from age 5–6 years onward showed no cross-linguistic differences (language–order interaction; $b = -0.59$, 90% CI = $[-1.19, -0.01]$, $BF = 0.14$; Fig. 5B). Turkish speakers strongly preferred (Figure)–Ground–MOTION ordering in their co-speech gestures starting at age 5–6 years ($b = -2.73$, 90% CI = $[-4.33, -1.39]$, $BF = 67.3$), language-specific pattern that was found in adult Turkish speakers. English speakers also showed an anecdotal to moderate preference for this ordering ($b = -1.67$, 90% CI = $[-2.91, -0.48]$, $BF = 2.3$). English speakers did, however, produce *proportionally* more (Figure)–MOTION–Ground orders (the English-specific pattern) than Turkish speakers. We thus observed language-specific ordering patterns in *speech* in both Turkish- and English-speaking children, and in *co-speech gesture* in Turkish-speaking children throughout development.

Silent gesture

Next, we asked whether children's gestures, when produced without speech, would show cross-linguistic similarities in the ordering of motion elements. We found that children learning Turkish or English did *not* show the differences we observed in children's speech or co-speech gesture (see Fig. 5C). Instead, both language groups showed an overall preference for (Figure)–Ground–MOTION ordering in silent gesture (English; $b = -2.79$, 90% CI = $[-4.25, -1.44]$, $BF > 100$, Turkish; $b = -3.01$, 90% CI = $[-3.76, -2.34]$, $BF > 100$). This preference was evident in children by age 5–6 years in Turkish ($b = -4.17$, 90% CI = $[-5.88, -2.67]$, $BF > 100$) and 7–8 years in English ($b = -5.41$, 90% CI = $[-8.32, -3.08]$, $BF > 100$) speakers, remaining relatively stable thereafter. The children's silent gestures did show a moderate cross-linguistic difference (language $\times$ order interaction; $b = -0.53$, 90% CI = $[-0.78, -0.31]$, $BF = 9.9$), which stemmed from the slightly more pronounced preference for (Figure)–Ground–MOTION ordering – the typical Turkish pattern – in Turkish speakers than in English speakers (see Table A.1 in Appendix for means and standard errors for each ordering type by age and language in speech, co-speech gesture, and silent gesture).

Discussion

Adults display systematic cross-linguistic differences in speech when they package and order the semantic elements of a motion event (Özçalışkan et al., 2016a, 2016b, 2018). These cross-linguistic differences also affect the organization of semantic elements in gesture, but only when those gestures are produced *with speech* (co-speech gesture), not when they are produced *without speech* (silent gesture). More specifically, adult speakers of Turkish and English *package* and *order* semantic elements of events differently, and in accordance with the language they speak, when describing those semantic elements in co-speech gesture. However, they package and order the same semantic elements similarly when describing them in silent gesture (Özçalışkan et al., 2016b, 2018; Tütüncü et al., 2023). Here, we found that children learning either Turkish or English display these adult patterns in co-speech gesture and silent gesture as early as ages 3–4 years.

Focusing first on *co-speech gesture*, we found that children learning English used more conflated gestures than children learning Turkish, who produced more separated gestures. Neither group produced many multiple-gesture combinations in their co-speech gestures, limiting our ability to draw strong conclusions about gesture ordering. But the children's packaging patterns in co-speech gesture strongly mirrored their spoken language and were found in the youngest groups (ages 3–4), providing support for an early influence of language on thinking during the speaking act.

Turning next to *silent gesture*, we found that children learning either English or Turkish both conflated manner and path within a single gesture (the English pattern) and did so at age 3–4 years. Children learning either English or Turkish also followed the (Figure)–Ground–MOTION gesture order (the Turkish pattern) in their multi-gesture combinations, beginning around age 5 to 6 years. Our silent gesture results thus provide evidence for an early limit on the effect that language has on thinking, even during communication.

Note that the cross-linguistic differences we observed in speech and gesture were robust, even with a relatively modest sample. The magnitude of the Bayes factors (>100 for speech and co-speech gesture) indicates that our alternative hypotheses were at least 100 times more likely to be true than the null hypotheses. The large effect sizes, as indexed by Bayes factors, thus provide strong evidence for early emerging cross-linguistic differences in co-speech gesture, along with cross-linguistic similarities in silent gesture. We explore the implications of our developmental findings for co-speech gesture and silent gesture in the next two sections.

What co-speech gesture tells us about the effects of language on thought

Packaging manner and path

We found an early effect of language on co-speech gesture in how children package manner and path semantic elements. Children learning Turkish or English followed their respective spoken languages' packaging strategies (manner and path separated for Turkish, conflated for English) in their gestures at age 3 to 4 years. Language can thus influence the *nonverbal* representation of an event during speech production at an early age, providing support for Slobin's (1996) thinking-for-speaking account and its early onset. Our findings on co-speech gesture also lend support to theories of gesture–speech integration. Under Kita and Özyürek's (2003) interface theory, gesture and speech are assumed to arise from two separate systems (an action generator for gesture, a message generator for speech). Nevertheless, the two work in tandem from conceptualization to articulation to convey intended meanings, constituting an integrated system (McNeill, 1992). Our finding that language-specific patterns appear in 3-year-olds' co-speech gestures indicate that gesture–speech integration begins early in development.

Co-speech gesture not only *reveals* the effects of language on thought, but it can also *shape* those thoughts. Three- to 4-year-old English-speaking children, when taught novel verbs accompanied by iconic gestures depicting manner, generalized significantly more verbs to novel events that depicted the same or similar types of action than when taught novel verbs accompanied by gestures that did *not* convey manner (Aussems & Kita, 2021, see also Mumford & Kita, 2014). Observing co-speech gesture can change thought. Note, however, that the children in this earlier study were all English speakers, for whom manner constitutes a frequently expressed

semantic component both in speech and in gesture. The fact that the majority of the co-speech gestures produced by the Turkish children in our study expressed only path information raises the possibility that the beneficial effects of the type of information conveyed in co-speech gesture might also vary by language – a possibility that needs to be explored in future research.

Ordering ground and motion

Our results for ordering in co-speech gesture are tentative because children rarely produced multiple semantic elements in co-speech gesture in either language. This pattern is consistent with the ‘one gesture per spoken clause’ preference observed in adult speakers (McNeill, 1992). The few strings children learning Turkish produced in co-speech gesture followed the ordering patterns in their speech (optional Figure–Ground–Motion), but the co-speech gestures that children learning English produced did not mirror their speech. The limited number of strings children used in co-speech gesture in either language prevents us from making broad conclusions based on these patterns.

Might the cross-linguistic differences in packaging that we observed in speech by age 3 to 4 years be evident in gesture even earlier, thus preceding and/or predicting upcoming changes in language-specific speech? Research examining children’s gesture–speech system at different language milestones (i.e., first words, first sentences, or first noun phrases) has found that children take their first step into a milestone in gesture alone or in gesture with speech, only later attaining the same milestone exclusively in speech (e.g., Cartmill et al., 2014; Iverson & Goldin-Meadow, 2005; Özçalışkan et al., 2017; Özçalışkan & Goldin-Meadow, 2005). However, this pattern is primarily found for deictic gestures (i.e., pointing at objects, which precedes producing nouns for the same objects). Deictic gestures emerge earlier in development than iconic gestures, the type of gesture we focus on here. Iconic gestures emerge around age 3 years, long after children have begun producing their first verbs conveying motion (Özçalışkan et al., 2014; Stites & Özçalışkan, 2017, 2021). The relatively late onset of iconic gestures makes it less likely that precursors of language-specific patterns in speech will be found in gesture. Future studies, however, can shed further light on this question by studying younger children using nonverbal tasks other than gesture (e.g., ordering pictures that depict ground, motion, or figure) as a way to test the effect of language on the nonverbal representation of events.

What silent gesture tells us about limits on the effects of language on thought

Packaging manner and path

In contrast to the differences in packaging found between Turkish and English learners’ co-speech gestures, both Turkish- and English-speaking children display a robust preference for conflating manner and path in their silent gestures. Conflation in silent gesture appeared in each group at age 3–4 years, and the preference remained unchanged over developmental time. Interestingly, even child speakers of English (who use conflation in their co-speech gestures) increased their use of conflation in their silent gestures.

What explains the early emergence of the conflated pattern in silent gesture, particularly in Turkish children, who followed a separated pattern almost exclusively in their co-speech gestures? The verbal expression of motion in Turkish requires that path be expressed in the main clause (*gir* ‘enter’), accompanied by manner in a

subordinate clause (*koşarak* ‘by running’), resulting in two clauses. The two-clause requirement in speech might create a heavier cognitive load for Turkish speakers than for English speakers, who need to produce only one clause (*run to*). In fact, Turkish speakers – adult and child – frequently leave out manner from their motion descriptions and only express path (Özçalışkan, 2009, 2016). Unlike speech, gesture allows expression of both manner and path at the same time in a relatively easy-to-produce form. Perhaps this is the reason that Turkish- and English-speaking children find it easy to adopt the conflated form in silent gesture and that Turkish- and English-speaking adults maintain the form in silent gesture.

The conflated form for manner and path is also found in the earliest stages of *homesigns*, gesture languages created by children who have no usable model for language. Homesigners are children whose hearing losses are so profound that they cannot make use of the spoken language input that surrounds them, and whose hearing parents have not exposed them to sign language. Despite this lack of linguistic input, the children create gesture systems that have many of the properties of natural language (Goldin-Meadow, 2003, 2023). Homesigners in both the United States and in Turkey use the conflated form to convey manner and path. However, the homesigners in both countries also produce a form that is partially sequenced – a conflated manner+path gesture produced along with either a path gesture (e.g., CLIMB+UP – UP) or a manner gesture (CLIMB+UP – CLIMB) (Goldin-Meadow et al., 2015). This mixed form is the first step in segmenting and sequencing the path and manner semantic elements, a step that precedes the fully segmented form (CLIMB – UP) in language emergence (Senghas et al., 2013).

Ordering ground and motion

Our ordering results in silent gesture echoed our packaging results. We found the same ordering preference in both Turkish- and English-speaking children when describing events using only their hands – (Figure)–Ground–Motion, the Turkish pattern. This preference emerged slightly later in English than in Turkish, possibly because English speakers had to switch their ordering preference from (Figure)–Motion–Ground (SVO) in speech to (Figure)–Ground–Motion (SOV) in silent gesture. The SOV ordering in silent gesture mirrors earlier work with adult speakers using their hands to describe motion events (Özçalışkan et al., 2018; Tütüncü et al., 2023) and to describe events in which an animate entity acts on an inanimate entity (Goldin-Meadow et al., 2008; Hall et al., 2013; Langus & Nespors, 2010; Meir et al., 2010; Schouwstra & de Swart, 2014).

Why do young speakers prefer to express ground before motion in their silent gestures? The three key components of a motion event include two entities – a *figure* and a *ground* – along with a *motion* that stipulates the relation between the two entities. When describing events in silent gesture, it might be communicatively more informative (i.e., providing the most information with the fewest tools; Grice, 1975) and/or cognitively less burdensome (Gentner, 1982; Goldin-Meadow et al., 2008) to set up figure and ground as anchors before conveying the motion that relates the two – resulting in (Figure)–Ground–Motion ordering.

Is the ordering of semantic elements in silent gesture unique to gesture? In other words, is the ordering unique to communication, or does it extend to other nonverbal representations of events? Earlier work (Gershkoff-Stowe & Goldin-Meadow, 2002; Goldin-Meadow et al., 2008) has found that adult native speakers of different languages order pictorial depictions of semantic elements as they would have ordered

the elements in silent gesture, picking up the picture depicting the action or motion *after* picking up the picture of the figure and the ground. Whether this default ordering is found in other non-communicative nonverbal behaviors is an important question for future research and can be explored by examining a broader range of cognitive tasks in children learning structurally different languages.

Interestingly, the ordering found in silent gesture – ground (or patient) preceding motion (or act) – is also found in the signs of homesigners aged 3 to 5 years in the United States (Gershkoff-Stowe & Goldin-Meadow, 2002; Goldin-Meadow, 2003), China (Goldin-Meadow & Mylander, 1998), and Turkey (Goldin-Meadow et al., 2015). Silent gesture thus appears to simulate the first step in creating a manual language (see Goldin-Meadow, 2015, for a review that compares packaging and ordering patterns in silent gesture to conventional sign languages and homesign systems).

In sum, we have found that at an early age, children learning languages that differ in how they organize motion events display language-specific patterns in co-speech gesture, but not in silent gesture. The close alignment between speech and gesture during communication highlights the integration of the two modalities and leads to cross-linguistic differences in co-speech gesture. These cross-linguistic differences reflect an early effect of language on thought during the act of speaking, observable in co-speech gesture. But the cross-linguistic similarities that we found in silent gesture suggest the possibility of a language of gesture that is *not* affected by the speaker's language. This language of gesture appears early in development in speakers of two structurally distinct languages; it is also evident in homesign systems around the globe. As such, these gestures are likely to reflect a basic cognitive structure that is recruited for communicating about events when no conventional system is available.

Data availability statement. The video records for the study consist of identifiable data; we cannot provide access to these data based on the confidentiality agreement we signed with the participants and the regulations of the institutional review board for research at our institution. However, we posted anonymized quantitative summaries of our data along with coding manuals at the link: https://osf.io/fse6r/?view_only=a90669bcf98b4dcd8e4e24b3aef1d32b. The photographs of all the three-dimensional stimulus scenes are available upon request. We do not have any computational models associated with the analysis.

Acknowledgments. This research was supported by a grant from the March of Dimes Foundation (#12-FY08-160) to Özçalışkan and Goldin-Meadow. We thank Andrea Pollard, Christianne Ramdeen, and Burcu Sancar for help with data collection, transcription, and coding.

Competing interest. The authors declare none.

References

- Allen, S., Özyürek, A., Kita, S., Brown, A., Furman, R., Ishizuka, T., & Fujii, M. (2007). Language-specific and universal influences in children's syntactic packaging of Manner and Path: A comparison of English, Japanese and Turkish. *Cognition*, 102, 16–48.
- Aussem, S., & Kita, S. (2021). Seeing iconic gesture promotes first- and second-order verb generalization in preschoolers. *Child Development*, 92(1), 124–141.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278.
- Bohn, M., Kachel, G., & Tomasello, M. (2019). Young children spontaneously recreate core properties of language in a new modality. *PNAS*, 116(51), 26072–26077.
- Cardini, F. (2010). Evidence against Whorfian effects in motion conceptualization. *Journal of Pragmatics*, 42(5), 1442–1459.

- Cartmill, E. A., Hunsicker, D., & Goldin-Meadow, S. (2014). Pointing and naming are not redundant: Children use gesture to modify nouns before they modify nouns in speech. *Developmental Psychology*, 50(6), 1660–1666. <https://doi.org/10.1037/a0036003>
- Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition*, 41(1–3), 83–121.
- Choi, S., & Lantolf, J. P. (2008). Representation and embodiment of meaning in L2 communication: Motion events in the speech and gesture of advanced L2 Korean and L2 English speakers. *Studies in Second Language Acquisition*, 30(2), 191–224.
- Chui, K. (2009). Linguistic and imagistic representations of motion events. *Journal of Pragmatics*, 41(9), 1767–1777.
- Chui, K. (2012). Cross-linguistic comparison of representations of motion in language and gesture. *Gesture*, 12(1), 40–61.
- Clay, Z., Pople, S., Hood, B., & Kita, S. (2014). Young children make their gestural communication systems more language-like: Segmentation and linearization of semantic elements in motion events. *Psychological Science*, 25(8), 1518–1525.
- Ekmekçi, Ö. (1986). Significance of word order in the acquisition of Turkish. In D. Slobin & K. Zimmer (Eds.), *Studies in Turkish linguistics* (pp. 265–272). John Benjamins.
- Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, 32(5), 429–492.
- Gennari, S. P., Sloman, S. A., Malt, B. C., & Fitch, W. (2002). Motion events in language and cognition. *Cognition*, 83(1), 49–79.
- Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj (Ed.), *Language development: Vol. 2, Language, thought and culture* (pp. 301–334). Erlbaum.
- Gershkoff-Stowe, L., & Goldin-Meadow, S. (2002). Is there a natural order for expressing semantic relations? *Cognitive Psychology*, 45(3), 375–412.
- Glaser, M. G., Williamson, R., & Özçalışkan, Ş. (2018). Do children understand iconic gestures about events as early as iconic gestures about entities? *Journal of Psycholinguistic Research*, 47(3), 741–754.
- Goldin-Meadow, S. (2003). The resilience of language: What gesture creation in deaf children can tell us about how all children learn language. In J. Werker & H. Wellman (Eds.), *Essays in developmental psychology series*. Psychology Press.
- Goldin-Meadow, S. (2015). The impact of time on predicate forms in the manual modality: signers, homesigners and silent gesturers. *Topics in Cognitive Science* 7, 169–184.
- Goldin-Meadow, S. (2023). *Thinking with your hands: The surprising science behind how gestures shape our thoughts*. Basic Books.
- Goldin-Meadow, S., & Mylander, C. (1998). Spontaneous sign systems created by deaf children in two cultures. *Nature*, 391, 279–281.
- Goldin-Meadow, S., Namboodiripad, S., Mylander, C., Ozyurek, A., & Sancar, B. (2015). The resilience of structure built around the predicate: Homesign gesture systems in Turkish and American deaf children. *Journal of Cognition and Development*, 16(1), 55–80.
- Goldin-Meadow, S., So, W. C., Özyürek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *PNAS*, 105(27), 9163–9168.
- Goodrich, B., Gabry, J., Ali, I., & Brilleman, S. (2020). rstanarm: Bayesian applied regression modeling via Stan. *R Package Version*, 2(19), 3.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics*, vol. 3, speech acts (pp. 41–58). Academic Press.
- Hall, M. L., Mayberry, R. I., & Ferreira, V. S. (2013). Cognitive constraints on constituent order: Evidence from elicited pantomime. *Cognition*, 129(1), 1–17.
- Hickmann, M., Taranne, P., & Bonnet, P. (2009). Motion in first language acquisition: Manner and path in French and English child language. *Journal of Child Language*, 36(4), 705–741.
- Hohenstein, J. M. (2005). Language-related motion event similarities in English- and Spanish-speaking children. *Journal of Cognition and Development*, 6(3), 403–425.
- Ibarretxe-Antuñano, I. (2004). Language typologies in our language use: The case of Basque motion events in adult oral narratives. *Cognitive Linguistics*, 15(3), 317–349.
- Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, 16, 368–371.

- Kita, S., & Özyürek, A. (2003). What does crosslinguistic 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(1), 16–32.
- Langus, A., & Nespore, M. (2010). Cognitive systems struggling for word order. *Cognitive Psychology*, 60, 291–318.
- Lee, M. D., & Wagenmakers, E. J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.
- Lewandowski, W., & Özçalışkan, Ş. (2021). How language type influences patterns of motion expression in bilingual speakers. *Second Language Research*, 37(1), 27–49.
- Lucy, J. (1992a). *Language diversity and thought: A reformulation of the linguistic relativity hypothesis (Studies in the Social and Cultural Foundations of Language, No. 12)*. Cambridge University Press.
- Lucy, J. (1992b). *Grammatical categories and cognition: A case study of the linguistic relativity hypothesis (Studies in the Social and Cultural Foundations of Language, No. 13)*. Cambridge University Press.
- Makowski, D., Ben-Shachar, M. S., Chen, S. H. A., & Lüdecke, D. (2019). Indices of effect existence and significance in the Bayesian framework. *Frontiers in Psychology*, 10, 2767.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.
- Meir, I., Lifshitz, A., Ilkbasaran, D., & Padden, C. (2010). The interaction of animacy and word order in human languages: A study of strategies in a novel communication task. In A. D. M. Smith, M. Schouwstra, B. de Boer, & K. Smith (Eds.), *Proceedings of the eighth evolution of language conference* (pp. 455–456). World Scientific Publishing Co.
- Mumford, K. H., & Kita, S. (2014). Children use gesture to interpret novel verb meanings. *Child Development*, 85(3), 1181–1189.
- Naigles, L. R., Eisenberg, A. R., Kako, E. T., Hightner, M., & McGraw, N. (1998). Speaking of motion: Verb use in English and Spanish. *Language and Cognitive Processes*, 13(5), 521–549.
- Özçalışkan, Ş. & Slobin, D. I. (2003). Codability effects on the expression of manner of motion in English and Turkish. In A. S. Özsoy, M. Nakipoglu-Demiralp, E. Erguvanli-Taylan & A. Aksu-Koç (Eds.), *Studies in Turkish Linguistics* (pp. 259–270). Istanbul: Bogaziçi University Press.
- Özçalışkan, Ş. (2007). Metaphors we move by: Children's developing understanding of metaphorical motion in typologically distinct languages. *Metaphor and Symbol*, 22, 147–168.
- Özçalışkan, Ş. (2009). Learning to talk about spatial motion in language-specific ways. In J. Guo, E. Lieven, S. Ervin-Tripp, N. Budwig, K. Nakamura & Ş. Özçalışkan (Eds.), *Cross-linguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin* (pp. 263–276). New York: Psychology Press.
- Özçalışkan, Ş. (2016). When do gestures follow speech in bilinguals' description of motion? *Bilingualism: Language and Cognition*, 19(3), 644–653.
- Özçalışkan, Ş., Adamson, L. B., Dimitrova, N., & Baumann, S. (2017). Early gesture provides a helping hand to spoken vocabulary development for children with autism, Down syndrome and typical development. *Journal of Cognition and Development*, 18(3), 325–337.
- Özçalışkan, Ş., Gentner, D., & Goldin-Meadow, S. (2014). Do iconic gestures pave the way for children's early verbs? *Applied Psycholinguistics*, 35, 1143–1162.
- Özçalışkan, Ş., & Goldin-Meadow, S. (2005). Gesture is at the cutting edge of early language development. *Cognition*, 96(3), B101–B113. <http://doi.org/10.1016/j.cognition.2005.01.001>
- Özçalışkan, Ş., & Goldin-Meadow, S. (2011). Is there an iconic gesture spurt at 26 months? In G. Stam & M. Ishino (Eds.), *Integrating gestures: The interdisciplinary nature of gesture* (pp. 163–174). John Benjamins.
- Özçalışkan, Ş., Lucero, C., & Goldin-Meadow, S. (2016a). Is seeing gesture necessary to gesture like a native speaker? *Psychological Science*, 27(5), 737–747.
- Özçalışkan, Ş., Lucero, C., & Goldin-Meadow, S. (2016b). Does language shape silent gesture? *Cognition*, 148, 10–18.
- Özçalışkan, Ş., Lucero, C., & Goldin-Meadow, S. (2018). Blind speakers show language-specific patterns in co-speech gesture but not silent gesture. *Cognitive Science*, 42(3), 1001–1014.
- Özçalışkan, Ş., & 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 Boston University conference on language development* (pp. 541–552). Cascadia Press.
- 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(4), 1040.

- Radford, A. (1990). *Syntactic theory and acquisition of English syntax: The nature of early child grammars in English*. Blackwell.
- Sapir, E. (1961). In D. G. Mandelbaum (Ed.), *Culture, language and personality. Selected essays*. University of California Press.
- Schouwstra, M., & de Swart, H. (2014). The semantic origins of word order. *Cognition*, 131(3), 431–436.
- Senghas, A., Özyürek, A., & Goldin-Meadow, S. (2013). Homesign as a way-station between co-speech gesture and sign language: The evolution of segmentation and sequencing. In R. Botha & M. Everaert (Eds.), *The evolutionary emergence of language: Evidence and inference* (pp. 62–76). Oxford University Press.
- Slobin, D. (1996). From ‘thought’ and ‘language’ to ‘thinking for speaking’. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge University Press.
- Slobin, D. I. (2004). The many ways to search for a frog: Linguistic typology and the expression of motion events. In S. Strömquist & L. Verhoeven (Eds.), *Relating events in narrative: Typological and contextual perspectives* (Vol. 2, pp. 219–257). Lawrence Erlbaum Associates.
- Slobin, D. I., & Bever, T. (1982). Children use canonical sentence schemas: A crosslinguistic study of word order and inflections. *Cognition*, 12(3), 229–265.
- Stites, L. J., & Özçalışkan, Ş. (2017). Who does what to whom: Children track story referents first in gesture. *Journal of Psycholinguistic Research*, 46(4), 1019–1032.
- Stites, L. J., & Özçalışkan, Ş. (2021). Time is at hand: Literacy predicts changes in children’s gestures about time. *Journal of Psycholinguistic Research*, 50, 967–983.
- Talmy, L. (1985). Semantics and syntax of motion. In J. Kimball (Ed.), *Syntax and semantics* (Vol. 4, pp. 181–238). Academic Press.
- Talmy, L. (2000). *Toward a cognitive semantics: Typology and process in concept structuring*. The MIT Press.
- Tütüncü, I., Emerson, S., Şengül, M., Kenesevic, M., & Özçalışkan, Ş. (2023). When gestures *do* or *do not* follow language-specific patterns of motion expression in speech: Evidence from Chinese, English, and Turkish. *Cognitive Science*, 47(4), e13261.
- Whorf, B. L. (1956). Language, Mind, and Reality. In J. B. Carroll, S. C. Levinson, & P. Lee (Eds.), *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf* (pp. 315–344). The MIT Press.

Appendix

Table A1. Means and standard errors for types of packaging and ordering in speech, co-speech gesture and silent gesture.

		SPEECH		CO-SPEECH GESTURE		SILENT GESTURE		
		separated	conflated	separated	conflated	separated	conflated	
PACKAGING ENGLISH	AGE	<i>M</i> (<i>SE</i>)	<i>M</i> (<i>SE</i>)	<i>M</i> (<i>SE</i>)	<i>M</i> (<i>SE</i>)	<i>M</i> (<i>SE</i>)	<i>M</i> (<i>SE</i>)	
	3–4	4.0 (0.68)	7.0 (0.32)	2.20 (0.79)	4.90 (0.85)	0.90 (0.60)	7.10 (0.60)	
	5–6	4.80 (0.92)	6.90 (0.46)	2.70 (1.57)	6.80 (1.01)	0.70 (0.47)	6.10 (0.99)	
	7–8	4.10 (0.86)	7.10 (0.46)	3.80 (1.2)	6.30 (0.79)	0.40 (0.22)	7.40 (0.34)	
	9–10	1.70 (0.24)	8.20 (0.35)	1.20 (0.88)	7.60 (1.01)	0.80 (0.53)	7.00 (0.68)	
TURKISH	11–12	5.30 (1.03)	6.80 (0.07)	4.30 (1.1)	5.20 (1.18)	0.90 (0.80)	6.90 (0.80)	
	3–4	7.50 (0.35)	1.30 (0.15)	4.90 (1.27)	2.40 (0.96)	1.40 (0.65)	5.70 (0.90)	
	5–6	9.20 (1.13)	3.20 (0.42)	6.40 (1.44)	4.20 (0.87)	0.20 (0.20)	7.60 (0.31)	
	7–8	9.80 (0.89)	2.30 (0.38)	6.20 (1.76)	4.20 (1.09)	0.60 (0.40)	7.00 (0.42)	
	9–10	7.30 (0.30)	2.80 (0.46)	7.80 (1.36)	1.50 (1.00)	1.20 (0.55)	6.00 (0.86)	
ORDERING	11–12	7.10 (0.76)	3.00 (0.30)	6.90 (1.19)	2.70 (1.01)	0.70 (0.40)	7.20 (0.42)	
		<i>F-G-M</i>	<i>F-M-G</i>	<i>F-G-M</i>	<i>F-M-G</i>	<i>F-G-M</i>	<i>F-M-G</i>	
	ENGLISH	3–4	0	8.60 (0.37)	0.40 (0.22)	0.10 (0.10)	0.10 (0.10)	1.10 (0.64)
	5–6	0	10.60 (1.29)	1.00 (0.79)	1.60 (1.07)	1.30 (0.84)	1.70 (1.05)	
	7–8	0	8.60 (0.31)	3.30 (1.13)	0.60 (0.43)	5.20 (1.01)	0	
TURKISH	9–10	0	9.20 (0.29)	3.60 (1.25)	1.10 (0.72)	4.90 (1.13)	1.00 (0.68)	
	11–12	0	10.70 (0.72)	2.10 (0.66)	0.60 (0.27)	5.30 (0.96)	0.20 (0.13)	
	3–4	9.20 (1.13)	0.20 (0.13)	1.50 (0.82)	0.10 (0.10)	3.20 (0.93)	1.10 (0.35)	
	5–6	12.30 (1.92)	0.20 (0.13)	0.90 (0.48)	1.10 (0.99)	6.20 (0.79)	0.10 (0.10)	
	7–8	10.70 (0.97)	0.20 (0.13)	1.90 (1.05)	0.10 (0.10)	5.60 (0.88)	0.30 (0.21)	
	9–10	9.50 (0.60)	0	1.80 (0.70)	0.10 (0.10)	6.60 (0.58)	0.30 (0.21)	
	11–12	9.20 (0.71)	0	1.60 (0.81)	0.10 (0.10)	6.50 (0.82)	0.50 (0.34)	

Abbreviations: F-M-G, Figure–MOTION–Ground ordering; F-G-M, Figure–Ground–MOTION ordering; M, mean; SE, standard error.

Cite this article: Özçalışkan, Ş., Lucero, C., & Goldin-Meadow, S. (2023). What the development of gesture with and without speech can tell us about the effect of language on thought, *Language and Cognition*, 1–22. <https://doi.org/10.1017/langcog.2023.34>