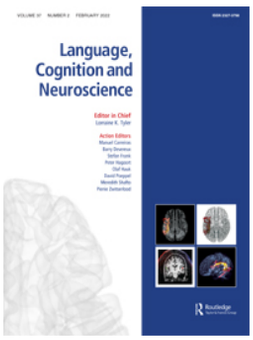


Routledge
Taylor & Francis Group

Language, Cognition and Neuroscience

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/plcp21

Cross-modal investigation of event component omissions in language development: a comparison of signing and speaking children

Beyza Sümer & Aslı Özyürek

To cite this article: Beyza Sümer & Aslı Özyürek (2022): Cross-modal investigation of event component omissions in language development: a comparison of signing and speaking children, Language, Cognition and Neuroscience, DOI: 10.1080/23273798.2022.2042336

To link to this article: https://doi.org/10.1080/23273798.2022.2042336

LANGUAGE, COGNITION AND NEUROSCIENCE
https://doi.org/10.1080/23273798.2022.2042336

Routledge
Taylor & Francis Group

REGULAR ARTICLE

OPEN ACCESS Check for updates

Cross-modal investigation of event component omissions in language development: a comparison of signing and speaking children

Beyza Sümer[a,b] and Aslı Özyürek [b,c]

[a]Department of Linguistics, University of Amsterdam, Amsterdam, The Netherlands; [b]Centre for Language Studies, Radboud University Nijmegen, Nijmegen, The Netherlands; [c]Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen, Nijmegen, The Netherlands

ABSTRACT

Language development research suggests a universal tendency for children to be under-informative in narrating motion events by omitting components such as Path, Manner or Ground. However, this assumption has not been tested for children acquiring sign language. Due to the affordances of the visual-spatial modality of sign languages for iconic expression, signing children might omit event components less frequently than speaking children. Here we analysed motion event descriptions elicited from deaf children (4–10 years) acquiring Turkish Sign Language (TİD) and their Turkish-speaking peers. While children omitted all types of event components more often than adults, signing children and adults encoded more Path and Manner in TİD than their peers in Turkish. These results provide more evidence for a general universal tendency for children to omit event components as well as a modality bias for sign languages to encode both Manner and Path more frequently than spoken languages.

ARTICLE HISTORY
Received 18 December 2020
Accepted 3 February 2022

KEYWORDS
Language acquisition; motion; sign language; spoken language; language modality

Introduction

Children constantly observe motion events around them (e.g. mother running up the stairs), in which a Figure experiences a change in location with respect to a Ground and accompanied by Manner (e.g. run) and/or Path (e.g. up). In describing such events, children have a bias to focus on certain event components in their speech and omit others, and therefore they are less informative in their narrations compared to adults (e.g. Bunger et al., 2012; Hickmann, 2003; Özyürek et al., 2008; Papafragou & Selimis, 2010; Papafragou et al., 2006). This robust tendency has been explained as resulting from various underlying cognitive mechanisms as well as factors related to linguistic knowledge. These include developmental differences between children and adults in terms of attentional capacity (Scerif et al., 2005), processing mechanisms involved in utterance preparation (e.g. Adams & Gatherhole, 2000; Levelt, 1989), pragmatic awareness (e.g. Papafragou et al., 2006) and language-specific lexical/syntactic knowledge (Allen et al., 2007; Bunger et al., 2012).

It is not yet well understood, however, whether the modality of language being acquired (i.e. sign vs speech) modulates this universal developmental trajectory. Pursuing this question in the domain of motion events can reveal new insights since the visual-spatial modality of sign language allows events to be mapped onto the signing space through visually motivated (i.e. iconic) language forms (e.g. Cuxac & Sallandre, 2007; Engberg-Pedersen, 2003; Galvan & Taub, 2006; Hoiting & Slobin, 2007; Perniss & Özyürek, 2008; Taub & Galvan, 2001). In the present study, we aim to understand the role of language modality in the development of motion event expressions under the assumption that visually motivated form-meaning mappings might encourage sign language acquiring children to express more event components than their speaking peers. It might, however, also be the case that learning how to narrate events by children will be insensitive to the effects of language modality since linguistic input in language development will be mapped onto already available concepts. This would suggest domain-general developmental factors (e.g. cognitive, pragmatic), rather than language modality as the force driving developmental linguistic patterns (e.g. Allen et al., 2007; Bunger et al., 2012; Clark, 2004; Gleitman, 1990; Jackendoff, 1996; Papafragou et al., 2006; Pinker, 1989).

CONTACT Beyza Sümer b.sumer@uva.nl Department of Linguistics, University of Amsterdam, Spuistraat 134, Amsterdam 1012 VB, The Netherlands
Supplemental data for this article can be accessed https://doi.org/10.1080/23273798.2022.2042336.

Development of encoding motion event components in spoken languages

Speakers of different languages express the Path and Manner components of motion events differently in terms of how they are packaged lexically and syntactically (e.g. Slobin, 1996; Talmy, 1985). For example, adult speakers of satellite-framed languages like English and Chinese tend to conflate Motion with Manner in the main verb of a sentence and express Path in a non-verb position (e.g. preposition). This pattern is demonstrated in sentence (1a) below, in which the verb "roll" describes the Manner of Motion and "down" describes its Path of Motion. In contrast, speakers of verb-framed languages like Spanish, Greek, and Turkish are more likely to conflate Motion with Path in the main verb and to encode Manner (if at all) in a linguistic form outside the main verb (e.g. adverb). In sentence (1b), the Path of Motion is encoded in the main verb (i.e. fall "*düş*") and the Manner of Motion as a subordinate clause (i.e. roll "*yuvarlan*"). In this example, the Manner verb is inflected with a connective (i.e. CONN),[1] Ground with an ablative case marker (i.e. ABL), and the Path verb with a past tense marker (i.e. PAST).

(1a)	The rock	rolled	down	the hill.	English (Talmy, 1985)	
	Figure	Manner + Motion	Path	Ground		

(1b)	Kaya	yuvarlan + arak	tepe + den	düş + tü.	Turkish
	Rock	roll + CONN	hill + ABL	fall+PAST	
	Figure	Manner	Ground	Path + Motion	

"The rock fell from the hill while rolling."

Despite such variation across languages in packaging event components, comparing children's event descriptions to those of adults yields a universal developmental picture whereby children's event descriptions express fewer event components regardless of the spoken language being acquired. English-acquiring children between 3- and 5-years of age, for example, produce fewer event descriptions that include *both* Path and Manner information than do adults (Bunger et al., 2012; Özyürek et al., 2008; Papafragou & Selimis, 2010), and this tendency extends until 8 years of age (Papafragou et al., 2006). Similar tendencies were also observed for the expression of event components other than Path and Manner, namely Ground. Hickmann (2003) examined the picture story narrations of adults and children (4–10 years old) in four different spoken languages (English, French, German, Mandarin Chinese). She found a strikingly similar developmental progression. Children showed a tendency to omit Grounds in their narratives and were less likely to express both Figure and Ground in adult-like ways until 10 years of age.

One explanation for children's omission of event components relates to the limitations in their attentional capacity compared to adults. Thus, children are challenged in simultaneously encoding the complete set of components that make up a complex event since they do not pay attention to all components (e.g. Scerif et al., 2005). It has also been proposed that linearising simultaneously occurring event components poses a challenge for speaking children by creating a cognitive demand (Berman & Slobin, 1994), thus causing them to omit components in their event descriptions. Another explanation may be children's insufficient mastery of pragmatic strategies (Bunger et al., 2012), leading them to miscalculate how much information needs to be encoded in a motion event description (Papafragou et al., 2006). A final explanation for the omissions could be that children have not yet conceptualised these components in an adult-like way. However, it has been found that children early on can discriminate Path versus Manner (e.g. Pulverman et al., 2008; Pulverman et al., 2006; Pulverman et al., 2003) and Figure versus Ground (e.g. Göksun et al., 2009) in motion events in non-linguistic tasks. In the case of Figure and Ground, Göksun et al. (2009) found that English-reared infants pay attention to Figures and Grounds quite early (11 months for Figures and 14 months for Grounds) in events with animated geometric shapes where a Figure is moving towards a Ground.

Whether such omissions also occur in children who use the visual modality, i.e. sign languages, for linguistic expressions has not been systematically tested. Compared to speaking children, signing children might have a greater advantage in learning to encode motion events because the modality of sign language allows event components to be visually and iconically mapped onto visible articulators to encode simultaneous aspects of events (Slonimska et al., 2020 for adult users of Italian Sign Language [LIS]). This iconic mapping possibility might allow signing children to omit fewer components than their speaking peers.

Encoding and development of motion event expressions in sign languages

Signers encode different components of a motion event by using their visible body articulators and signing space, which allow them to exploit the iconic and spatial affordances of the visual-spatial modality (e.g. Engberg-Pedersen, 1993; Galvan & Taub, 2006; Supalla, 1982; Tang et al., 2007; Taub & Galvan, 2001;

Zwitserlood, 2003). This is exemplified in (2), where a signer of American Sign Language (ASL) encodes the Figure (one animal) with a two-legged classifier (CL), expressed by the bent index and middle fingers of her left hand, that incorporates Manner (walking) and Path (to) information, while her right hand represents the Ground (the other animal), located and held in the signing space (Arık, 2009).

(2)

LH: TWO ANIMAL $CL(animal)_{loc}$ $CL(animal)_{mot}$ $CL(animal)_{mot}$
RH: ANIMAL $CL(animal)_{loc}$ HOLD
"There are two animals facing each other. The one on the left is walking towards to the one on the right."

Linguistic strategies for motion expressions in sign languages can be considered multi-morphemic complex structures that consist of discrete morphemes for Manner and/or Path of Motion (e.g. wiggling of fingers and change of location), and in the form of classifiers (inverted handshapes to represent each animal) for Figure and Ground, combined simultaneously in a classifier predicate (Supalla, 1982; Zwitserlood, 2012). In contrast to spoken languages, these morphemes have visually motivated relationships to their referents (e.g. Talmy, 2003), such as an animal's walking legs mapped onto an inverted V-handshape etc. Furthermore, sign languages seem to be quite homogenous in having visually motivated forms (Aronoff et al., 2003).

In line with the idea that such motivated form-meaning mapping might facilitate the encoding of more event components in sign than spoken languages, comparisons of ASL signers with English speakers (Galvan & Taub, 2006) and signers of French Sign Language (LSF) with French speakers (Sallandre et al., 2018) have revealed that while describing a motion event, signers were found to encode more information than speakers, mainly by providing more information for *both* Manner and Path. Thus, it might be interesting to ask whether signing children also encode such motion event components more frequently than their speaking peers and do not exhibit the same universal tendency for omission relative to adults.

Earlier studies with signing children alone suggest that even though they focus on certain components while omitting others early on, very soon they begin to express many of the event components (Supalla, 1982). Newport and Supalla (1980) and Newport (1981) have investigated the acquisition of certain movement morphemes in three ASL-acquiring children between the ages of 3;6 and 5;11. They report that the youngest child (age 3;6) in their study was able to produce simple Path movements (e.g. linear, arc). These early productions, elicited through short vignette narrations, did not convey information about the Manner component. However, a few months later, a small proportion of the productions conveyed both Path and Manner information (Newport, 1981). More recent studies also seem to suggest a facilitating role of the visual-spatial modality of sign languages on the acquisition of these event components. Sallandre et al. (2018) found that deaf children (aged 5–10 years) acquiring LSF encoded both Path and Manner from early on, as opposed to French-acquiring children, who frequently produced Path-only descriptions and fewer Path + Manner utterances in their narrations of videos showing motion events with various Paths and Manners. Additionally, language forms that express Path and Manner emerge early in signing children's vocabularies (e.g. Fenson et al., 1994; Naigles et al., 1998). Finally, two studies on the gestural repertoires of deaf homesigners who have not been exposed to a conventional sign language (Özyürek et al., 2015; Zheng & Goldin-Meadow, 2002) report the expression of motion event components such as Manner and Path in children between the ages of 3–6, thus suggesting the visual modality plays a role in facilitating the expression of event components.

However, a signing advantage in expressing Path and Manner does not seem to generalise to learning to encode the Ground component of a motion event. Despite the iconic and visual affordances of the signed modality, signing children have a strong tendency to consistently omit Ground from their descriptions. In an earlier work, Supalla (1982) found a few expressions of Ground in motion event descriptions of a 3;6-year-old ASL-acquiring child, but these were very infrequent. Morgan et al. (2008) also report the case of a BSL-acquiring child between 2;0 and 2;6 who tended to refer to real objects around him as Grounds rather than encoding them via lexical signs or classifier predicates. Ground omission was also reported for older signing children by Slobin et al. (2003) for ASL-acquiring children (between ages 5–12), De Beuzeville (2006) for Australian Sign Language (AUSLAN)-acquiring children (about age 9); Tang et al. (2007) for children acquiring Hong Kong Sign Language (HKSL) (about age 12), and Engberg-Pedersen (2003) for Danish Sign Language (DSL)-acquiring children (about age 13). In these studies, however, definitions of Ground are not clear or consistent since the authors draw their conclusions either from spontaneous signing data (Morgan et al., 2008; Slobin et al., 2003) or narrations of picture stories (Engberg-Pedersen,

2003; Tang, 2003), which might contain several anchor points as possible Grounds.

Despite lacking direct comparisons between signing and speaking children, the findings of the above-mentioned studies seem to suggest some developmental differences between signed and spoken expressions used in event descriptions and for different event components. For Path and Manner, it seems that sign languages allow both components to be encoded more frequently than in speech (e.g. Galvan & Taub, 2006; Sallandre et al., 2018), which points toward a strong effect of the language modality being acquired. For Grounds, however, there seems to be a universal bias in children to omit them, regardless of the language modality they are acquiring (De Beuzeville, 2006; Engberg-Pedersen, 2003; Morgan et al., 2008; Slobin et al., 2003; Supalla, 1982; Tang, 2003; Tang et al., 2007). However, studies comparing all these components across a sign and spoken language in a systematic way are missing.

Our goal in this paper is to conduct the first systematic study of how different event components (Path, Manner, Ground) are encoded by directly comparing signing and speaking children who complete the same tasks under similar language production conditions, i.e. vignette descriptions. We would like to examine whether there are similar or different biases in the encoding or omission of different event components, which could shed further light on our understanding of the role of domain-general developmental factors and/or language modality in the domain of how events are mapped onto language during development.

The present study

The main purpose of the present study is to investigate whether and in which ways language modality modulates learning to encode different semantic components of a motion event in Turkish Sign Language (Türk İşaret Dili [TİD]) and Turkish, a verb-framed spoken language. We also aim to understand whether the acquisition of some event components is more sensitive to the effects of language modality than others. In order to be in line with the previous studies with signers and speakers (Galvan & Taub, 2006; Sallandre et al., 2018; Taub & Galvan, 2001), we focus on the comparison of sign versus speech only, thus the use of co-speech gestures during event narrations is beyond the scope of the current paper.[2]

We predict that the visual-spatial modality of TİD might allow for the encoding of more event components than Turkish does, for both adults and children, mirroring the findings in earlier studies (Galvan & Taub, 2006; Sallandre et al., 2018; Taub & Galvan, 2001). Further evidence for this prediction comes from the acquisition of locative expressions by signing and speaking children (aged 8 years), in which the former group was found to be more informative in providing "left/right" information than their speaking peers, even when gestures were taken into account (Karadöller et al., 2020; under review; Sümer, 2015). If this tendency generalises to motion events, this sign advantage might also manifest itself in the development of event narrations. Thus, it is also possible that TİD-signing children might encode event components more frequently and become adult-like earlier than their Turkish-speaking peers. Alternatively, both groups of children may exhibit similar developmental patterns and omit these components despite the differences in language modality, thus suggesting that language development simply scaffolds on domain-general developmental tendencies (e.g. Allen et al., 2007; Bunger et al., 2012; Clark, 2004; Gleitman, 1990; Jackendoff, 1996; Papafragou et al., 2006; Pinker, 1989). A third possibility is that the patterns of event component encoding may differ depending on the type of the event component. Considering the previous findings with (home)signing children (Fenson et al., 1994; Naigles et al., 1998; Özyürek et al., 2015; Sallandre et al., 2018; Zheng & Goldin-Meadow, 2002), we might see a sign advantage for the development of Path and Manner components, but not for the expression of Ground (e.g. Hickmann, 2003).

Method

Participants

We recruited 10 adult Turkish speakers and 10 adult deaf signers of TİD, and 20-Turkish-acquiring children and 20 TİD-acquiring children, where for each modality 10 were school age and 10 were preschool age (see Table 1 below for demographic information). All deaf participants in this study had two deaf parents and thus had been exposed to TİD since birth.

The age group categories for the children were based on the age for starting primary school in Turkey. At the time of data collection, children started primary school

Table 1. The number (N) of participants, the mean (M) and the standard deviation (SD) of their age, and their gender (F = Female; M = Male).

Group	*N*	M_{AGE}	SD_{AGE}	Gender
TİD-signing adults	10	31;2	9;7	7 F, 3 M
Turkish-speaking adults	10	36	10;4	7 F, 3 M
TİD-acquiring children				
School-age children	10	8;3	0;1	1 F, 9 M
Preschool-age children	10	5;2	1;3	5 F, 5 M
Turkish-acquiring children				
School-age children	10	8;2	1	7 F, 3 M
Preschool-age children	10	5;3	0;1	3 F, 7 M

Figure 1. A still from a vignette that shows a tomato rolling from the top of a hill (Source), away from the triangle (Source), towards a tree (Goal) on a hill/slope (Surface), for four possible Grounds.

at age 7, so children younger than 7 were categorised as preschool-age children and children aged 7 and up were categorised as school-age children. Seven deaf children in the school-age group attended a primary school for the deaf and three were in mainstream schools for the hearing. As for the preschool-age group, three of them were full-time (five days a week) and four were part-time (two days a week) attenders of a preschool education programme for the deaf. The rest stayed at home. It is important to note that TİD was not systematically taught at the schools for the deaf and thus was not part of the curriculum at the time of data collection. Furthermore, due to the lack of access to spoken language input, deaf children learned very little Turkish at school. For the hearing children, all of them in the school-age group were receiving formal education. Five of the hearing children in the preschool-age group attended a preschool education programme five days a week.

Stimulus materials

The data for the current study were collected by eliciting descriptions of eight short vignettes that depict motion events whereby a Figure is changing its location with respect to a Ground, along a Path in a certain Manner. In these vignettes, Figures are either people (involving two animated Lego figures and two real people; a total of four vignettes) (used originally in Goldin-Meadow et al., 2008) or geometric-shaped characters, which are all animated (two vignettes with tomato and two vignettes with triangle shape[3]; a total of four animated vignettes) (Allen et al., 2007; Özyürek et al., 2015; Özyürek et al., 2001). Grounds, i.e. the goal or source of the Motion or the surface upon which the Motion takes place, included trees (two vignettes), hills/slopes (four vignettes), a triangle (two vignettes), a tomato (two vignettes), a car (one vignette), and a girl (one vignette).

Some of the vignettes have more than one Ground. For example, in Figure 1 below, there are four possible Grounds, namely triangle, tree, slope/hill, and the top of the slope/hill. These Grounds can function as the Goal (tree), Surface (slope/hill), or Source of Motion

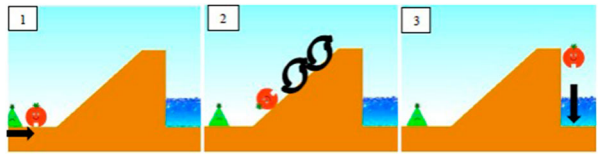


Figure 2. Stills from a vignette that shows three motion events.

(triangle, top of the slope/hill). For the present study, we did not differentiate between the three Ground types, and when any of these Grounds was mentioned in the vignette description, it was analysed as expressing a Ground. Critically, for this study, the same criteria were used for both signing and speaking groups.

Each of the vignettes featuring geometric characters (thus only four in number) shows three movement events: an entry event, a target motion event, and a closing event. For example, the first still in Figure 2 depicts a triangle entering the scene from the left and bumping into a tomato (i.e. the entry event). In the second still, the tomato rolls up the hill (i.e. the target motion event). In the third still, the tomato falls into the water and bobs up and down (i.e. the closing event). The other four vignettes show people engaged in a motion event and do not have entry and/or closing events, which makes their descriptions less complex than the tomato and triangle events (Figure 3) (see the Appendix for stills from all vignettes used in the study).

Procedure

In data collection sessions, signers/speakers were asked to sit opposite the addressee, who was a deaf or hearing adult confederate, depending on the language condition. The signer/speaker and addressee sat on opposite sides of a low table so the hand movements from the waist up could easily be seen, and a laptop sat on the table, with the screen facing the signer/speaker and away from the addressee. Participants were shown a series of short video vignettes on the laptop screen and then asked to describe what happened in each vignette to their interlocutors, who purportedly had not seen them before. The task of the addressee was to listen to/watch the descriptions and ask clarification questions, if

Figure 3. A still from a vignette that shows real people and does not have an entry/closing event.

necessary, after the participants finished their descriptions.

Data coding

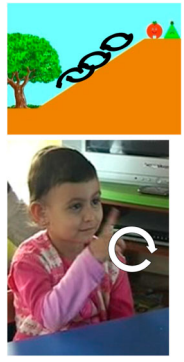
There was a total of 480 vignette descriptions (240 from each language). To exclude uncontrollable variation, descriptions in which participants repeated the description were not taken into account. Thus, one description from each participant was coded in the current study. For the vignettes with the tomato and triangle, participants sometimes did not describe the target motion event, but rather described the entry and/or the closing events only (see Figure 2). Such cases were not included in the current analyses. For the simple vignettes with fictive and real human figures, the whole description was analysed (Figure 3).

ELAN, an annotation software specifically designed for multimodal language data (Wittenburg et al., 2006), was used to annotate the signed descriptions. For each vignette description in TİD, all signs were transcribed using Turkish and English glosses on separate tiers for the left and right hand. The glosses were written by two hearing researchers who have good knowledge of TİD and checked by a deaf native signer of TİD. For Turkish, the data in each vignette description were coded by a Turkish-speaking research assistant.

For all vignette descriptions, we first identified the descriptions containing Motion encodings. We further coded the data for Path and Manner components and Ground omissions within the vignette descriptions in which Motion had been mentioned. Thus, in vignettes where Motion was mentioned, there were descriptions that express Path only, Manner only or both Path and Manner. In these descriptions we further identified the encoding of other components such as Figure and Ground. Descriptions lacking mention of Motion (e.g. introducing only the Figure and the Ground such as "there is a boy and a girl") were not included in the analysis.

In (3a), a child signer of TİD only encodes Motion with Manner (circular movement made with the index finger); her description lacks Path, Figure and Ground information. In (3b), a child speaker of Turkish refers to Motion with Path as well as Ground. Descriptions can also consist of all components (Motion, Path, Manner, Figure, Ground). In (3c), a child signer of TİD mentions the tomato (Ground, still 1) and triangle (Figure, still 3) and also indicates that the triangle hops up towards the tomato (Motion with both Manner and Path, still 4). Here, the signer holds his right hand in place (indicated as HOLD in the gloss) to refer to the tomato (Ground), which functions as the anchor point in the signing space. Similarly, in (3d), a child speaker of Turkish encodes all event components in their description.

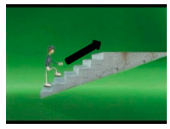
(3a) Encoding Motion with Manner only in TİD



(4;2)

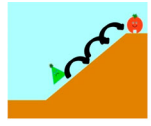
LH:
RH: ROLL
"[Someone] is rolling."

(3b) Encoding Ground and Motion with Path in Turkish



Merdiven + den çık + tı (4;5).
Stair + ABL ascend + PAST
"[Someone] went up the stair."

(3c) Encoding Motion with Path and Manner as well as Figure and Ground in TİD



(7;11)

LH CL(triangle)$_{loc}$ CL(triangle)$_{hop\_up}$
RH: TOMATO CL(tomato) .....HOLD.....
"There is a tomato here. There is a triangle here. Triangle hops up to the tomato."

(3d) Encoding Motion with Path and Manner as well as Figure and Ground in Turkish

Adam dağ + dan yuvarlan + arak aşağı in + di (6;6).
Man mountain + ABL roll + CONN down fall + PAST
"The man went down the mountain rolling."

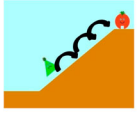
Encoding Path and Manner in TİD and Turkish

The mention of Manner and/or Path components fell into three categories. The first category, "Both Path and Manner", refers to the cases whereby participants encoded both Path and Manner. For example, in (4a, still 5), an adult signer of TİD encodes the Motion of the Figure (i.e. triangle) on her left hand with a slanted upwards movement (i.e. Path) combined with small up and down movements (i.e. Manner). In (4b), an adult speaker of Turkish uses a Path verb (i.e. ascend "*çık*") as well as a Manner verb (i.e. roll "*yuvarlan*").

The next category, "Path only", indicates the encoding of only Path information; these are also the cases that will be analysed as "Manner omissions" in the following sections. In (4c), the right hand of a child signer of TİD shows the upwards movement without further encoding Manner, Figure or Ground information in her description. In (4d), the description of an adult Turkish speaker encodes Motion and Path (i.e. fall "*düş*") as well as Figure, while information on Manner and Ground is lacking.

The final category, "Manner only", denotes use of only a Manner element (i.e. Path omissions). In (4e), an adult signer of TİD encodes in her description Ground (i.e. stone step, stills 1 & 2) and Motion plus Manner (index fingers of both hands with circular movements, still 3) information without referring to a Figure or Path of Motion. In (4f), a child speaker of Turkish refers to Manner of Motion (i.e. roll "*yuvarlan*") only, thus excluding from their description information about Ground, Figure and Path of Motion.

(4a) Encoding both Path and Manner in TİD



LH: TRIANGLE STONE SLOPE. -----HOLD----- $CL(triangle)_{jump\_up}$
RH: TRIANGLE STONE SLOPE $BALL_{loc}$ -----HOLD-----
"There is a triangle. There is a stone slope. The ball is here. The triangle is jumping up towards the ball."

(4b) Encoding both Path and Manner in Turkish

Üçgen	domates + e	doğru	yuvarlan + arak	yukarı	çık+tı. (adult)
Triangle	tomato + DAT	towards	roll + CONN	up	ascend + PAST

"The triangle went up to the tomato while rolling."

(4c) Encoding Motion with Path only (Manner omission) in TİD

(8;2)

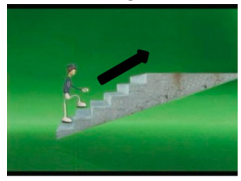
LH:
RH: MOVE_UP_SLANTING
"[Something] is slanting up."

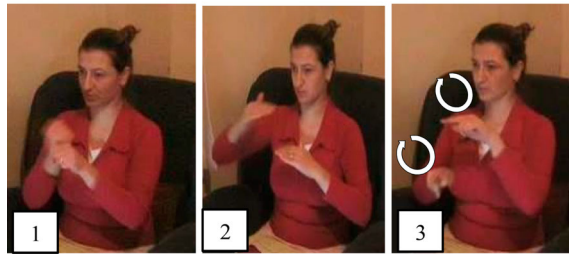
(4d) Encoding Motion with Path only (Manner omission) in Turkish

Domates	düş + tü. (adult)
Tomato	fall + PAST

"The tomato fell down."

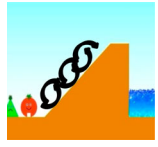
(4e) Encoding Motion with Manner only (Path omission) in TİD





LH: STONE STEP WALK
RH: STONE STEP WALK
"There is a stone step. [Someone] walks."

(4f) Encoding Motion with Manner only (Path omission) in Turkish



Yuvarlan + dı. (6;6)
Roll + PAST
"[Something] rolled."

Ground omissions in TİD and Turkish

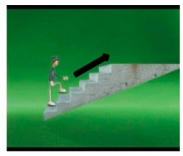
We further examined all narrations in which Motion was encoded and Ground was omitted. Even though we also coded for omissions of Figure, we did not compare them to the omission of other elements such as Ground, Path or Manner for two reasons. First, the number of narrations with Figure omissions was low: of all descriptions where Motion was encoded,

Figure encoding was lacking in only .07 of them in TİD (.03 for adults, .03 for school-age children, and .01 for preschool-age children) and .09 of them in Turkish (.06 for adults, none for school-age children, and .03 for preschool-age children), which did not allow for appropriate statistical comparisons across groups. Secondly, previous work mainly focused on Ground omissions by either speaking or signing children. Here we analyse this developmental pattern across these two groups of children in a systematic way and by using direct comparisons. To be in line with previous studies and to be able to compare our findings with them directly, we examined only Ground omissions (but not Figure omissions) and compared them to Manner and Path omissions in the descriptions where Motion was encoded.

Descriptions with Ground omissions in this analysis did not refer to any Ground in any way (i.e. no lexical signs or classifier predicates). In these cases, the signers mentioned the Motion (with Path and/or Manner), and the Figure could either be absent (see 4c) or mentioned (using either lexical signs or via classifier constructions). As noted earlier, some vignettes had two or more Grounds (Figure 1), and some of them could function as Source, Surface or Goal of Motion. Here we did not differentiate between them and coded any mention of Ground as Ground included (i.e. not omitted). If there was more than one Ground, all Grounds were coded, but the number of Grounds expressed was not included in the relevant analysis.

An example is (5) in which a TİD signer describes the vignette in which a man is going up the stairs. As shown in the still 2, index and middle fingers of his right hand, which are in the downwards position, wiggle and move up in the signing space, thus referring to both Path and Manner of Motion. He never mentions the Ground (i.e. the stairs) in his description. Such expressions were considered as Ground omissions because there is no explicit mention as to whether the Figure was moving on a Ground, away from a Ground, or towards a Ground. Note that signers in this study often encoded both the man and the stairs using their lexical signs, which were followed by a classifier construction, in which a slightly tilted back of a hand represents the Ground (i.e. stairs in this example) while the upside-down wiggling index and middle fingers of the other hand move up on it. In (4d), an adult Turkish speaker describes the vignette where the tomato is rolling down the hill/slope without mentioning any of the possible Grounds (i.e. the triangle, the hill/slope, or the tree).

(5) Ground omission in TİD



LH:
RH: MAN $CL(man)_{walk\_up}$
"There is a man. He is walking up."

Results

To understand whether and to what extent language modality modulates encoding motion events by adults and children in TİD and Turkish, we used a mixed-effects logistic regression model (Jaeger, 2008) with random intercepts for Participants and Items. To this end, we used the lme4 package (Bates et al., 2015) in R (version 1.1.26; R Core team, 2020) with the optimiser *bobyqa* (Powell, 2009). This approach allowed us to take into account the random variability that is due to having different participants and items. We did not include random slopes in any of the models because doing so either failed to increase model fit or resulted in convergence failure. For the different analyses reported below, a step-wise variable selection procedure was conducted for each model, and non-significant predictors were removed to obtain the most parsimonious model. To compare models, likelihood ratio tests that compared the goodness of fit were performed using the *anova* function in the *base* package (R Core Team, 2020). In this way, the final model was selected by checking whether the *p*-value from the likelihood ratio test was significant.

Encoding motion in TİD and Turkish

All the analyses of Path and Manner encoding as well as Ground omissions were performed for the vignette descriptions in which the Motion is expressed. Therefore, we first calculated the subject-based mean proportions of Motion encoding (with Path only, Manner only, or both) in all vignette descriptions ($n = 240$ for each language). We conducted a mixed-effects logistic regression model with mention of Motion in vignette descriptions as the binary dependent variable (0 = no, 1 = yes) by adding the step-wise fixed effect factors Age (Adult, School age, Preschool age) and Language (TİD, Turkish) to the

Table 2. Details for the model predicting whether Motion is mentioned in a vignette description.

Fixed effects	*β*	*SE*	*z*	*p*
Intercept	11.58	119.03	.097	.92
$Age_{school\text{-}age}$ vs Age_{adult}	−9.04	119.03	−.076	.94
$Age_{preschool\text{-}age}$ vs Age_{adult}	−9.31	119.03	−.078	.94
$Language_{TR}$ vs $Language_{TID}$	−17.63	238.08	−.074	.94
$Age_{school\ age}$ vs Age_{adult} $Language_{TR}$ vs $Language_{TID}$	15.50	238.08	.065	.95
$Age_{preschool\ age}$ vs Age_{adult} $Language_{TR}$ vs $Language_{TID}$	16.37	238.08	.069	.95
Random Intercepts	*var*	*SD*		
Participant (Intercept)	.97	.98		
Item (Intercept)	1.58	1.25		

For the fixed effects, estimates (β), standard errors (SE), z-values and *p*-values are given. For the random effects, variance (var) and standard deviations (SD) are reported.
Significance codes: * 0.05, ** 0.01, *** 0.001.
Encoding motion ~ Age * Language + (1 | participant) + (1 | item).

baseline model (including participants and items as random intercepts). The fixed effect of Language was analysed with the numeric contrasts (Helmert contrast), and for the fixed effect of Age we used treatment coding.

The model did not reveal any main effects of Age and Language, and there was no interaction between them (Table 2). In both languages, children encoded Motion as frequently as adults. Additionally, TİD signers and Turkish speakers were similar to each other in how frequently their narrations included the mention of Motion ($M = .94$ for TİD vs $M = .79$ for Turkish) (Figure 4).

Encoding Path and Manner in TİD and Turkish

In descriptions where Motion is mentioned ($n = 226$ for TİD and $n = 189$ for Turkish), we further analysed the expression of Path and Manner. To this end, we first used a similar mixed-effects logistic regression model with the mention of both Path and Manner in vignette descriptions as the binary dependent variable (0 = no, 1 = yes) by adding the step-wise fixed effect factors Age (Adult, School age, Preschool age) and Language (TİD, Turkish) to the baseline model (including participants and items as random intercepts). The fixed effect of Language was analysed with the numeric contrasts (Helmert contrast), and for the fixed effect of Age we used treatment coding. This analysis yielded main effects of Age and Language without any interaction between them (Table 3). Both child age groups, regardless of the language modality being acquired, produced motion narrations that mentioned both Path and Manner less frequently than the adults did. Moreover, Path and Manner were encoded more frequently in TİD than in Turkish (Figure 5).

As a further analysis, we also compared each age group across language by using the package *emmeans* (Length, 2019; Searle et al., 1980). We found that TİD signers in each age group encoded both Path and Manner more than the corresponding age groups of Turkish speakers (Adults: $\beta = 1.76$, $SE = .52$, $z = 3.38$, $p < .01$; School-age children: $\beta = 2.42$, $SE = .55$, $z = 4.41$, $p < .001$; Preschool-age children: $\beta = 1.62$, $SE = .52$, $z = 3.11$, $p < .05$).

In the cases where signers and speakers did not mention both Path and Manner, their descriptions lacked either Path or Manner. As shown in Table 4 below, in their target event descriptions, Turkish speakers mainly dropped Manner, which is expected considering that Turkish is a verb-framed language, where the main verb is reserved for Path information and Manner encoding is optional.

As mentioned earlier, in order to be in line with previous studies that compare only sign vs speech, the analysis of co-speech gestures in Turkish speakers is outside the scope of this study (Galvan & Taub, 2006; Sallandre et al., 2018; Taub & Galvan, 2001).

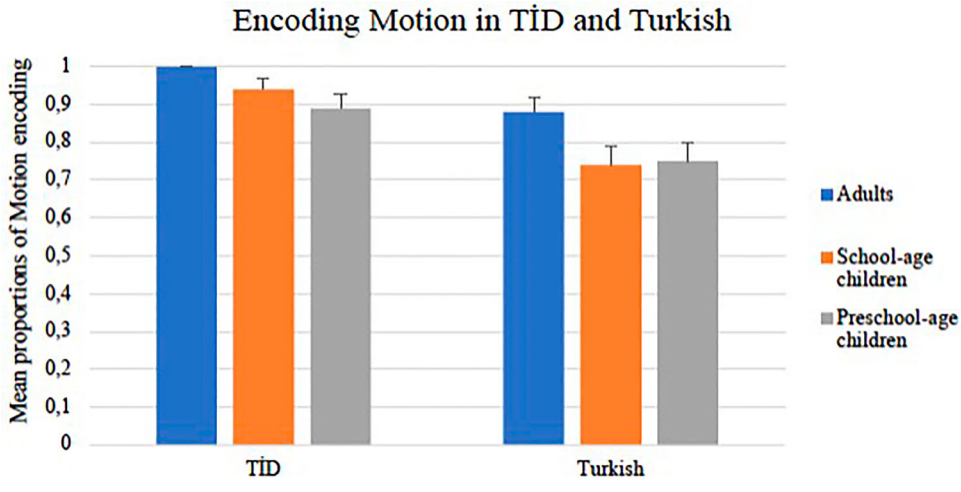


Figure 4. Mean proportions of Motion encoding by TİD signers and Turkish speakers across age groups.

Table 3. Details for the model predicting whether both Path and Manner will be encoded in a motion event narration.

Fixed effects	*β*	*SE*	*Z*	*p*
Intercept	.95	.35	2.70	<0.01**
$Age_{school\text{-}age}$ vs Age_{adult}	−1.34	.37	−3.58	<0.001***
$Age_{preschool\text{-}age}$ vs Age_{adult}	−1.41	.37	−3.84	<0.001***
$Language_{Turkish}$ vs $Language_{TID}$	−1.76	.52	−3.38	<0.001***
$Age_{school\ age}$ vs Age_{adult} $Language_{Turkish}$ vs $Language_{TID}$	−.65	.74	−.88	.38
$Age_{preschool\ age}$ vs Age_{adult} $Language_{Turkish}$ vs $Language_{TID}$	.16	.72	.20	.84
Random Intercepts	*Var*	*SD*		
Participant (Intercept)	.46	.68		
Item (Intercept)	.46	.68		

For the fixed effects, estimates (β), standard errors (SE), z-values and *p*-values are given. For the random effects, variance (var) and standard deviations (SD) are reported.
Significance codes: * 0.05, ** 0.01, *** 0.001.
Encoding both Path and Manner ~ Age * Language + (1 | participant) + (1 | item)

However, we still checked the event narrations of Turkish speakers in each age group to see if they encoded Path or Manner in their co-speech gestures when their speech lacks this information. Of 103 narrations without the mention of the Manner, only 11 included the use of co-speech gestures expressing Manner (two from school-age children and nine from preschool-age children). Of nine descriptions that lack Path encoding in speech, Path was encoded by a co-speech gesture in only one (preschool-age) child.

Ground omissions in TİD and Turkish

For descriptions where Motion is mentioned ($n = 226$ for TİD and $n = 189$ for Turkish), we further analysed the proportion of Ground omissions in a mixed-effects logistic regression model with the Ground omission as the binary dependent variable (0 = no, 1 = yes) by adding the step-wise fixed effect factors Age (Adult, School age, Preschool age) and Language (TİD, Turkish) to the baseline model (including participants and items as random intercepts). The fixed effect of Language was analysed with the numeric contrasts (Helmert contrast), and for the fixed effect of Age we used treatment coding. As a result, we observed a main effect of age but not language, with no interaction between the two factors (Table 5). Both child age groups, regardless of the language modality being acquired, omitted Grounds more frequently than adults. Moreover, TİD and Turkish narrations were similar with respect to Ground omissions (Figure 6).

Table 4. Mean proportions and (SE)s of omitting Manner and Path information in target event descriptions with Motion encoded by different age groups across languages.

	TİD		Turkish	
	Manner omission	Path omission	Manner omission	Path omission
Adults	.10 (.04)	.07 (.03)	.47 (.06)	.00 (.00)
School-age children	.24 (.05)	.09 (.03)	.69 (.06)	.08 (.04)
Preschool-age children	.15 (.04)	.28 (.05)	.55 (.06)	.18 (.05)
Total	.16 (.02)	.14 (.02)	.56 (.04)	.08 (.02)

Comparing types of omissions in motion event encodings in TİD and Turkish

In the previous sections, we reported how frequently signers and speakers encoded Motion as well as both Path and Manner in descriptions in which Motion is

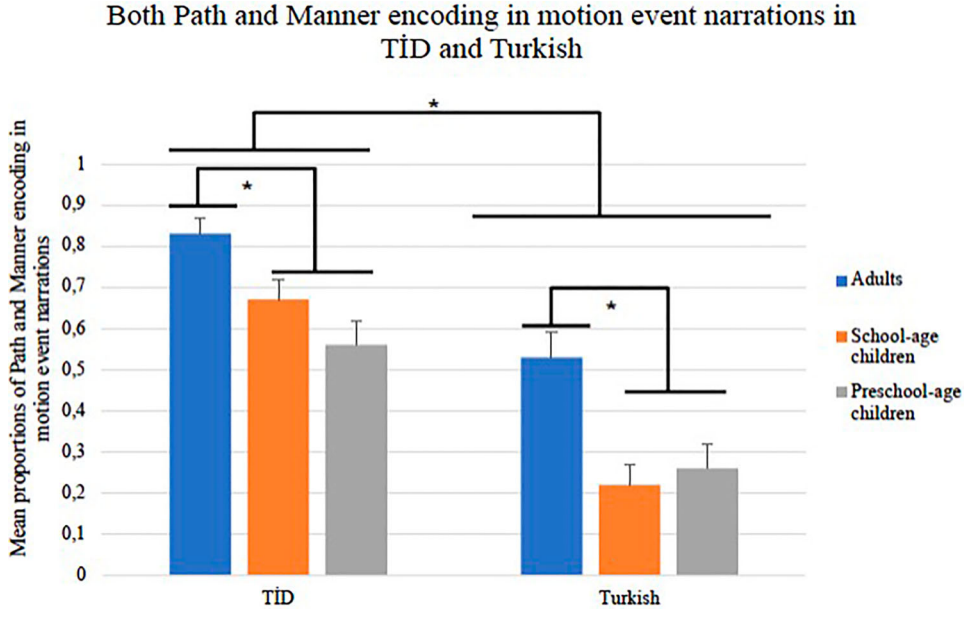


Figure 5. Mean proportions of both Path and Manner encoding by TİD signers and Turkish speakers across age groups.

Table 5. Details for the model predicting whether Ground is omitted in a motion event narration.

Fixed effects	*β*	*SE*	*Z*	*p*
Intercept	−3.52	.58	−6.10	<0.001***
$Age_{school\text{-}age}$ vs Age_{adult}	1.92	.53	3.60	<0.001***
$Age_{preschool\text{-}age}$ vs Age_{adult}	2.68	.53	5.03	<0.001***
$Language_{TR}$ vs $Language_{TID}$	−.78	.91	−.86	.39
$Age_{school\text{-}age}$ vs Age_{adult} $Language_{Turkish}$ vs $Language_{TID}$	.48	1.06	.46	.65
$Age_{preschool\text{-}age}$vs Age_{adult} $Language_{Turkish}$ vs $Language_{TID}$	.24	1.03	.24	.83
Random Intercepts	*var*	*SD*		
Participant (Intercept)	.32	.56		
Item (Intercept)	.75	.86		

For the fixed effects, estimates (β), standard errors (SE), z-values and *p*-values are given. For the random effects, variance (var) and standard deviations (SD) are reported.
Significance codes: * 0.05, ** 0.01, *** 0.001.
Omitting Ground ~ Age * Language + (1 | participant) + (1 | item).

encoded, and we also reported a separate analysis showing how frequently Grounds were omitted in these descriptions. In this section, we present our findings on how descriptions with Manner omissions compare to descriptions with Path omissions in TİD and Turkish. We also report how Ground omissions compare to omissions of Path omissions and Manner omissions. Note that Motion is encoded in all these descriptions.

As shown in Table 6 below, TİD signers omitted Path or Manner in quantitively similar ways ($t = .47$, df = 225, $p = .63$), while Turkish speakers omitted Manner more than Path ($t = 10.20$, df = 188, $p < .01$), which is not surprising since Turkish is a verb-framed language where Manner encoding is optional. Further statistical comparisons across these languages also revealed that both signers and speakers omitted Path in similar amounts ($t = -1.97$, df = 411.63, $p = .05$). However, TİD and Turkish differed for Manner omissions: speakers omitted Manner more often than signers ($t = 9.20$, df = 341.93, $p < .01$).

We also compared whether omissions of Ground are more frequent than omissions of Path or Manner in TİD and Turkish since it might be the case that the presence of multiple Grounds in vignettes could have created competition for the participants, who would have to choose among several possible candidates for Ground. As reported above, there was no effect of language on Ground omission. That is, the frequency with which signers and speakers dropped them from their motion event descriptions was similar. In TİD, signers' descriptions were similar in terms of Ground versus Manner omissions ($t = -1.57$, df = 225, $p = .12$), though signers showed a stronger tendency (although at a marginal level) to drop Grounds relative to Path in their descriptions ($t = -2.11$, df = 225, $p = .04$). In Turkish, speakers dropped Grounds more often than Path ($t = -2.38$, df = 188, $p = .02$) but less often than Manner ($t = 9.04$, df = 188, $p < .01$).

Table 6. Mean proportions and (SE)s of omitting Ground, Path, and Manner information in target event descriptions with Motion encoded in TİD and Turkish.

	TİD	Turkish
Ground omissions	.23 (.03)	.15 (.03)
Path omissions	.14 (.02)	.08 (.02)
Manner omissions	.16 (.02)	.56 (.04)

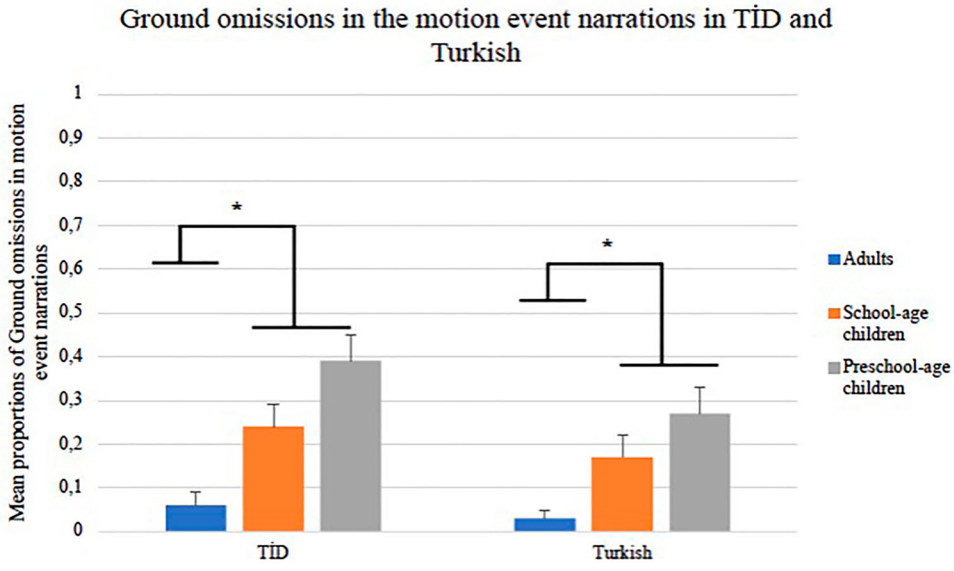


Figure 6. Mean proportions of Ground omissions by TİD signers and Turkish speakers across age groups in event descriptions in which Motion is mentioned.

Discussion

In this study, we compared language development between signing and speaking children in the domain of motion event expressions. We systematically collected data to test whether the reported tendency for speaking children to be less informative in their event descriptions than adults also holds for sign language acquiring children by using the same tasks under similar conditions.

We entertained the following predictions. The visual-spatial modality of sign languages may have a facilitating effect on learning to encode event components since signers were found to encode them more frequently than speakers (Galvan & Taub, 2006; Sallandre et al., 2018; Taub & Galvan, 2001). If this is the case, then TİD-signing children might not show omissions and become adult-like earlier than their Turkish-speaking peers due to the iconicity available in the signing space and the language forms used for event representations in sign languages. However, following the classic language acquisition models in which language builds upon domain-general principles that are widely shared by members of different language communities and form the basis of language development in children (e.g. Clark, 2004; Gleitman, 1990; Jackendoff, 1996; Papafragou & Selimis, 2010; Papafragou et al., 2002; Pinker, 1989), it is also possible that both groups of children will exhibit similar developmental patterns in encoding these components and omit event components more than adults. Finally, we might expect to observe different effects of language modality for different event components. Considering previous findings with signing children (Fenson et al., 1994; Naigles et al., 1998; Özyürek et al., 2015; Sallandre et al., 2018; Zheng & Goldin-Meadow, 2002), we might see a sign advantage for the development of Path and Manner components but not for the expression of Ground, since children have been found to show a tendency to omit Ground in their event narrations regardless of language modality (Engberg-Pedersen, 2003; Hickmann, 2003; Morgan et al., 2008; Slobin et al., 2003; Tang, 2003). This prediction would then suggest that learning to encode Ground is strongly driven by domain-general factors and thus insensitive to the effects of modality, while learning Path-Manner encoding can be modulated by language-specific factors such as the iconicity available in linguistic forms.

Our study revealed that children presented similar tendencies in their event encoding regardless of the language they were acquiring. Both signing and speaking children encoded Motion as frequently as the adults of the same language but used both Path and Manner less and omitted Ground more than the adults in that language. These findings suggest that with regard to the expression of Manner, Path and Ground components, domain-general developmental factors (e.g. cognitive, pragmatic) rather than language modality play a role in the development of event narration (e.g. Clark, 2004; Gleitman, 1990; Jackendoff, 1996; Papafragou et al., 2002; Papafragou & Selimis, 2010; Pinker, 1989). Our findings also provide further support for Supalla's (1982) claim earlier that rather than representing motion events as holistic units in their productions, signing children's acquisition of motion event components happens in a piece-by-piece (i.e. event component units) fashion – similar to the piecemeal acquisition of language forms in spoken languages. The only modality effect we found was that signing children might be more advantaged in at least Path and Manner encoding since they referred to them more frequently than their speaking peers – albeit less frequently than the adults of the same language. While some of these findings are in line with previous research, some are not, as we discuss further below.

Encoding motion

We found that while describing target events, signers and speakers did not differ in how frequently they encoded Motion in their event narrations. This result contradicts the findings of earlier studies claiming that the visual-spatial modality of sign languages encourages signers to encode Motion – the most salient component of a motion event – more than speakers. For example, studying the Motion information in the "Frog, where are you?" (a wordless picture story, Mayer, 1969) narrations elicited in ASL and English, Galvan and Taub (2006) found that ASL signers specifically mentioned instances of Motion more often than English speakers did. They attributed this finding to the affordances of the visual-spatial modality in expressing Motion in an analogue way to the real event itself.

However, most of the earlier studies analysed longer discourse by signers – unlike the present study, which examined short vignette descriptions. In previous studies (Berman & Slobin, 1994), signers might have encoded Motion more frequently while describing events from a longer and static representation of events than from shorter animated videos. Corroborating evidence for the notion that the way in which events are presented (i.e. visually or non-visually) affects linguistic expression also comes from a recent study which showed that speakers provide richer linguistic information about motion events when presented with videos of these events (Mamus et al., 2021). It is thus possible that the speakers in our study behaved similarly to signers in terms of how much information they

encoded when describing the motion events since these events were shown in animated videos. Further comparisons of longer versus shorter motion event narrations as well as use of picture stories versus animated videos to elicit data might also reveal an effect of how frequently Motion is encoded in event narrations. More in line with our findings with children and adults is an earlier study that used the same stimuli and found that homesigning deaf children (i.e. deaf children who are not exposed to a conventional sign language) could still convey as much Motion information as their speaking peers when they depicted motion events with their gestures (Gentner et al., 2013). Thus, encoding Motion might be a fundamental cognitive-domain task that is insensitive to language input or modality of language.

Encoding both Path and Manner

Most relevant to our research question is the finding that the richer encoding of Path and Manner in TİD did not help signing children to encode both to a degree that is similar to what is seen in signing adults. On the developmental level, our results show that children, regardless of the language modality being acquired, encoded both Path and Manner in their motion event descriptions less frequently than adults, even at the age of nine. Thus, encoding both types of information seems to be challenging for children and they have a tendency to omit one of these components, as earlier studies have also found (Özyürek et al., 2008; Papafragou et al., 2006). This might be due to the difficulties in representing both of these components in a motion event description. It seems that despite the visually motivated mappings of events on linguistic forms in sign languages, signing children derive no benefit from this modality advantage and follow a developmental path similar to that of their speaking peers. This finding supports the view that domain-general developmental factors (e.g. cognitive, pragmatic) and human experience play a strong role in first language acquisition (e.g. Clark, 2004; Gleitman, 1990; Jackendoff, 1996; Papafragou et al., 2002; Papafragou & Selimis, 2010; Pinker, 1989).

Our findings also show that even though both signing and speaking children used both Path and Manner less frequently than adults, signing children still encoded both of these components more frequently than their speaking peers. This could indeed be an effect of the visual-spatial modality of sign languages. However, it is also possible that this could be the result of sign languages' being closer to satellite-framed languages in encoding both Path and Manner rather than the effect of the visual-spatial modality per se. That is, it is possible that the TİD pattern resembles the pattern found in satellite-framed languages (e.g. English), which allows speakers to combine Path and Manner more easily (i.e. in one verbal clause), whereas Turkish, being a verb-framed language, allows speakers to focus on Path more than on Manner. A similar pattern of Path and Manner encoding was also reported for LSF, where signers also combine both Path and Manner more frequently than French speakers, who tended to produce more Path-only responses than LSF signers (Sallandre et al., 2018).

It is important to note the disagreement regarding how to classify sign languages according to Talmy's (1985) classification of languages as verb-framed or satellite-framed. Investigating the signed narrations of motion events in ASL, Supalla (1990) treats the Manner verb as the main verb, thus suggesting that it is a satellite-framed language, in which signers produce constructions similar to English sentences like "a boy jumps up and down in circle". However, Slobin and Hoiting (1994) argue that the Manner verb cannot be the main verb in ASL because it is optional and they go on to describe sign languages as "complex verb-framed", suggesting a new category in Talmy's typology. In their view, sign languages are more akin to verb-framed languages because the visual-spatial modality requires the mention of spatial relations (including Path) when encoding Motion. Encoding Motion is also complex because it enables the simultaneous representation of other components (such as Figure and Manner). Other factors, such as the type of Path, might also influence how different components of Motion are encoded in sign versus spoken languages. Sallandre et al. (2018), for example, report differences in encoding both Path and Manner in LSF (satellite-framed) versus English (satellite-framed) via indirect comparisons, depending on the type of the Path of the Motion. For the motion events that include "up" or "down" as their Path, LSF signers encoded both Path and Manner more frequently than English speakers. But for the "across" Path type, the pattern reversed: English speakers produced more Path and Manner combinations than LSF signers.

Regardless of the typological classification of Motion encoding in sign languages (verb- versus satellite-framed), there seems to be evidence for the role of language modality rather than language typology in learning to encode Motion. In an earlier study, Bunger et al. (2012) found that only 26% of all the motion event descriptions elicited from English-acquiring children (4 years of age) included both Path and Manner. In our study, 26% of the Turkish-speaking children and 56% of the TİD-signing children between the ages of 4–6, used both Manner and Path. Speaking children seem to be challenged more than their signing peers

in encoding both Path and Manner despite being exposed to a satellite-framed language such as English. A direct comparison of, for example, ASL and English can be informative in this regard. Özyürek et al. (2015) also provided corroborating evidence for the facilitating effect of language modality in encoding both components by providing evidence from Turkish homesigning children of similar ages. Their data showed the tendency in these children to express both Path and Manner more than Turkish-acquiring children. This finding itself is interesting given that many studies of similarly-aged speaking children describing motion events in speech have found that they tend to mention only one component of a motion event rather than mentioning both (e.g. Allen et al., 2007; Bunger et al., 2012; Hickmann, 2003; Özyürek et al., 2008; Papafragou & Selimis, 2010). Thus, it might be easier to convey both Path and Manner in the manual modality for children, which supports an iconic mapping between form and meaning – even though the visual-spatial modality does not eliminate the general tendency in children's Path and Manner omissions. Thus, for children the effect of language modality is not totally neutral in learning to express motion events, since it may play a lesser role in acquiring the ability to conceptualise all components of a motion event.

A further look into our data shows a tendency for preschool-age signing children to produce responses with an omitted Path (thus encoding only Manner) more than responses with an omitted Manner (thus encoding Path only), whereas this pattern was reversed for speaking children. Preschool-age speaking children dropped Manner more than Path in their descriptions. The Manner tendency found in signing children's productions is in line with the findings of a previously reported study on LSF, where it was suggested that iconicity highlights Manner due to the signer's ability to enact the Figure's action (i.e. Figure embodiment), which results in a motion expression that typically includes the Manner component (Sallandre et al., 2018). This finding, however, stands in contrast to what is observed in the data coming from homesigners, who showed a stronger preference for Path gestures than for Manner gestures in their productions for similar motion events (Goldin-Meadow et al., 2008; Özyürek et al., 2015; Zheng & Goldin-Meadow, 2002) and in the data coming from ASL-acquiring children (Newport, 1981; Newport & Supalla, 1980). At this point, there is not enough evidence to support either of the conclusions about the preference of Path over Manner for modality.

The preference for the descriptions with Manner omitted (Path only) by Turkish-speaking children suggests an early tuning into the adult pattern observed in Turkish, which, as a verb-framed language, pays more attention to Path information than Manner information. A similar tendency was also observed for the acquisition of other verb-framed languages, where children preferred Path-only descriptions more than Manner-only descriptions from 4 years of age onwards, and the frequency of encoding both components was low but steadily increasing with age (Allen et al., 2007; Hickmann, 2003; Hickmann et al., 2009; Papafragou & Selimis, 2010). Thus, while omission of Manner or Path might be a general tendency, which one is preferred might be modulated by modality and language type – a possibility requiring more research.

Ground omissions

Signing and speaking children in both age groups omitted Grounds in their descriptions more frequently than adults. This finding seems to be in line with the previous studies that report the omission of Ground to be a pervasive feature of children's motion event narrations (Engberg-Pedersen, 2003; Hickmann, 2003; Morgan et al., 2008; Slobin et al., 2003; Supalla, 1982; Tang et al., 2007). This finding also suggests that while the visual-spatial modality encourages the mention of Path and Manner, we do not see such an effect for the mention of Ground. Despite the findings on the early recognition of Grounds (e.g. Göksun et al., 2009), the omission of Grounds might be related to stimuli-related factors. In the current study, the presence of several Grounds in the vignettes might have led to this difference since participants, especially children, could not decide which Ground to include, thus not mentioning any of them in their descriptions. It is also important to note that the Path and Manner of Motion were highlighted in the vignettes, thus driving more attention to their mention than to the other components.

Comparing omissions in motion event encodings in TİD and Turkish

Comparing the omission of different event components (Ground, Path, Manner) has also revealed further insights into the developmental patterns observed in TİD and Turkish. These two languages differ in terms of which event component is omitted most frequently: in TİD, Ground was omitted more often than were Path and Manner, which were mainly mentioned by the signers. However, in Turkish Manner was the most omitted component, an unsurprising finding given that Turkish is a verb-framed language, in which Manner information is optional. This observed pattern could therefore be an

effect of typological difference in encoding Motion, with Turkish being a verb-framed language and TİD being more akin to a satellite-framed language. However, it might also be due to the visual-spatial modality of sign languages as well as linguistic forms (i.e. classifier constructions) employed mostly to encode Motion, which encourage mention of the Path and Manner of a Motion (together with Figure information) in one construction, usually articulated with one hand. Adding Ground information in such a construction usually means employing the other hand, which increases the complexity of the narration, thus discouraging the signers from additionally encoding Ground in their narrations. These interpretations pertain to the comparative omissions of each component, but Grounds were equally likely to be omitted in signers and speakers.

Conclusion

In sum, we investigated the role of language modality in learning to encode different components of a motion event in a sign (TİD) and a spoken (Turkish) language. Specifically, we investigated children's omission of event components and whether this is modulated by the language modality being used. Our findings in general can be interpreted as indicating a neutral role of language modality in the acquisition of motion event encodings and for the omission of event components, suggesting that children learn to encode events in a piecemeal fashion. We have also confirmed a bias for both Manner and Path being encoded by signing children more often than speaking children, possibly due to the affordance of using the visual-spatial modality for linguistic expressions. Our study therefore provides evidence from the perspective of sign language acquisition in support of the claim that the development of linguistic expressions of motion events is modulated by both domain-general and language-specific factors, conforming to previous claims (e.g. Allen et al., 2007).

Notes

1. Full versions of the abbreviations used for morphological categories in the whole paper are: "CONN" for connectives, "ABL" for ablative case marker, "PAST" for past tense marking, "DAT" for dative case marker, "CL" for classifier predicates.
2. However, in our analysis of the elicited descriptions we checked whether co-speech gestures, if used at all, changed omission patterns.
3. Participants were free to choose their own names for the characters. Here, we refer to them as tomato and triangle.

Acknowledgements

We thank our deaf assistants Sevinç Akın and Şule Kibar and hearing assistant Hükümran Sümer for their help collecting, annotating, and coding data. We would also like to thank Prof Pamela Perniss (University of Cologne) and Dr Inge Zwitserlood (Radboud University) for their intellectual contributions to this study.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work has been supported by a H2020 European Research Council (ERC) Starting Grant and Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO VICI) grant awarded to the second author.

ORCID

Aslı Özyürek http://orcid.org/0000-0002-0914-8381

References

Adams, A. M., & Gatherhole, S. E. (2000). Limitations in working memory: Implications for language development. *International Journal of Language & Communication Disorders*, *35*(1), 95–116. https://doi.org/10.1080/136828200247278

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*(1), 16–48. https://doi.org/10.1016/j.cognition.2005.12.006

Arık, E. (2009). *Spatial language: Insights from signed and spoken languages* [Unpublished doctoral dissertation]. Purdue University, The USA.

Aronoff, M., Meir, I., Padden, C., & Sandler, W. (2003). Classifier constructions and morphology in two sign languages. In K. Emmorey (Ed.), *Perspectives on classifier constructions in sign languages* (pp. 53–84). Lawrence Erlbaum Associates.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Berman, R. A., & Slobin, D. (1994). *Relating events in narrative: A crosslinguistic developmental study*. Lawrence Erlbaum Associates.

Bunger, A., Trueswell, J. C., & Papafragou, A. (2012). The relation between event apprehension and utterance formulation in children: Evidence from linguistic omissions. *Cognition*, *122*(2), 135–149. https://doi.org/10.1016/j.cognition.2011.10.002

Clark, E. V. (2004). How language acquisition builds on cognitive development. *Trends in Cognitive Sciences*, *8*(10), 472–478. https://doi.org/10.1016/j.tics.2004.08.012

Cuxac, C., & Sallandre, M. A. (2007). Iconicity and arbitrariness in French Sign Language: Highly iconic structures, degenerated iconicity and diagrammatic iconicity. In E. Pizzuto, P. Pietrandrea, & R. Simone (Eds.), *Verbal and signed languages:*

Comparing structures, constructs and methodologies (pp. 13–33). Mouton de Gruyter.

De Beuzeville, L. (2006). *Visual and linguistic representation in the acquisition of depicting verbs: A study of native signing deaf children of Auslan* [Australian sign language]. Unpublished doctoral dissertation, University of Newcastle, Australia.

Engberg-Pedersen, E. (1993). *Space in Danish Sign Language: The semantics and the morphosyntax of the use of space in a visual language*. Hamburg.

Engberg-Pedersen, E. (2003). How composite is a fall? Adult's and children's descriptions of different types of falls in Danish sign language. In K. Emmorey (Ed.), *Perspectives on classifier constructions in sign languages* (pp. 311–332). Lawrence Erlbaum Associates.

Fenson, L., Dale, P. S., Reznick, J. S., & Bates, E. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, *59*(5), (Serial No. 242). https://doi.org/10.2307/1166093

Galvan, D., & Taub, S. (2006). The encoding of motion information in American Sign Language. In S. Strömqvist, & L. Verhoven (Eds.), *Relating events in narrative: Typological and contextual perspectives* (pp. 191–217). Lawrence Erlbaum Associates.

Gentner, D., Özyürek, A., Gürcanli, Ö, & Goldin-Meadow, S. (2013). Spatial language facilitates spatial cognition: Evidence from children who lack language input. *Cognition*, *127*(3), 318–330. https://doi.org/10.1016/j.cognition.2013.01.003

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, *1*(1), 3–55. https://doi.org/10.1207/s15327817la0101_2

Göksun, T., Hirsh-Pasek, K., & Golinkoff, R. M. (2009). Processing figures and grounds in dynamic and static events. In J. Chandlee, M. Franchini, S. Lord, & G. Rheiner (Eds.), *Proceedings of the 33rd annual Boston University conference on language development* (pp. 199–210). Cascadilla Press.

Goldin-Meadow, S., Chee So, W., 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*(27), 9163–9168. https://doi.org/10.1073/pnas.0710060105

Hickmann, M. (2003). *Children's discourse: Person, space, time across languages*. Cambridge University Press.

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. https://doi.org/10.1017/S0305000908009215

Hoiting, N., & Slobin, D. I. (2007). From gestures to signs in the acquisition of sign language. In S. D. Duncan, J. Cassell, & E. T. Levy (Eds.), *Gesture and the dynamic dimension of language: Essays in honor of David McNeill [Gesture Studies 1]* (pp. 51–65). John Benjamins. https://doi.org/10.1075/gs.1.06hoi.

Jackendoff, R. (1996). The architecture of the linguistic-spatial interface. In P. Bloom, M. Peterson, L. Nadel, & M. Garrett (Eds.), *Language and space* (pp. 1–30). The MIT Press.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or Not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446. https://doi.org/10.1016/j.jml.2007.11.007

Karadöller, D. Z., Sümer, B., Ünal, E., & Özyürek, A. (2020, November 5–8). *Sign advantage for children: Signing children's spatial expressions are more informative than speaking children's speech and gestures combined*. Talk presented at the 45th Annual Boston University (virtual) Conference on language development (BUCLD 45), Boston, MA, USA.

Karadöller, D. Z., Sümer, B., Ünal, E., & Özyürek, A. (under review). Sign advantage for children: Signing children's spatial expressions are more informative than speaking children's speech and gestures combined.

Length, R. (2019). emmeans: Estimated marginal means, aka least-squares means [Computer software]. https://CRAN.R-project.org/package=emmeans.

Levelt, P. (1989). *Speaking: From intention to articulation*. The MIT Press.

Mamus, E., Speed, L., Özyürek, A., & Majid, A. (2021). Sensory modality of input influences encoding of motion events in speech but not co-speech gestures. In T. Fitch, C. Lamm, H. Leder, & K. Teßmar-Raible (Eds.), *Proceedings of the 43rd annual conference of the Cognitive Science Society (CogSci 2021)* (pp. 376–382). Cognitive Science Society.

Mayer, M. (1969). *Frog, where are you?* Dial Press.

Morgan, G., Herman, R., Barriere, I., & Woll, B. (2008). The onset and mastery of spatial language in children acquiring British Sign Language. *Cognitive Development*, *23*(1), 1–19. https://doi.org/10.1016/j.cogdev.2007.09.003

Naigles, L. R., Eisenberg, A. R., Kako, E. T., Highter, M., & McGraw, N. (1998). Speaking of motion: Verb use in English and spanish. *Language and Cognitive Processes*, *13*(5), 521–549. https://doi.org/10.1080/016909698386429

Newport, E. L. (1981). Constraints on structure: Evidence from American Sign Language and language learning. In W. A. Collins (Ed.), *Minnesota symposia on child psychology*, Vol. 14 (pp. 1–22). Lawrence Erlbaum Associates.

Newport, E. L., & Supalla, T. (1980). The structuring of language: Clues from the acquisition of signed and spoken language. In U. Bellugi, & M. Studdert-Kennedy (Eds.), *Signed and spoken language: Biological constraints on linguistic form* (pp. 247–260). Dahlem Konferenzen. Verlag Chemie.

Özyürek, A., Furman, R., & Goldin-Meadow, S. (2015). On the way to language: Event segmentation in homesign and gesture. *Journal of Child Language*, *42*(1), 64–94. https://doi.org/10.1017/S0305000913000512

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*. 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*(4), 1040–1054. https://doi.org/10.1037/0012-1649.44.4.1040

Papafragou, A., Massey, C., & Gleitman, L. (2002). Shake, rattle, 'n' roll: The representation of motion in language and cognition. *Cognition*, *84*(2), 189–219. https://doi.org/10.1016/S0010-0277(02)00046-X

Papafragou, A., Massey, C., & Gleitman, L. (2006). When English proposes what Greek presupposes: The cross-linguistic encoding of motion events. *Cognition*, *98*(3), B75–B87. https://doi.org/10.1016/j.cognition.2005.05.005

Papafragou, A., & Selimis, S. (2010). Event categorisation and language: A cross-linguistic study of motion. *Language*

and Cognitive Processes, *25*(2), 224–260. https://doi.org/10.1080/01690960903017000

Perniss, P. M., & Özyürek, A. (2008). Representations of action, motion and location in sign space: A comparison of German (DGS) and Turkish (TID) sign language narratives. In J. Quer (Ed.), *Signs of the time: Selected papers from TISLR 8* (pp. 353–376). Signum Press.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. MIT Press.

Powell, M. J. (2009). *The BOBYQA algorithm for bound constrained optimization without derivatives* (Cambridge NA Report NA2009/06). University of Cambridge.

Pulverman, R., Golinkoff, R. M., Hirsh-Pasek, K., & Buresh, J. (2008). Infants discriminate manners and paths in non-linguistic dynamic events. *Cognition*, *108*(3), 825–830. https://doi.org/10.1016/j.cognition.2008.04.009

Pulverman, R., Hirsh-Pasek, K., Golinkoff, R. M., Pruden, S., & Salkind, S. (2006). Conceptual foundations for verb learning: Celebrating the event. In K. Hirsh-Pasek, & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 134–159). Oxford University Press.

Pulverman, R., Sootsman, J. L., Golinkoff, R. M., & Hirsh-Pasek, K. (2003). The role of lexical knowledge in non-linguistic event processing: English-speaking infants' attention to manner and path. In *B. Beachley, A. Brown, & F. Conlin (Eds.), Proceedings of the 27th annual Boston University conference on language development* (pp. 662–673). Cascadilla Press.

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Sallandre, M. A., Schoder, C., & Hickmann, M. (2018). Motion expressions in the acquisition of French Sign Language. In M. Hickmann, E. Veneziano, & H. Jisa (Eds.), *Sources of variation in first language acquisition: Languages, contexts, and learners* (pp. 365–390). John Benjamins.

Scerif, G., Karmiloff-Smith, A., Campos, R., Elsabbagh, M., Driver, J., & Cornish, K. (2005). To look or not to look? Typical and atypical development of oculomotor control. *Journal of Cognitive Neuroscience*, *17*(4), 591–604. https://doi.org/10.1162/0898929053467523

Searle, S. R., Speed, F. M., & Milliken, G. A. (1980). Population marginal means in the linear model: An alternative to least squares means. *The American Statistician*, *34(4)*, 216–221.

Slobin, D., & Hoiting, N. (1994). Reference to movement in spoken and signed languages: Typological considerations. In *Proceedings of the twentieth annual meeting of the Berkley Linguistics society: General session dedicated to the contributions of Charles J. Fillmore* (pp. 487–505). https://doi.org/10.3765/bls.v20i1.1466

Slobin, D. I. (1996). From "thought and language" to "thinking for speaking". In J. Gumperz, & S. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge University Press.

Slobin, D. I., Hoiting, N., Kuntze, M., Lindert, R., Weinberg, A., Pyers, J., Anthony, M., Biederman, Y., & Thumann, H. (2003). A cognitive/functional perspective on the acquisition of "classifiers". In K. Emmorey (Ed.), *Perspectives on classifier constructions in signed languages* (pp. 271–296). Lawrence Erlbaum Associates.

Slonimska, A., Özyürek, A., & Capirci, O. (2020). The role of iconicity and simultaneity for efficient communication: The case of Italian Sign Language (LIS). *Cognition*, *200*, 104246. https://doi.org/10.1016/j.cognition.2020.104246

Sümer, B. (2015). *Acquisition of spatial language by signing and speaking children: A comparison of Turkish Sign Language (TID) and Turkish* [Unpublished doctoral dissertation]. Radboud University Nijmegen, The Netherlands.

Supalla, T. (1990). Serial verbs of motion in ASL. In S. Fischer, & P. Siple (Eds.), *Theoretical issues in sign language research* (pp. 127–152). University of Chicago Press.

Supalla, T. R. (1982). *Structure and acquisition of verbs of motion and location in American Sign Language* [Unpublished doctoral dissertation]. UCSD, The USA.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description (Vol: 3 grammatical categories and the lexicon)* (pp. 36–149). Cambridge University Press.

Talmy, L. (2003). The representation of spatial structure in spoken and signed language. In K. Emmorey (Ed.), *Perspectives on classifier constructions in sign languages* (pp. 169–195). Lawrence Erlbaum Associates.

Tang, G. (2003). Verbs of motion and location in Hong Kong Sign Language: Conflation and lexicalization. In K. Emmorey (Ed.), *Perspectives on classifier constructions in sign languages* (pp. 143–165). Lawrence Erlbaum Associates Inc.

Tang, G., Sze, F., & Lam, S. (2007). Acquisition of simultaneous constructions by deaf children of Hong Kong Sign language. In M. Vermeerbergen, L. Leeson, & O. Crasborn (Eds.), *Simultaneity in signed languages* (pp. 283–316). John Benjamins.

Taub, S., & Galvan, D. (2001). Patterns of conceptual encoding in ASL motion descriptions. *Sign Language Studies*, *1*(2), 175–200. https://doi.org/10.1353/sls.2001.0006

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). *Elan: A professional framework for multimodality research*. Proceedings of LREC 2006. Fifth International Conference on language resources and evaluation. http://www.lrec-conf.org/proceedings/lrec2006

Zheng, M., & Goldin-Meadow, S. (2002). Thought before language: How deaf and hearing children express motion events across cultures. *Cognition*, *85*(2), 145–175. https://doi.org/10.1016/S0010-0277(02)00105-1

Zwitserlood, I. (2003). *Classifying hand configurations in Nederlandse Gebarentaal (sign language of the Netherlands)* [Unpublished doctoral dissertation]. University of Utrecht, The Netherlands.

Zwitserlood, I. (2012). Classifiers: Meaning in the hand. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign language: An international handbook* (pp. 158–186). Mouton de Gruyter.