

First-language acquisition in a multimodal language framework: Insights from speech, gesture, and sign

First Language

1–38

© The Author(s) 2024



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/01427237241290678

journals.sagepub.com/home/fla**Dilay Z. Karadöller*** 

Max Planck Institute for Psycholinguistics, The Netherlands; Middle East Technical University, Turkey

Beyza Sümer* 

University of Amsterdam, The Netherlands

Aslı Özyürek

Max Planck Institute for Psycholinguistics, The Netherlands; Donders Institute for Brain Behavior and Cognition, The Netherlands

Abstract

Children across the world acquire their first language(s) naturally, regardless of typology or modality (e.g. sign or spoken). Various attempts have been made to explain the puzzle of language acquisition using several approaches, trying to understand to what extent it can be explained by what children bring to language-learning situations as well as what they learn from the input and the interactive context. However, most of these approaches consider only speech development, thus ignoring the inherently multimodal nature of human language. As a multimodal view of language is becoming more widely adopted for the study of adult language, a multimodal approach to language acquisition is inevitable. Not only do children have the capacity to learn spoken and sign language equally easily, but spoken language acquisition consists of learning to coordinate linguistic expressions in both modalities, that is, in both speech and gesture. To provide a step forward in

*Joint first authorship.

Corresponding author:

Beyza Sümer, Amsterdam Center for Language and Communication, University of Amsterdam, Spuistraat 134, Room Number 6.40, Amsterdam, 1012 VB, The Netherlands.

Email: b.sumer@uva.nl

this direction, this article aims to synthesize findings from research studies that take a multimodal perspective on language acquisition in different sign and spoken languages, including the development of speech and accompanying gestures. Our review shows that while some aspects of language acquisition seem to be modality-independent, others might differ according to the affordances of each modality when used separately as well as together (either in sign, speech, and/or gesture). We argue that these findings need to be integrated into our understanding of language acquisition. We also identify which areas need future research for both spoken and sign language acquisition, taking into account not only multimodal but also cross-linguistic variation.

Keywords

Multimodality, language acquisition, sign language, gesture

There is a growing consensus in the language sciences that human language is inherently multimodal in nature and in its primary face-to-face context (e.g. Hagoort & Özyürek, 2024; Holler & Levinson, 2019; Perniss, 2018). When both the auditory and visual modalities are accessible, languages use both vocal and visual articulators (i.e. head, hands, face, torso). When hearing and speech are not accessible (i.e. in the case of deafness), the visual modality alone can be recruited to subserve all structures and functions of language, as in the case of sign languages. That is, the human language faculty can be expressed in both modalities and even simultaneously.

An interesting consequence of this multimodal nature of language is that no language uses only speech and that all languages used in face-to-face communication involve the visual modality. That is, expressions in the visual modality (e.g. signs, co-speech gestures) are intrinsic to both sign and spoken languages. Co-speech gestures are meaningful and visible communicative movements that universally accompany spoken languages and are aligned with expressions in speech. Deaf communities use systematic and conventionalized movements of the hands, face, and body for all levels of linguistic expression in sign languages (e.g. Goldin-Meadow & Brentari, 2017; Özyürek & Woll, 2019; Perniss et al., 2015). While most theories of adult language have been based on structures expressed mostly in the spoken (and/or written) modality, it is now becoming accepted that a new multimodal language framework is needed to account for not only speech but also sign languages and co-speech gestures used in spoken languages (e.g. Hagoort & Özyürek, 2024; Kita & Emmorey, 2023; Özyürek & Woll, 2019; Vigliocco et al., 2014). For example, there is mounting evidence in the study of adult spoken language that co-speech gestures recruit language-specific brain areas of speech and sign (Özyürek, 2018; Özyürek & Woll, 2019). Yet, such a view of adult language also needs to be integrated into our understanding of language acquisition, which is mostly dominated by investigations of linguistic structures expressed in speech only. In order to contribute toward such a

multimodal perspective on language acquisition in this article, we review studies that have begun to reveal what role multimodality plays in first-language acquisition both for sign languages and for co-speech gestures.

In particular, our review focuses on modality-specific aspects of visual-spatial expressions and how they are acquired as co-speech gestures and in sign languages. Our chosen focus reflects our belief that these modality-specific aspects need to be integrated into our models of language and its acquisition. Expressions in the visual modality, as in co-speech gestures or sign languages, share some unique language features and differ from expressions in speech. These features include visual indexicality (e.g. pointing to a bottle saying ‘this’), visual iconicity (e.g. a drinking gesture to mean bottle or the drinking action itself), and simultaneity in the form of using speech with gestures in spoken languages and the concurrent use of different articulators (hands, face, and body) to represent multiple referents or arguments in sign languages (Slonimska, 2022; Slonimska et al., 2020). These modality-specific features allow different ways to express information (e.g. in indexical, iconic, and simultaneous ways) than those of the sequential and arbitrary nature of most expressions in speech. Thus, how they are recruited during language acquisition allows them to provide different insights into the language-acquisition process.

These modality-specific features of expressions in co-speech gestures and signs might have different consequences for language acquisition than for speech alone. For example, as infants begin to express their first words in speech they can already point to objects and places related to what is expressed in speech (e.g. pointing to a bottle and saying ‘drink’) showing that children can express two arguments instead of one when the multimodal utterance is taken into account. Similarly, a signing baby might point to a bottle by hand and produce a negation sign by shaking their head at the same time. Taking such visual modality-specific expressions into account in language acquisition might portray a different (or similar) account of how children go about acquiring a language (Goldin-Meadow, 2014). For example, some expressions or language structures might appear earlier in the visual modality or in different ways than in speech, such as in the development of vocabulary, syntactic, morphosyntactic, and pragmatic expressions or narrative structure (e.g. Karadöller et al., 2022; Sümer, 2015; Sümer et al., 2014, 2017), as we will discuss further in the current article.

As multimodal language acquisition is already a relatively big field, in this review, we will only focus on the aspects of first-language acquisition where modality-specific aspects of visual expressions might modulate (or not) the language-acquisition process when pointing, iconicity, and simultaneity are considered. For example, acquiring iconic structures in the visual modality with or in place of arbitrary spoken language structures might occur early in the acquisition process. Furthermore, simultaneity might provide children with different ways to express linguistic structures that are different from the sequential structures of spoken language, which needs to be accounted for in language acquisition.

This review focuses on typical first-language acquisition in production, leaving out second-language acquisition, multimodal input, interactive features of the acquisition

process, atypical cases of language acquisition, as well as language comprehension due to space limitations. However, we believe that a unified multimodal view of language acquisition might be foundational for other types and aspects of language acquisition. Below, we first synthesize the literature on language acquisition when both speech and co-speech gestures are taken into account, and then we move on to sign language acquisition, focusing on pointing, iconicity, and simultaneity.

Gestures in spoken language acquisition

Gestures are the actions of the hands, body, and face used as part of communication and mostly accompanying speech (Kendon, 2004). Gesture use is universal, and all children and adults in every speaking community are considered to use gestures as no evidence to the contrary has yet been encountered. Interestingly, the universality of gesture use in communication and in language acquisition is independent of the frequency of gestures in the input (e.g. evidence for the similar trajectory of the emergence of pointing in different cultures, Liszkowski et al., 2012). Furthermore, the ability to gesture does not need gestural input, as children and adults who are congenitally blind also use gestures to accompany speech (Iverson & Goldin-Meadow, 1997; Mamus et al., 2023; Özçalışkan et al., 2016). Gestures have also been shown to play a crucial role in communication and cognition and are part of adults' language structures—albeit they show variation across individuals (Özer & Göksun, 2020) and languages (Özyürek et al., 2008).

Despite being an essential part of cognition, communication, and language, linguistic studies and theories often overlook the role of gestures in spoken language acquisition and focus solely on examining spoken linguistic structures. Initial attempts to understand the role and extent of gestures in language have focused on whether gestures are entirely independent of speech (i.e. the free imagery hypothesis, de Ruiter, 1998, 2000; Krauss et al., 2000) or whether they are closely linked to speech and are limited to expressing information expressed in speech in visual-spatial ways (e.g. Butterworth & Hadar, 1989). More recently, the synthesis of findings from many languages has shown that gestures are not entirely independent of language; and at the same time, they are linked to imagery and are embodied (*embodied cognition*: the notion that cognition is deeply rooted in the body's interaction with the world; Hostetter & Alibali, 2019; Kita & Özyürek, 2003; Özyürek, 2018). Gestures may even be a precursor to language and thought processes (i.e. the gesture-for-conceptualization hypothesis, Kita et al., 2017; also see the article by Avçılar et al., 2022 for a discussion). Therefore, achieving a complete picture of language acquisition is impossible without understanding the role and functions of gestures in language acquisition.

Taking these functions of co-speech gestures into account, in this part of the review, we synthesize the accumulated evidence on the role of gestures in language acquisition, focusing on pointing and iconicity. We review these roles in the context of first-word acquisition, sentence production, vocabulary development, and the expression of


complex conceptual domains such as relational language. Before this, we briefly review the types of different gestures accompanying speech.

Types of gestures accompanying speech

Gestures have many types and communicative functions (Kendon, 2004; McNeill, 1992, 2005). These types include pointing, iconic, emblem, and beat gestures. Although pointing and iconic gestures are typically meaningful only in the context of the speech, emblem gestures have a meaning on their own and can be used in the absence of speech. Finally, beat gestures are the rapid hand movements produced following the rhythm of speech (McNeill, 1992).

Pointing gestures are typically formed by extending the arm, hand, and index finger. They are often used to indicate locations, directions, persons, and objects and are often accompanied by demonstrative and/or personal pronouns (e.g. Peeters & Özyürek, 2016) or by a full noun, or sometimes they are used in isolation (see the articles by Azar et al., 2019; Kita, 2003 for more information). For example, a pointing gesture accompanying the sentence ‘I want the apple over there’, which refers to one apple among many others, conveys the information about which apple is being referred to in the sentence and its location in the immediate environment. Pointing to an empty gesture space might be associated with pronouns or locations in narrative space.

Iconic gestures have semantic content, mostly specified in the particular context of the speech with which they co-occur. They are used to represent an object, person, place, or event visually through the shape or movement of the hand or the relative positioning of the two hands (e.g. McNeill, 1992). Iconic gestures bear visually-motivated similarities between the handshapes they depict and the objects, events, event components, or locations they represent (Furman et al., 2014; Göksun et al., 2010; Karadöller et al., 2022). For example, a stirring motion made with the hand while talking about cooking is visually and structurally similar to the stirring action (see Figure 2 for more iconic gesture examples). Although iconic gestures are visually motivated, the meaning they express is mostly based on the speech they accompany. Indeed, experimental studies have shown that the meaning of these gestures is quite ambiguous in the absence of speech (Krauss et al., 1995).

In emblem gestures, a relatively arbitrary relationship exists between the gesture form and its meaning. Due to this relationship, emblem gestures are similar to words in spoken languages. For example, there is no similarity between the conventional gesture () and the meaning it represents in the Turkish context, i.e. ‘delicious’). Therefore, even when emblem gestures do not accompany speech, the meaning they represent is known by most language users (McNeill, 1992).

Finally, beat gestures (Figure 1) serve important roles in pragmatics, information structure, and phrasal organization. They are used to emphasize parts of speech at the information structure level to express prominence on these parts (McNeill, 1992).



Figure 1. Example of a Beat Gesture.

Source: The example is taken from McNeill (1992).

Although all these gesture types are present in everyday multimodal communication and may play a role in language acquisition, the role of pointing and iconic gestures has been more extensively studied than that of beat and emblem gestures (but see the work of Esteve-Gibert et al., 2022; Rohrer et al., 2022). In the sections to come, we review what is known about the role of pointing and iconic gestures in language acquisition, including how they are simultaneously used with speech, by focusing on the emergence of first words, first sentences, vocabulary development, and the acquisition of specific terminology to communicate about specific domains such as spatial and causal relationships, motion event descriptions, and narrative recall.

Role of pointing gestures in spoken language acquisition

Pointing gestures are essential to children's first-language acquisition (e.g. Kirk et al., 2022; McGillion et al., 2017). Pointing as a communicative act appears around infants' first birthday, regardless of the culture they are born into (Bates, 1979; Bates & Dick, 2002; R  ther & Liszkowski, 2023). Infants use pointing to signal their needs and interests to their caregivers (e.g. Guevara & Rodr  guez, 2023; Tomasello et al., 2007), and points carry an antecedent role in delivering their social motivation for communication (e.g. Matthews et al., 2012). Pointing is used across a range of diverse cultural settings, likely constituting a gestural universal (Liszkowski et al., 2012). Moreover, using

pointing gestures for communication is a crucial milestone for language development, as they are the most reliable indicators that the first words are on their way (e.g. Colonnese et al., 2010). Delays in index-finger pointing have been associated with language delays (e.g. Lüke et al., 2017) and even atypical development (e.g. Sansavini et al., 2019).

Considering the importance of pointing gestures in language development, Rütter and Liszkowski (2023) investigated the ontogenetic origins of pointing gestures by following dyadic interactions between 31 infants and their parents when children were 8–13 months of age through a ‘decorated room paradigm’ (Liszkowski et al., 2012; Liszkowski & Tomasello, 2011). In this paradigm, parents and infants wander around a room decorated with 20 interesting objects (e.g. photos of animals, flowers, etc.) hung on the walls and ceilings. Parents are asked to look at the objects together with their infant while holding them on their hips. Researchers video-recorded these dyadic interactions and later coded for the presence of parent pointing and infants’ earlier emerging showing, hand-pointing, and point-following frequencies. Results showed an increase in all pointing behaviors as a function of age. Moreover, caregiver pointing and infants’ earlier pointing behaviors longitudinally predicted the age of emergence of index-finger pointing. This study advances our understanding of the ontogeny of human pointing by demonstrating the role of early interactional behaviors between parents and children in the emergence of index-finger pointing (see also the study by Liszkowski & Rütter, 2021 for extended discussion).

Once children start using pointing gestures, these gestures predict future language skills. Several studies on infant pointing have demonstrated that children use pointing gestures before they say their first words (e.g. Iverson & Goldin-Meadow, 2005). Before children can say the word ‘ball’, they point to a ball to indicate their need and desire to communicate about it. For instance, Iverson and Goldin-Meadow (2005) followed 10 children longitudinally between the ages of 10 and 24 months by videotaping their 30-minute interactions with their caregivers at home every month. Researchers later coded these interactions for the presence of communicative gestures and speech. Results showed that children’s early pointing gestures have a positive relationship with the early words that they subsequently utter. Pointing use seems to provide alternative ways for children to refer to objects in the visual modality when they do not yet produce words for those objects.

Pointing gestures are the most dominant communication strategy before words appear, but they are not restricted to this period. Children continue to use pointing gestures after they begin to express their first words. Children use speech-pointing gesture combinations before forming two-word utterances or combining words into sentences (Butcher & Goldin-Meadow, 2000; Capirci et al., 1996; Goldin-Meadow & Butcher, 2003; Greenfield & Smith, 1976; Iverson & Goldin-Meadow, 2005; Özçalışkan & Goldin-Meadow, 2005, 2009). For instance, before they start using two words (e.g. ‘my ball’), they point to themselves and say ‘ball’ to express that the ball belongs to them. In the same study, Iverson and Goldin-Meadow (2005) further explored whether children use speech-gesture combinations either in a supplementary or complementary fashion and whether using these multimodal compounds relates to using two-word utterances. Results showed that all children used speech-gesture combinations of either nature for a couple of months before forming two-word utterances. Overall, these studies showed that pointing gesture

use is not only a precursor to the emergence of spoken words but also an indicator of the formation of two-word utterances (e.g. Iverson & Goldin-Meadow, 2005; Özçalışkan & Goldin-Meadow, 2005, 2009).

Pointing gesture use also indicates children's future vocabulary size (e.g. Acredolo & Goodwyn, 1988; Goldin-Meadow & Butcher, 2003; Rowe & Goldin-Meadow, 2009a, 2009b; Rowe et al., 2006; Salo et al., 2018; but see the article by Blake et al., 2003). In most of these studies, gestures that children used were mostly but not exclusively in the form of pointing. Although few in number, some children started to use iconic gestures in addition to points. In one of these studies, Rowe and colleagues (2006) recorded daily home activities of 53 parent-child dyads every 4 months when children were aged between 14 and 34 months. They also administered the Peabody Picture Vocabulary Test (PPVT) at 42 months. Results showed that children's gesture use at 14 months predicts their vocabulary size at 42 months. These results were robust and held even after controlling for children's early vocabulary knowledge and demographics such as SES (Socio-Economic Status) and parental gesture use.

Nevertheless, more recent studies investigating the potential correlates of infant pointing frequency have underlined the importance of SES (e.g. Ger et al., 2023; Rowe & Goldin-Meadow, 2009a) and parental gesture use (Ertaş et al., 2023). For instance, Ger and colleagues (2023) followed 56 Turkish-speaking infants monthly from 8 to 12 months of age to measure the frequency of index-finger pointing along with several candidate correlates, such as SES, mothers' pointing production, and infants' point-following to targets in front of and behind them. Results showed that infants' pointing frequency increases as they age, and their abilities to track and follow their parent's pointing (i.e. point-following abilities) were positively associated with their own pointing frequency. Moreover, high-SES infants had a steeper increase and a higher pointing frequency than the low-SES infants from 10 months onwards, possibly due to higher-SES parents having more responsive and less controlling behaviors (Koşkuşlu et al., 2021; see also the article by Liszkowski & Rütger, 2021 for alternative mechanisms that potentially moderate the relationship between SES and infant pointing such as caregiver-responsiveness). Maternal pointing frequency was not associated with infant pointing frequency at any age. Another study (Ertaş et al., 2023) investigated the same infants at a later time point in development (14 and 18 months of age) and measured infants' and parents' index-finger-pointing frequency as well as their direct (lexical processing efficiency via looking while listening task) and indirect (parent-completed Turkish Communicative Development Inventory; Aksu-Koç et al., 2019) vocabulary proficiency. They found that infants' index-finger-pointing frequency at 14 months predicted their vocabulary development prospectively at 18 months when measured directly via a looking while listening task. Moreover, neither maternal pointing nor infants' own pointing predicted their vocabulary when measured indirectly via parent-completed inventory. These recent studies extend the evidence on the relationship between pointing and language acquisition with more direct measures of vocabulary (Ertaş et al., 2023), shedding more light on the relationship between early pointing and later language abilities (Ger et al., 2023).

The contribution of pointing gestures to language acquisition continues during preschool (see the study by Cameron & Xu, 2011 for the use of pointing in narrative recall)

and school-age periods (see the study by Karadöller et al., 2019, 2022 for the use of pointing gestures to abstract space in describing spatial relationships between objects). During these periods, the relationship between pointing and speech seems to change from pointing being a precursor to speech to pointing being a facilitator. To investigate the role of gestures in narrative recall, Cameron and Xu (2011) investigated 30 children aged between 49 and 58 months in recalling narratives and their use of pointing gestures. Results showed that children who were instructed to point to the locations on the map while retelling the narrative narrated more details about the story than children who were asked to keep their hands still. Cameron and Xu (2011) discussed the possible underlying mechanisms for using gestures to lighten the cognitive load (e.g. Goldin-Meadow et al., 2001; Wagner et al., 2004) and facilitate lexical search (Krauss et al., 1996). Overall, these results underline the scaffolding role of pointing gestures in preschool children's narrative recall.

Finally, pointing gestures to abstract spaces are also used as a visual-spatial compensatory mechanism to missing speech. For example, school-aged children use pointing gestures to provide spatial information that is missing in their speech when expressing information in cognitively challenging domains, such as when describing relative locations for objects that are in a left-right relationship to one another other (Karadöller et al., 2019, 2022). However, pointing gestures are not the only visual expressions used by children; they also employ iconic gestures, which we review in the next section.

Role of iconic gestures in spoken language acquisition

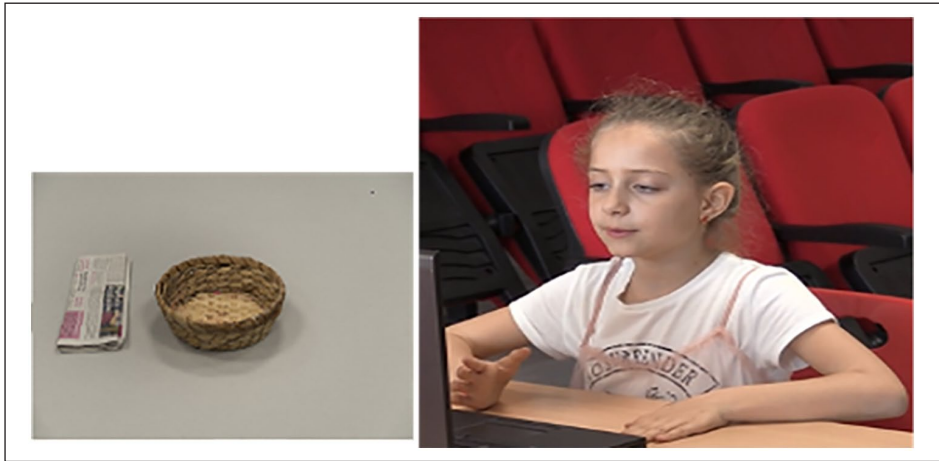
As children develop in their language use and cognition, they communicate beyond indexicality and can gesture about objects and concepts. Sometime between 17 and 36 months of age, they begin to use iconic gestures that depict semantic and visual characteristics of objects, actions, and events (e.g. Capirci et al., 2005; Furman et al., 2014; Özçalışkan et al., 2014). A tight link has been established between early action representations and the appearance of gestures and/or first words (Capirci et al., 2005; see also the article by Volterra et al., 2017). A longitudinal study investigating action, gesture, and word use of Italian-speaking children between 10 and 23 months of age showed that children demonstrate their understanding of meanings by initially practicing actions associated with specific objects (e.g. pushing a little car) and then using iconic gestures and/or words semantically related to these actions (Capirci et al., 2005).

The order between iconic gesture use and the appearance of words related to these gestures seems to be somewhat different from that for pointing. Although pointing gestures consistently precede children's first words, the appearance of iconic gestures does not precede speech but rather develops together with the children's use of verbs that correspond to these gestures (Özçalışkan et al., 2014). For instance, Özçalışkan and colleagues (2014) investigated spontaneous speech and gesture use of 40 American English-speaking children from 14 to 34 months of age in their daily interaction with their mothers. Later, they coded these interactions for the presence of gestures representing actions and action verbs in children's repertoire. They found that children produce their first iconic gestures 6 months later from their first verbs related to these actions.

Moreover, iconic gestures increased in frequency together with the use of verbs, and over time, these gestures began to convey meanings that were not yet expressed in speech. Therefore, rather than preceding verbs, gestures possibly allow children to represent actions iconically while they produce these verbs. Iconic gestures might also compensate for the lack of specific semantic knowledge in verbs representing these actions only after children acquire a verb system.

Acquiring a verb system is not the only challenge children face in their language-acquisition journey. They need to learn language terms to talk about conceptually-challenging domains such as expressing causal relationships (e.g. Göksun et al., 2010) or spatial relations between objects (e.g. Karadöller et al., 2022). Iconic gesture use precedes the acquisition of terms to communicate about complex conceptual domains, such as spatial relations between objects, events, and/or event components, causal relationships, and so on. Several studies have demonstrated that children capitalize on using iconic gestures to compensate for the lexical terminology they lack when communicating about these domains. As a result, they use gestures more frequently than adults (e.g. Austin & Sweller, 2014; Church et al., 2000; Karadöller et al., 2022; Sauter et al., 2012). The main factor underlying children's high frequency of gesture use is that gestures support speech by creating an alternative communication channel (Krauss et al., 2000; Melinger & Levelt, 2004). This is especially the case when the development of language structures for complex conceptual themes lags behind cognitive development for these concepts (Church & Goldin-Meadow, 1986; Perry et al., 1992). In these instances, gestures convey conceptual information that language structures cannot yet express. This is often seen in the expression of conceptually-challenging categories such as spatial relationships or causal events (Austin & Sweller, 2014; Calero et al., 2019; Furman et al., 2014; Göksun et al., 2010; Karadöller et al., 2019, 2022; Karadöller, Sümer, & Özyürek, 2021; Karadöller, Sümer, Ünal, & Özyürek, 2021). For instance, children use gestures to convey information that is missing from their speech when they cannot express spatial terms such as 'Left-Right' (e.g. Karadöller et al., 2019, 2022; Karadöller, Sümer, & Özyürek, 2021; Karadöller, Sümer, Ünal, & Özyürek, 2021) or the grammar structures used to talk about causal events (e.g. Göksun et al., 2010).

Even though children start to communicate about spatial relations between objects using spatial terms very early in life, such as 'In-On' appearing at age 2, followed by 'Front-Behind' at the age of 5 years (e.g. Johnston & Slobin, 1979), using spatial terms to refer to left-right relations between objects does not usually appear until children are 8 years or older (Clark, 1973). Children do not often use 'Left-Right' in the early stages of their language development; instead, they use somewhat ambiguous spatial terms, such as 'Side' or 'Next to' (Figure 2). In these instances, they resort to gestures to indicate the exact relationship they want to convey, which is missing in their speech (Karadöller et al., 2019, 2022; Karadöller, Sümer, & Özyürek, 2021; Sümer, 2015). These findings reveal that gestures support their communication about spatial relationships that are known at a conceptual level but not yet expressed in speech. Moreover, these findings also highlight the importance of multimodality's crucial role and function in language acquisition.



Right Hand: Iconic/Placement (basket)
Left Hand: Iconic/Placement (newspaper)

Sepet-in	yan-in-da	gazete	var
basket-GEN	side-POSS-LOC	newspaper	there_is

‘There is a newspaper at the side of the basket’.

Figure 2. Spatial Description of an 8-Year-Old Turkish-Speaking Child Using a General Spatial Term (Side) in Speech and Iconic Hand Placement Gestures that Depict More Specifically the Relative Locations of Objects to Each Other.

Notes. The participant introduced gestures sequentially. The gesture indicating the basket with the right hand was performed when the participant mentioned the basket in her speech. The gesture indicating the newspaper with the left hand was performed when the participant mentioned the newspaper. Both gestures remained in the gesture space until the end of the description. The spatial description conveys the exact spatial information between the objects only when combined information in speech and gestures is considered.

Similar to what has been found for gestures aiding the communication of spatial relations between objects, studies show that children use gestures to compensate for missing information when describing causal relationships related to events. For instance, Göksun and colleagues (2010) asked 64 preschoolers aged between 2.5 and 5 years to describe an action in which the experimenter pushed a ball across a small pool with a stick. Results show that younger children produced noncausal sentences and location gestures to refer to the goal of the action. However, older children also used gestures that conveyed missing information in speech about the instrument and the direction of the action. These results underscore not only a close relationship between action representations (e.g. Capirci et al., 2005) and children’s manifestation of these representations via the use of gestures but also the role of gestures in carrying causal information before children can form complete sentences to describe causal relationships related to these events.

In addition to gestures aiding the communication of conceptually-challenging domains, narrative development stands as another domain that demonstrates the role of iconic gestures in language acquisition, especially when referring to character viewpoints. Although previous work on character introduction in narratives suggests that

children start to introduce characters around 6 years of age (e.g. Bower & Morrow, 1990), recent studies investigating multimodal referent introduction in narratives found that children can already understand character viewpoint and can introduce characters via gestures before doing so in speech (e.g. Demir et al., 2015; Stites & Özçalışkan, 2017). For instance, Stites and Özçalışkan (2017) asked 36 children between 4 and 6 years of age to watch and tell a cartoon that depicted Sylvester climbing up a drainpipe to catch Tweety, then swallowing a bowling ball thrown by Tweety and rolling into a bowling alley with the ball in his stomach. Results showed that 5-year-olds can introduce the character with a pronoun along with iconic gestures that show the character from the character's point of view (e.g. raising the body and bringing the hands to the side of the face like cat claws when saying, 'He is climbing up the pipe'). These results provide multimodal evidence that children acquire character viewpoints in narratives and can introduce characters together with their perspectives via speech-gesture combinations before expressing them exclusively in their speech.

Overall, gestures accompanying speech are not only an alternative communication channel that supports the acquisition of linguistic structures (e.g. Alibali et al., 2009; Church et al., 2000; Colletta et al., 2010; Karadöller et al., 2022), but they are also a communication strategy that is frequently used by adults (McNeill, 1992). As children progress in their cognitive skills and can employ linguistic terminology to refer to their thought processes, the role and functions of gestures may shift from being complementary (i.e. conveying information missing in speech) to being co-expressive (i.e. conveying information that is already present in speech) (e.g. Karadöller et al., 2019, 2022). The frequency and the role of gestures (complementary or co-expressive) in adulthood may vary across individuals (see the article by Özer & Gökşun, 2020 for a review).

Finally, it has been shown that iconic gestures show variability in different languages, varying with different typological structures (see the article by Ünal et al., 2023 for a review). Several studies demonstrated a language-specific patterning of gesture use early in development as a function of age, especially when conveying information related to event components that show cross-linguistic variability (e.g. Furman et al., 2006, 2010, 2014; Gullberg et al., 2008; Özyürek et al., 2008). Semantic elements of motion events (agent, action, figure, path, manner, etc.) are mapped onto lexical and syntactic structures differently across languages (Talmy, 2000). These differences lead to variations in the constructions preferred by the speakers of different languages (Kita, 2009; Ünal et al., 2023). Iconic gestures have also been found to be sensitive to cross-linguistic differences. For instance, in caused-motion events (e.g. she put the toy away), the main semantic elements of caused-motion (action and path) can be encoded in the verb (e.g. *kaldır*- put.away) in *verb-framed* languages such as Turkish. Elements of the same event can also be represented by encoding the action in the verb (e.g. put) and path in the satellite (e.g. away) in *satellite-framed* languages such as English. Previous research has shown Turkish-speaking children's speech and iconic gestures to manifest this linguistic specificity (Furman et al., 2014). Specifically, Turkish-speaking children encoded the semantic elements of caused-motion in the verb in their speech. Moreover, their gesture also reinforced or supplemented their speech by encoding these semantic elements in language-specific ways, although these types of gestures were scarce in English-speaking children's repertoire. The authors argue that the language children learn has the potential

to shape the amount and the type of gestures they use in their first years of life from 14 and 19 months onward, respectively (Furman et al., 2014). These results suggest that the emergence of iconic gestures (i.e. involving actions) might be facilitated in languages where verbs are more central and allow for verbs and iconic gestures to emerge together early in development.

Cross-linguistic differences in motion event descriptions (e.g. Boy run down the stairs) are also present in how manner and path are encoded. For instance, Japanese and Turkish encode the manner and the path of a motion event in separate verbs and clauses (e.g. *koşarak indi*, run descent) in speech, and speakers of these languages also use separate gestures for them. Whereas in other languages, such as English, these elements are encoded in a single clause (using a verb and a satellite) (e.g. run down) in speech, and speakers of English also use a single gesture to represent these elements (e.g. Kita & Özyürek, 2003). Özyürek and colleagues (2008) investigated 3-, 5-, and 9-year-old children and adults from two typologically different languages (English and Turkish) in encoding event components via speech and their iconic gesture use. Participants watched and retold 10 videoclips of motion events (e.g. Tomato Man slides up a hill after being pushed by Green Man). Although children use adult-like linguistic structures in speech to encode these elements from 3 years onwards, their adult-like gestural representations take longer to emerge (about 9 years of age). That is, both groups of children start in similar ways (using separated gestures for manner and path), but English-speaking children, over time, start using conflated gestures. These findings demonstrate that language-specific encoding of events develops earlier in speech than in gesture. This development requires considerable time until gestures are shaped by language over time, especially for languages where more than one event component is expressed in one clause.

Interim summary: gestures in spoken language acquisition

Gestures are integral to language (Kendon, 2004) and are used by speakers across all communities (Liszkowski et al., 2012), including blind speakers who have never seen gestures (Mamus et al., 2023). In this part of the article, we summarized the role and functions of gestures in first-language acquisition by synthesizing the evidence focusing on pointing and iconic gestures. This synthesis demonstrated several pieces of evidence that both pointing and iconic gestures have an important role in providing a visual-spatial scaffold for the emergence of words and language structures. It also showed that since gestures are closely linked to spoken languages, they are embedded into and develop in tandem with spoken components of language as children age.

Many studies reviewed here show clear evidence for gestures preceding spoken language acquisition. First, pointing gestures have been found to be a precursor to and predictor of several domains of language acquisition, such as the appearance of first words, first two-word utterances (e.g. Iverson & Goldin-Meadow, 2005), and vocabulary size (Rowe et al., 2006). Moreover, iconic gestures have been found to be a precursor to the acquisition of terms and linguistic structures concerning conceptually-challenging relational domains such as spatial (e.g. Karadöller et al., 2022) and causal (e.g. Göksun et al., 2010) relationships, and narrative recall (e.g. Stites & Özçalışkan, 2017).

Gestures, however, cannot always be considered a precursor to speech as, in some cases, they seem to develop in tandem with speech and are sensitive to language specificity. Empirical evidence for gestures to follow speech development centers mostly around iconic gestures. Specifically, iconic gestures for actions have been found to follow the development of the verbs or emerge together, allowing children to express the semantic specificity of the actions (e.g. Özçalışkan et al., 2014). However, the emergence of iconic gestures might be facilitated in languages where verbs are more central and where verbs and iconic gestures might emerge together early in development (Furman et al., 2014). Moreover, iconic gestures encoding motion event components are closely tied to the cross-linguistic variation in encoding these components (e.g. Özyürek et al., 2008) and might look language-specific after language specificity emerges in speech (but see the article by Özçalışkan et al., 2023 for showing language specificity for cases when speech and gesture might develop together).

Therefore, the development of pointing and iconic gestures goes hand in hand with language development. When gestures are considered together with what is expressed in speech, it becomes clear that children have more expressive capabilities than can be expressed in their speech. Thus, what children know and can express as they acquire language might be very limiting when one focuses only on speech. Gestures also reveal language-specific developmental trajectories, such as the early action and verb bias that we see in the development of Turkish causal events. This part of the review therefore calls for revisiting current language and cognitive development theories across different languages with a multimodal viewpoint.

Sign language acquisition

The studies discussed earlier in this article provide robust evidence for the intricate relationship between gesture use and first-language acquisition at different stages and for different domains of language. Speaking children employ all mediums available to them to communicate their message and interact with their interlocutors. For deaf children, however, the visual modality is the primary modality accessible to them, and they mainly use their hands and body (including eye gaze and facial expressions) to construct meaning and produce language structures. Thus, the linguistic organization of sign languages hinges on the affordances of the visual modality in a systematic way. In this part of our review, we shift our focus to sign languages and review the evidence on whether and how the visual affordances of sign languages (especially pointing, iconicity, and simultaneity) modulate sign language acquisition for different domains such as vocabulary and morpho-syntax as used in encoding spatial relations and narrating events.

In many respects, the trajectory of sign language acquisition is similar to that of speech. For example, similar to the appearance of first words, first signs appear at around 12 months of age; a vocabulary spurt typically occurs roughly at 18 months; two sign combinations appear at 24 months; the 500-sign stage is reached by 36 months; grammar emerges between 2 and 3 years of age; and the acquisition of discourse functions continues up to school age (Chamberlain & Mayberry, 2000; Lillo-Martin & Henner, 2021; Morgan & Woll, 2002). This parallelism between sign and speech development is observed if children interact with fluent language users, who can provide them with rich

linguistic input in sign language from early on. However, this is only the case for a minority of deaf children (around 10%) since most of them are born into families where there are no proficient signers (Mitchell & Karchmer, 2004). To make this part of our review comparable to speaking children's multimodal language development as described above, we will focus on sign language acquisition in deaf children who receive early sign language input. Despite the universal patterns of language development shown for early signing children and speaking children, there has also been accumulating evidence showing how modality-specific aspects of sign languages, such as pointing, iconicity, and simultaneity, modulate different domains of language acquisition.

Pointing in sign language acquisition

The use of pointing is one of the first ways to communicate about the world and is key to achieving joint-attention behavior, a necessary component of language acquisition (e.g. Colonnese et al., 2010; Goldin-Meadow, 2007; Tomasello et al., 2007). All children, deaf or hearing, point at objects, people, places, or sometimes an event to get the attention of their interlocutors. Indeed, several studies have reported that signing children use pointing more frequently than their non-signing peers (Cormier et al., 1998 for American Sign Language (ASL); Hoiting, 2009; Hoiting & Slobin, 2007 for Sign Language of the Netherlands (NGT), Kanto et al., 2015, 2024 for Finnish Sign Language (FinSL), Morgenstern et al., 2010 for French Sign Language (LSF); Wille et al., 2018 for Flemish Sign Language (VGT)). For example, the study by Morgenstern and her colleagues (2010) presented evidence for the earlier and more frequent use of pointing by making direct comparisons between a deaf child acquiring LSF, a hearing bimodal bilingual child acquiring French and LSF, and a hearing child acquiring French across 7–24 months of age. Overall, it was the deaf child who produced more pointings than her hearing peers. Furthermore, the children who were acquiring LSF (both the deaf child and hearing child) started to produce them 3 months earlier than the French-acquiring hearing child, whose first pointings emerged around 11 months.

However, deaf children point in different ways than speaking children, and in the case of sign languages, pointing progressively gains grammatical functions such as personal reference (i.e. pronouns) (Cooperrider et al., 2021; Cormier et al., 2013), unlike what we see for pointing in speaking children. For example, deaf children usually point toward their chest to mean 'I' (first-person singular) or to their addressee for 'you' (second-person singular). One prediction could be that such transparency (e.g. the direct link between pointing and the reference 'I') might facilitate the acquisition of pronominal pointing behavior in deaf children. Despite the earlier and more frequent use of pointing in sign compared to speech development, several studies with deaf children acquiring sign language reported a discontinuity in using pointing to refer to people (i.e. as a pronoun): deaf children freely pointed at people and objects up to 12–15 months of age. However, they later stopped pointing at people (including themselves) while continuing to point at objects and places till 18 months of age. When these children started to point at people again at 20 months of age, it was mainly self-pointing, and pointing at other people (including their addressees) emerged at 24–25 months of age. The reappearance of pointing to people has been interpreted as showing that pointing to people might have a

special status in a child's linguistic system, that is, pronominal function (Caët et al., 2017; Hatzopoulou, 2011 for Greek Sign Language (GSL); Petitto, 1987 for ASL; Pizzuto, 1990 for Italian Sign Language (LIS)). Thus, signing children need to differentiate functions of pointing (e.g. deictic, pronominal), which might be a reason for the observed discontinuity in the use of this type of pointing in sign language acquisition.

Another interesting observation about pointing to people by deaf signing children relates to their predictive role for sign language development in other domains of language. More specifically, the disappearance of pointing (especially self-pointing) is related to acquiring other linguistic aspects. When signing children stopped pointing at people, they started using more sign tokens (i.e. number of signs) as well as more sign types (i.e. sign categories such as predicates [e.g. WANT], lexical signs, classifier constructions, etc.; Caët et al., 2017; Petitto, 1987). Furthermore, when they re-started using self-points with pronominal functions, they could produce more complex syntax, combining an average of three signs or more (Blondel & Tuller, 2008; Caët et al., 2017; Petitto, 1987). For example, at about 2 years of age, an LSF-acquiring deaf child was able to produce combinations such as SELF-POINT EAT POINT, meaning 'I want to eat that [cheese]' (Morgenstern et al., 2010). Self-points, therefore, seem to be progressively integrated into deaf children's syntactically complex utterances.

Most of the research on pointing in signing children focuses on pointing to refer to people. However, the few studies that explored pointing from an interactional perspective in parent-child contexts report its use to refer to objects. For example, Wille et al. (2018) report the use of pointing and lexical signs by a deaf girl acquiring VGT at around 18 months of age. Furthermore, Kanto et al. (2024), comparing data from KODAs (hearing Kids of Deaf Adults) and hearing children with hearing parents, observed differences in the use of pointing and language development, that is, the more KODAs produced signs in FinSL at the age of 1, the more they produced pointing at the same age. On the other hand, the more hearing children produced words in Finnish at the age of 3 years, the less they produced pointing at the same age. Furthermore, the frequency of pointing of the KODAs was rather stable across the different data points, even though their language skills developed over time, whereas the pointing frequency of hearing children decreased over time while their language skills developed further.

The studies reviewed here depict an intriguing developmental trajectory for pointing in sign languages, as in sign languages, pointing is used for several functions, unlike in spoken languages. The affordances of the visual modality of sign languages for the representation of referents for pointing to external objects as well as for pronominal functions seem to encourage the earlier and more frequent use of pointing by signing children compared to their speaking peers.

Iconicity and simultaneity in sign language acquisition

In sign languages, many (but not all) signs are visually motivated since they depict their referents, actions, or events in iconic ways (Karadöller et al., 2022; Karadöller, Sümer, & Özyürek, 2021; Perniss et al., 2010, 2015; Taub, 2001). Sign languages also allow for the simultaneous use of multiple articulators in order to depict more complex relations (e.g. spatial relations between objects) or actions (Slonimska et al., 2022).

These modality-specific features of sign languages potentially modulate the acquisition of vocabulary and linguistic structures.

Iconicity in sign language forms is mainly defined as the visually-motivated mapping between a language form and its referent (Perniss et al., 2010; Slonimska & Özyürek, in press; Taub, 2001). This link is modulated by relevant cognitive abilities (e.g. being able to make analogies between language forms and their referents), conceptual knowledge, and world experience (Emmorey, 2014; Pizzuto & Volterra, 2000; Sevcikova Sehyr & Emmorey, 2019). In this part of the review, we first illustrate iconicity and simultaneity in the form of lexical signs and linguistic structures. Next, we review the literature on whether and how they modulate sign language acquisition.

Iconic signs bear visually-motivated resemblance(s) to what they refer to in the real world. However, how these signs resemble their referents shows variation across sign languages because iconicity in a lexical sign can be motivated by different properties of a referent. We see different correspondences between visually perceptible features of the signs and their referents (i.e. imagistic iconicity, Russo, 2004; Taub, 2001). For example, the lexical sign for DOG in Turkish Sign Language (TİD) highlights the barking of a dog (Figure 3[a]), while in British Sign Language (BSL), it depicts an image of a dog standing on its feet (Figure 3[b]) or, in ASL, how to call a dog (Figure 3[c]).

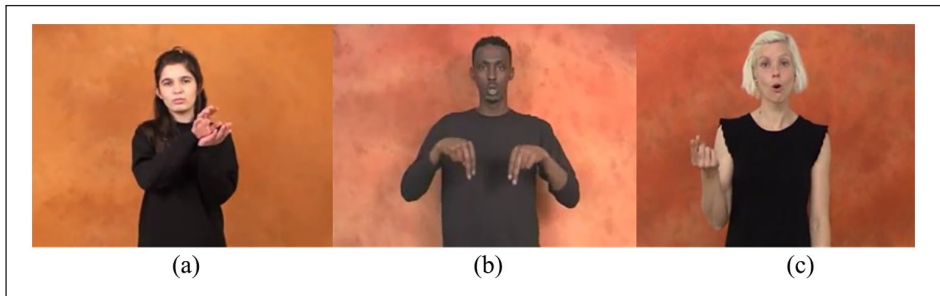


Figure 3. Lexical signs for DOG in (a) Turkish Sign Language (TİD), (b) British Sign Language (BSL), and (c) American Sign Language (ASL).

Source: Pictures are taken from www.spreadthesign.com

Iconicity in sign languages is not only restricted to lexical signs but also observed in linguistic structures that depict the location or motion of referents in space. One of the language forms with this type of iconicity is classifier constructions (see the article by Schembri, 2003 for the use of terminology). In these forms, the handshapes and their location/movement in relation to each other in signing space are associated with the size, shape, and location/motion of the referents being encoded (Brennan, 1990; Emmorey, 2002; Perniss, 2007). For instance, a signer of LIS in Figure 4 depicts a pen located to the right of a book (from his own perspective) using a classifier construction. While his left hand represents the flat surface of the book, his right hand shows the elongated shape of the pen, both of which reflect the perceptual features of these referents. The signer locates both of his hands in the signing space to depict the spatial

relations between them simultaneously (Slonimska, 2022). Along with spatial relations between objects, these forms can be used to depict different components of an event, such as agents involved and/or their movement by employing multiple body articulators simultaneously to refer both to their perceptual features and their relations with respect to each other (i.e. diagrammatic iconicity, Cuxac, 1999; Perniss, 2007; Slonimska, 2022).



Figure 4. Description of a spatial scene in which there is a pencil next to a book by a signer in Italian Sign Language (LIS).

Source: The example is taken from Slonimska (2022).

The third picture shows the choice of different classifier handshapes used for book and pencil and the spatial relations between them.

Despite a considerable body of research showing similar language-acquisition trajectories for both signing and speaking children under comparable input conditions (e.g. Lillo-Martin & Henner, 2021), it has also been questioned whether the prevalence of iconic forms in sign languages modulates language development for signing children differently compared to their speaking peers. Research investigating how visual affordances of sign languages, such as iconicity and simultaneity, modulate their acquisition has mainly focused on three areas: vocabulary use, expressing location and motion, and narrating events.

In early studies of sign language acquisition, the potential effects of iconicity on early vocabulary development were defined as being either weak or none (e.g. Orlansky & Bonvillian, 1984). Nevertheless, even these early studies suggested that signing children produce their first signs as well as their first 10 signs 2–4 months earlier than speaking children (e.g. Bonvillian et al., 1990). These findings were initially framed as a biological advantage for signing over speaking children since the hand muscles involved in producing signs mature earlier than the muscles in the vocal tract (Meier & Newport, 1990). However, despite this early maturation of muscles of main articulators, sign language acquiring children still make several phonological errors in the signs that they produce, with handshape in particular being notoriously challenging to acquire (Conlin et al., 2000; Gu et al., 2022; Karnopp, 2008). Furthermore, recent studies have challenged the ‘biological advantage’ view by emphasizing the facilitating role of iconicity

on vocabulary development (e.g. Caselli & Pyers, 2017; Novogrodsky & Meir, 2020; Sümer et al., 2017; Thompson et al., 2012) and the development of linguistic structures to communicate in other domains such as spatial relations (e.g. Karadöller et al., 2022; Perniss & Vigliocco, 2014; Sümer, 2015; Sümer et al., 2014, 2017).

The current line of sign language acquisition research has yielded strong insights into the predictive role of iconicity for early vocabulary development. Studies that examined parental report data (MacArthur-Bates Communicative Development Inventory [CDI]) on the production of lexical signs have revealed that signs with greater iconicity are acquired earlier than less iconic ones (Caselli & Pyers, 2017 for ASL; Novogrodsky & Meir, 2020 for Israeli Sign Language [ISL]; Sümer et al., 2017 for TİD; Thompson et al., 2012 for BSL). Most of these studies, however, report a clear link between age and the effect of iconicity on lexical sign development. Namely, older children benefit from iconicity more than younger ones during vocabulary learning (Novogrodsky & Meir, 2020; Sümer et al., 2017; Thompson et al., 2012). However, the facilitating effect of iconicity for early vocabulary development is still observed even at the age of 10–12 months. Greater benefits from iconicity in older children could be due to greater cognitive understanding or enhanced environmental and cultural experiences that help them understand and establish motivated links between meaning and form.

It is possible that, despite being iconic, lexical signs can also be complex, involving simultaneity and several movements (Figure 5[a]). Sometimes lexical signs might also have a non-manual component, as shown in Figure 5(b), in which a signer of Danish Sign Language (DSL) is producing the lexical sign for NABO (neighbor) while also mouthing the Danish word for it. Interestingly, simultaneity involving combinations of movements does not seem to pose a challenge for deaf children's sign acquisition (Gu et al., 2022; Tomaszewski et al., 2022). In contrast, signing children are reported to use mouthing/mouth actions together with manual articulators very infrequently (Wille et al., 2018).

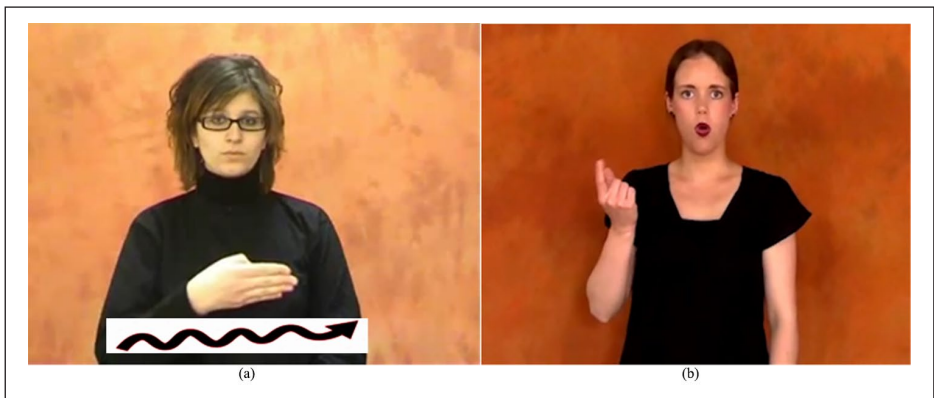


Figure 5. Lexical sign for (a) FISH in French Sign Language (LSF), which consists of two movements (path movement in signing space from right to left and small repeated movements of the wrist) and (b) NEIGHBOR in Danish Sign Language (DSL), in which the signer mouths the Danish word NABO while articulating the sign.

Source: Taken from www.spreadthesign.com.

There might be different sources for the iconicity in a language form (e.g. Ortega et al., 2017; Padden et al., 2014; Thompson, 2011). Lexical signs for tools, for example, may employ a certain set of handshapes (i.e. handling), while the ones for bulky objects such as houses will employ another set (i.e. object). Thus, one group will be more action-based since they depict how we handle/hold objects, while the other group will be based on the perceptual features of their referents (e.g. whether the referent is round or flat). It is also possible that the same objects can be expressed by two different lexical signs that differ in the type of iconicity (Figure 6). Ortega and his colleagues (2017) were interested in whether and in which ways having two variants that differ in how iconicity is represented (action-based vs perceptual-based) modulates the acquisition of lexical signs. Here, they showed pictures of objects to 10 TID signing adults and two groups of deaf children (4–6 years and 7–9 years) and asked them to produce lexical signs for these objects. The analysis of lexical sign productions revealed differences in the preferences of the use of different variants across groups. More specifically, adults preferred perceptual-based variants (Figure 6[b]) more frequently than action-based (Figure 6[a]) ones, while both groups of children mainly used action-based variants.



Figure 6. Action-based (a) and perceptual-based (b) variants for TOOTHBRUSH in TID.

One might argue that action-based signs are preferred by signing children because they are more iconic than perceptual-based ones. Thus, the higher degree of iconicity in action-based signs, rather than the motoric/embodied representations in them, drives early lexical sign acquisition. If this is the case, iconicity would explain such preferences of signing children better than an embodied view of language acquisition. Caselli and Pyers (2019) addressed the same issue by focusing on the interaction between degree and type of iconicity in lexical signs by analyzing parental reports from 58 ASL signing deaf parents with deaf children. They did not find an effect of iconicity type but rather observed a strong effect of the degree of iconicity on modulating early lexical sign acquisition. However, the average age of children in their study was 24 months, thus including

a much younger group of children compared to the ones in the study by Ortega et al. (2017). It is likely that these ASL-acquiring children have not yet had enough motoric/bodily experiences, cognitive capacity, or world knowledge to fully build the links between the form of the signs and their meanings. Thus, at younger ages, signing children may benefit more from the degree of iconicity regardless of its type (i.e. action or perceptual). When they get older, they are better able to distinguish types of iconicity and benefit more from those related to their motoric/bodily experiences. This explanation aligns with other studies showing that iconicity seems more advantageous for lexical sign acquisition when children are older than 24 months of age (Sümer et al., 2017; Thompson et al., 2012 see also the article by Ortega et al., 2017 for a review).

So far, the above-mentioned studies have mainly focused on the acquisition of lexical signs and how they are modulated by iconic features in these signs. However, iconicity is not only restricted to the form of lexical signs but is also seen in classifier constructions, which are mainly employed to encode spatial relations between objects and motion event descriptions. In these descriptions, signers across several sign languages mainly employ classifier constructions (Emmorey, 2003; Perniss, 2007), as depicted in Figure 4. Despite their iconic properties, the production of these language forms can be challenging for signing children since they need to select appropriate classifier handshapes (e.g. a flat—but not a round—handshape to refer to *BOOK* in Figure 4) as well as simultaneously coordinate manual articulators in signing space to refer to each referent. Thus, this domain can be quite relevant to understanding the effects of iconicity in sign language development due to its visual properties and structural complexities, thereby allowing us to explore whether and how these factors (iconicity and simultaneity) play a role in the acquisition of these forms by signing children.

Despite the universal tendency reported for the emergence of spatial terms in speech (i.e. spatial relations without a viewpoint such as ‘In’ and ‘On’ are produced earlier than the ones with a viewpoint such as ‘Left’ and ‘Right’), recent studies with signing children present quite a different picture and reveal a strong role of language modality in the order of their acquisition. Studying spatial descriptions (‘In’, ‘On’, ‘Under’) elicited from 20 deaf children (between 4 and 9 years of age) acquiring TİD and comparing them to those elicited from hearing children acquiring Turkish, Sümer (2015) and Sümer et al. (2014, 2020) found that both signing and speaking children were on par when they acquired adult target forms. However, the pattern was surprisingly different for expressing spatial terms that require taking a viewpoint (i.e. ‘Left’, ‘Right’). While encoding the spatial relation between the objects located on a lateral axis (e.g. pen left to paper), even the signing children at the age of 4 years were adult-like in how frequently they used locative forms such as classifier constructions (Figure 7). This was not the case for similarly-aged speaking children, who were still struggling to provide adult-like target forms for similar descriptions in their speech, even when their co-speech gestures were considered (Karadöller et al., 2022). Furthermore, TİD acquiring children, even at the age of 3–4 years, were able to produce simultaneous classifier constructions while describing object locations, in contrast to the difficulties that were suggested by earlier studies on the acquisition of these constructions (e.g. Engberg-Pedersen, 2003; Newport & Meier, 1985; Newport & Supalla, 1980; Slobin et al., 2003; Supalla, 1982; Tang et al., 2007). However, it is important to note that these earlier

studies focused on the expression of motion events, which consist of more event components compared to encoding location. In the next paragraph, we shift our attention to the domain of motion events.



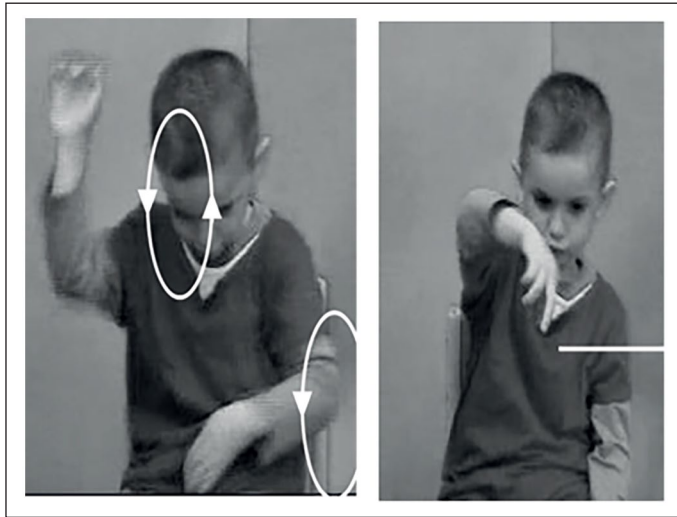
LH:
RH: PAPER

PEN
CL(paper)

CL(pen)
----- HOLD -----

Figure 7. Description of a spatial scene in which there is a pen next to a paper by a 5-year-old TID signer.

The use of classifier constructions while expressing motion events (e.g. a man walking toward a car) has been found to be challenging for signing children, who are not usually able to use them in adult-like ways until around 12 years of age (e.g. Engberg-Pedersen, 2003; de Beuzeville, 2006; Kantor, 1980; Schick, 1990; Slobin et al., 2003; Supalla, 1982; Tang et al., 2007). These challenges have been mainly reported for the use of appropriate classifier handshapes (Kantor, 1980) and the simultaneous representation of different event components, such as manner (e.g. walk) and path (e.g. toward) (Newport, 1981; Newport & Supalla, 1980; Supalla, 1982). Signing children tend to omit certain components at the early stages of language acquisition, but they express them simultaneously soon after (Newport, 1981; Newport & Supalla, 1980; Sallandre et al., 2018; Supalla, 1982). For example, 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 hearing children acquiring French, who frequently produced path-only descriptions and fewer path + manner utterances in their motion event encodings with various paths and manners. However, when these children encoded these types of events, they had a tendency to express path and manner components separately, as shown in Figure 8. Here, a 5-year-old deaf child is describing a ‘swim across’ event by expressing manner (swim) and path (across) in two different forms. These findings suggest that signing children do not acquire these constructions in an analog or holistic fashion, despite the visual resemblance between their linguistic forms and the real motion event (Newport, 1981; Newport & Meier, 1985; Newport & Supalla, 1980; Supalla, 1982).



LH:	SWIM	
RH:	SWIM	ACROSS

Figure 8. Description of a motion event by a 5-year-old signer of LSF, in which he encodes manner and path components in different language forms.

Simultaneity is not only confined to encoding path and/or manner in classifier constructions but also involves the expression of ground (bigger and usually stationary objects) and figure (smaller objects whose location and/or motion is encoded with respect to grounds) in expressing motion events (e.g. Aronoff et al., 2003; Perniss, 2007; Sümer, 2015; Supalla, 1982). In recent work, Sümer and Özyürek (2022) observed that TİD acquiring children (aged 4–10 years) omitted one of these components of a motion event more frequently than signing adults (Figure 9). Reporting similar findings from similar-aged Turkish acquiring children, the authors suggest this tendency to be a pervasive feature of children’s motion event descriptions (Engberg-Pedersen, 2003; Hickmann, 2005; Morgan et al., 2008; Slobin et al., 2003; Supalla, 1982; Tang et al., 2007) and might be derived from general cognitive tendencies in children regardless of the language modality that they acquire.



LH:

RH: MAN

CL(man_walking)

Figure 9. Description of a motion event (man walking toward a car) by a 4-year-old signer of TID, in which he drops the mention of the car, which is the ground in this event description.

Finally, another domain where the effects of iconicity and simultaneity have been studied is narrative development. In narrating events, several factors, such as providing background information (e.g. Berman, 2001; Berman & Slobin, 1994) and marking the identifiability of the referents that are introduced and maintained during the discourse, are taken into account (e.g. Küntay, 2002). Studies with children acquiring speech revealed that the younger the children are, the less information they provide to set the scene in their narratives (Berman, 2001; Peterson, 1990; Peterson & McCabe, 1983; Umiker-Sebeok, 1979). However, these studies are conducted with children who are required to translate events presented usually in visual-spatial form (as in the case of picture-story narrations) into sequential segments of verbal output, thus causing a particular kind of cognitive demand (Berman & Slobin, 1994). Thus, it might be interesting to ask whether signing children encode more information in their narrations than their speaking peers due to the similarity between the referent and the linguistic form. Previous studies that compared these two groups of children proposed similar developmental trajectories: using linguistic forms for reference in an unclear manner up to the age of 3 years and not being able to maintain the characters introduced at the beginning consistently throughout the narration (Anthony, 2002; Hickmann et al., 1996; Karmiloff-Smith, 1981; Morgan, 2000, 2002; Morgan & Woll, 2003; Niederberger, 2004; Sümer, 2016; Vercaingne-Menard et al., 2001). Furthermore, signing children (around 3–4 years of age), similar to their speaking peers, usually start with direct quotations (i.e. using one's words verbatim); only later do they use role-shift, which is a sign language-specific strategy in which they

take the perspective of the person whose thoughts, utterances, or actions are being reported (Ely & McCabe, 1993; Lillo-Martin, 2012; Lillo-Martin & Quadros, 2011).

Interim summary: sign language acquisition

The affordances of the visual modality, such as indexicality, iconicity, and simultaneity, seem to influence sign language development in modality-specific ways at different stages of acquisition and for different language domains. There are also areas of language acquisition where these modality-specific aspects do not affect the trajectory.

It seems that the transparent nature of the link established between the pointing form and its meaning encourages signing children to employ it earlier and more frequently than speaking children (Cormier et al., 1998; Hoiting, 2009; Hoiting & Slobin, 2007; Morgenstern et al., 2010). Pointing is also employed more frequently by signing parents than by non-signing parents, which makes it more frequent in the input that signing children receive. Thus, an abundance of pointing in language input for these children might also modulate their acquisition in language-specific ways (e.g. Kanto et al., 2024; Morgenstern et al., 2010; Wille et al., 2018). However, when it comes to acquiring different functions of pointing (deictic vs personal reference), signing children do not seem to benefit immediately from this transparent link but instead require time to figure out how to embed the function of pointing into grammatical structures. When the pointing behavior is re-analyzed by signing children and integrated into their linguistic system, it seems to play a facilitating role in the acquisition of more lexical sign tokens and types as well as in the acquisition of more complex syntax (e.g. Caët et al., 2017; Hatzopoulou, 2011; Kanto et al., 2024; Morgenstern et al., 2010; Petitto, 1987; Wille et al., 2018). However, it is also important to keep in mind that deaf parents produce more pointing signs while interacting with their deaf or hearing children. The frequency of their pointing behavior seems to be stable across time as opposed to hearing parents with hearing children (Kanto et al., 2024; Wille et al., 2018). Thus, it is quite possible that more-frequent and extended exposure to pointing signs modulates their acquisition by signing children.

We also observe an influence of the visual modality in other domains of language acquisition. For example, iconicity in lexical sign forms seems to facilitate signing children's vocabulary development, especially after 16–18 months of age (Caselli & Pyers, 2017; Novogrodsky & Meir, 2020; Sümer et al., 2017; Thompson et al., 2012). Such facilitating effects were also observed for object locations and especially for acquiring spatial terms with viewpoint (i.e. 'Left', 'Right') in sign language, which are notoriously challenging for speaking children (Karadöller et al., 2021, 2022; Sümer, 2015; Sümer & Özyürek, 2020; Sümer et al., 2014). Signing children also, in general, prefer to express both path and manner of motion early on, which might be considered a general modality effect through iconicity.

However, we do not see convincing evidence for such a facilitating role of iconicity for all domains. For example, the acquisition of spatial relations without viewpoint ('In', 'On', 'Under') seems to be similar for both signing and speaking children (Sümer, 2015; Sümer & Özyürek, 2020; Sümer et al., 2014). When we look at event descriptions in general, signing children do not acquire different aspects of these events holistically but

go through stages of segmentation of, for example, path and manner (Newport, 1981; Newport & Meier, 1985; Newport & Supalla, 1980; Sallandre et al., 2018; Supalla, 1982). Similarly, encoding other event components such as figure and ground seems to be related to children's general cognitive development rather than the language modality being acquired (Engberg-Pedersen, 2003; Hickmann, 2005; Morgan et al., 2008; Slobin et al., 2003; Sümer & Özyürek, 2022; Supalla, 1982; Tang et al., 2007).

Finally, regarding narrative development, signing children also seem to be on par with speaking children in their timeline and milestones of learning how to use different linking elements and cohesion strategies to form meaningful and connected narratives (Anthony, 2002; Ely & McCabe, 1993; Hickmann et al., 1996; Karmiloff-Smith, 1981; Lillo-Martin & Quadros, 2011; Morgan, 2000 & 2002; Morgan & Woll, 2003; Niederberger, 2004; Sümer, 2016; Vercaingne-Menard et al., 2001).

Conclusion and discussion

Human language is inherently multimodal in nature. That is, language can be expressed in either the auditory or the visual modality, and when both are accessible, language users prefer to use them simultaneously and/or in coordination with each other. Visual-spatial expressions used in sign languages employ linguistic structures and functions, and gestures accompanying speech are considered integral to spoken languages. Recent research on adult language encompasses a unified multimodal view of language where expressions in the visual modality can be explained in both spoken and sign languages. This unified multimodal language view also needs to be adapted to language acquisition. The studies reviewed earlier illustrate the role of modality-specific aspects of language as factors modulating language acquisition. This is the case not only for sign language but also when considering spoken language acquisition and gestures. Below, we discuss the specifics of this modulatory effect for both spoken languages, considering co-speech gestures, and for sign languages regarding pointing, iconicity, and simultaneity. We also point out which aspects of the language-acquisition process seem to be modality-independent. Finally, we conclude by suggesting new avenues for research in these domains and an even broader focus by encompassing all aspects and domains of language.

Pointing is an integral and frequently-used component of both spoken and sign language-acquisition processes (Cooperrider et al., 2021; Cormier et al., 2013; Sandler & Lillo-Martin, 2006), employed in similar as well as different ways in both types of languages. In spoken language acquisition, children's pointing gestures are considered to be one of the first communicative acts (e.g. Rütter & Liszkowski, 2023), which later facilitates early vocabulary (e.g. Iverson & Goldin-Meadow, 2005) and syntactic (e.g. the appearance of the two-word stage; Goldin-Meadow & Butcher, 2003) development. In sign language acquisition, pointing has also been found to be the most frequent and dominant communication form for deaf children in the early stages of their language acquisition and even more so than for hearing children (Cormier et al., 1998; Hoiting & Slobin, 2007; Morgenstern et al., 2010). However, deaf children undergo a period in which the use of pointing decreases while the functions of pointing change and become more pronominal (Caët et al., 2017; Hatzopoulou, 2011; Petitto, 1987; Pizzuto, 1990). Moreover, during this period, more lexical signs (both in terms of token and type) appear

in deaf children's lexicon, and deaf children start using complex syntax and pointing signs with different functions (Blondel & Tuller, 2008; Caët et al., 2017; Petitto, 1987). This also looks similar to the emergence of the combined use of points with verbs in spoken language development and expressions of spatial relations in speech, and it also considers the facilitating effects of pointing in speaking children for vocabulary size. However, no research has yet directly compared how pointing combines different grammatical aspects in sign and spoken language acquisition.

Iconicity also plays a prominent role in language acquisition through gestures in both sign and spoken language. It aids language acquisition while producing signs/gestures to refer to individual referents as well as locative/spatial relations among several referents. In spoken languages, iconic gestures might appear in tandem or later than verbs, at around 15 months of age. In sign languages, children are reported to benefit from iconicity as early as 10 months and even more after 30 months and older. This shows that children need to be cognitively mature enough to benefit from and understand iconic representations in gestures and signs. When cognitively apprehended, iconic representations in signs seem to provide an advantage for signing children when learning new signs (i.e. action representations). Sign language vocabularies usually consist of many iconic signs, which vary in degree and type. These aspects of iconicity (i.e. degree, type) might be an additional factor in modulating the acquisition of certain lexical signs compared to others. In spoken languages, iconic gestures might specify the semantic content of the verbs produced by children or ease their production when they emerge together.

Iconicity also seems to facilitate the acquisition of more complex structures in both spoken and sign languages. In spoken languages, iconic gestures precede linguistic forms on the relational level when expressing cognitively challenging concepts (e.g. spatial relations with viewpoints, causal relations, and narrative recall). In sign languages, relational concepts that are linguistically expressed via iconic forms such as classifier constructions (e.g. spatial relations between objects, motion events, and narratives) can be acquired earlier than their spoken language counterparts.

It is also important to note that iconicity does not always facilitate the acquisition of all domains. For example, while the acquisition of some components of motion events (i.e. manner and path) can be easier to express in sign than in spoken languages, this does not hold for other components (i.e. figure and ground), which show similar acquisition patterns in both spoken and sign languages. It is possible that certain semantic elements might be more suited for the visually-motivated mapping of concepts to forms than others.

Future directions

It is clear that a multimodal view of language acquisition needs to develop a unified framework to understand how the visual modality modulates the language-acquisition process in sign as well as spoken language, considering co-speech gestures. The above-reviewed studies show that modality-specific expressions, both in sign and gesture, play a significant role in the acquisition of many language structures from words to grammar to narrative structures. In order to understand modality specificity for spoken and sign language acquisition, we need to shift from comparing sign

language development to that of spoken language alone. Rather, we should compare sign language development to that of speech and gesture combined.

The fact that pointing and iconicity in both spoken and sign languages facilitate certain aspects of the language-acquisition process is in line with claims that children rely on sensorimotor simulations to acquire vocabulary and develop conceptual knowledge and interactions with the external world, including the physical, social, and broader cultural environment (Morgenstern, 2024, Reggin et al., 2023; Wellsby & Pexman, 2019, for reviews). In these frameworks, pointing and iconicity in gesture and sign help to simulate, conceptualize (Kita et al., 2017), and ground the concepts in the here and now and the child's body. Further theoretical and empirical studies are needed to understand how pointing and iconicity facilitate the acquisition of certain concepts and language structures at different levels. Of course, it is also possible that mapping of referents and actions to the visual modality might show slightly different trajectories in sign versus speech-plus-gesture. In sign, the visual modality carries the full load of language structure, whereas gesture develops in tandem with speech, and sign and gesture may exchange different loads during development. There is little research in this direction (see the articles by Karadöller et al., 2022, 2024 for the comparisons across TİD and Turkish speech-gesture combinations and Morgenstern et al., 2010 for comparisons across LSF and French speech-gesture combinations). Future research that compares speech-gesture combinations to sign language acquisition can delineate these aspects of language development from a unified multimodal perspective.

Conducting studies in diverse spoken and sign languages (including sign languages with varying sizes of communities, sociolinguistic structures, and histories; German, 2023; Lutzenberger et al., 2023) is also important. This is essential to uncover whether visual patterns might develop in different or similar trajectories in spoken and sign languages with different typologies (e.g. word order, path-satellite framed, pro-drop vs no-drop languages).

Finally, although there is ample evidence on the role of iconicity and pointing in language acquisition from a production perspective, this role is relatively less studied for language comprehension and processing and needs further research. Moreover, in all aspects of language, production, comprehension, and processing, the visual-spatial modality plays a role in other domains, such as phonology which was not reviewed in this article. Focusing on all aspects of language and drawing evidence from each domain will likely unravel many novel ways in which the visual modality is integral to language acquisition.


Author contributions

Dilay Z. Karadöller: Conceptualization; Writing—original draft; Writing—review & editing.

Beyza Sümer: Conceptualization; Writing—original draft; Writing—review & editing.

Aslı Özyürek: Conceptualization; Supervision; Writing—original draft; Writing—review & editing.

ORCID iDs

Dilay Z. Karadöller  <https://orcid.org/0000-0001-5160-7679>

Beyza Sümer  <https://orcid.org/0000-0002-1605-4551>

References

- Acredolo, L. A., & Goodwyn, S. (1988). Symbolic gesturing in normal infants. *Child Development*, 59(2), 450–466.
- Aksu-Koç, A., Acarlar, F., Küntay, A., Maviş, İ., Sofu, H., Topbaş, S., Turan, F., & Aktürk-Ari, B. (2019). Türkçe İletişim Gelişimi Envanteri (TİGE) el kitabı (Inventory of Communication Development in Turkish). In Detay publishing (1st ed.).
- Alibali, M. W., Evans, J. L., Hostetter, A. B., Ryan, K., & Mainela-Arnold, E. (2009). Gesture–speech integration in narrative: Are children less redundant than adults? *Gesture*, 9(3), 290–311.
- Anthony, M. (2002). *The role of American Sign Language and ‘conceptual wholes’ in facilitating language, cognition, and literacy* [Unpublished doctoral dissertation, University of California, Berkeley].
- 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.
- Austin, E. E., & Sweller, N. (2014). Presentation and production: The role of gesture in spatial communication. *Journal of Experimental Child Psychology*, 122, 92–103.
- Avçılar, G., Bilir, B., Cihankaya, Ş., Coşkun, Z., Piri, İ., & Ünal, E. (2022). How gestures support children's linguistic-communicative and cognitive development. *Studies in Psychology*, 41(3), 789–816.
- Azar, Z., Backus, A., & Özyürek, A. (2019). General and language specific factors influence reference tracking in speech and gesture in discourse. *Discourse Processes*, 56(7), 553–574.
- Bates, E. (1979). *The emergence of symbols: Cognition and communication in infancy*. Academic Press.
- Bates, E., & Dick, F. (2002). Language, gesture, and the developing brain. *Developmental Psychobiology*, 40(3), 293–310.
- Berman, R. A. (2001). Setting the narrative scene: how children begin to tell a story. In K. E. Nelson, A. Aksu-Koc, & C. E. Johnson (Eds.), *Children's language*, 10 (pp. 1–27). Lawrence Erlbaum Associates.
- Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study*. Lawrence Erlbaum Associates.
- Blake, J., Osborne, P., Cabral, M., & Gluck, P. (2003). The development of communicative gestures in Japanese infants. *First Language*, 23(1), 3–20.
- Blondel, M., & Tuller, L. (2008). Pointing in bimodal bilingual acquisition: a longitudinal study of LSF-French bilingual child. In J. Quer (Ed.), *Leading research in sign language: Selected papers from TISLR 2004* (pp. 275–292). Signum Verlag.
- Bonvillian, J. D., Orlansky, M. D., & Folven, R. J. (1990). Early sign language acquisition: Implications for theories of language acquisition. In V. Volterra, & C. J. Erting (Eds.), *Springer series in language and communication, Vol 27. From gesture to language in hearing and deaf children* (pp. 219–232). Springer.
- Bower, G. H., & Morrow, D. G. (1990). Mental models in narrative comprehension. *Science*, 247, 44–48.
- Brennan, M. (1990). *Word formation in BSL* [Unpublished doctoral dissertation, Stockholm University].
- Butcher, C., & Goldin-Meadow, S. (2000). Gesture and the transition from one- to two-word speech: When hand and mouth come together. In D. McNeill (Ed.), *Language and gesture* (pp. 235–258). Cambridge University Press.
- Butterworth, B., & Hadar, U. (1989). Gesture, speech, and computational stages: A reply to McNeill. *Psychological Review*, 96(1), 168–174.

- Caët, S., Limousin, F., & Morgenstern, A. (2017). A functional approach to self-points and self-reference in a deaf signing child and the (dis)continuity issue in child language. *Language, Interaction and Acquisition*, 8(1), 117–140.
- Calero, C. I., Shalom, D. E., Spelke, E. S., & Sigman, M. (2019). Language, gesture, and judgment: Children's paths to abstract geometry. *Journal of Experimental Child Psychology*, 177, 70–85.
- Cameron, H., & Xu, X. (2011). Representational gesture, pointing gesture, and memory recall of preschool children. *Journal of Nonverbal Behavior*, 35, 155–171.
- Capirci, O., Contaldo, A., Caselli, M. C., & Volterra, V. (2005). From action to language through gesture: A longitudinal perspective. *Gesture*, 5(1–2), 155–177.
- Capirci, O., Iverson, J. M., Pizzuto, E., & Volterra, V. (1996). Communicative gestures during the transition to two-word speech. *Journal of Child Language*, 23, 645–673.
- Caselli, N., & Pyers, J. (2017). The road to language learning is not entirely iconic: Iconicity, neighborhood density, and frequency facilitate acquisition of sign language. *Psychological Science*, 28, 979–987.
- Caselli, N., & Pyers, J. (2019). Degree and not type of iconicity affects sign language vocabulary acquisition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(1), 127–139.
- Chamberlain, C., & Mayberry, R. I. (2000). Theorizing about the relationship between ASL and reading. In C. Chamberlain, J. Morford, & R. I. Mayberry (Eds.), *Language acquisition by eye* (pp. 221–260). Lawrence Erlbaum Associates.
- Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition*, 23(1), 43–71.
- Church, R. B., Kelly, S. D., & Lynch, K. (2000). Immediate memory for mismatched speech and representational gesture across development. *Journal of Nonverbal Behavior*, 24, 151–174.
- Clark, E. V. (1973). Non-linguistic strategies and the acquisition of word meanings. *Cognition*, 2(2), 161–182.
- Colletta, J. M., Pellenq, C., & Guidetti, M. (2010). Age-related changes in co-speech gesture and narrative: Evidence from French children and adults. *Speech Communication*, 52(6), 565–576.
- Colonesi, C., Stams, G. J. J. M., Koster, I., & Noom, M. J. (2010). The relation between pointing and language development: A meta-analysis. *Developmental Review*, 30(4), 352–366.
- Conlin, K. E., Mirus, G. R., Mauk, C. E., & Meier, R. (2000). The acquisition of first signs: Place, handshape, and movement. In C. Chamberlain, J. P. Morford, & R. I. Mayberry (Eds.), *Language acquisition by eye* (pp. 51–69). Lawrence Erlbaum Associates.
- Cooperrider, K., Fenlon, J., Keane, J., Brentari, D., & Goldin-Meadow, S. (2021). How pointing is integrated into language: Evidence from speakers and signers. *Frontiers in Communication*, 6, 567774.
- Cormier, K., Mauk, C., & Repp, A. (1998). Manual babbling in deaf and hearing Infants: A longitudinal study. In *Proceedings of the Twenty-ninth Annual Child Language Research Forum* (pp. 55–61). CSLI Publications.
- Cormier, K., Schembri, A., & Woll, B. (2013). Pronouns and pointing in sign languages. *Lingua*, 137, 230–247.
- Cuxac, C. (1999). French Sign Language: Proposition of a structural explanation by iconicity. *Proceedings of the International Gesture Workshop on Gesture-Based Communication in Human-Computer Interaction*, 165–184.
- de Beuzeville, L. (2006). *Visual and linguistic representation in the acquisition of depicting verbs: A study of native signing deaf children of Auslan* [Unpublished doctoral dissertation, University of Sydney].

- Demir, Ö. E., Levine, S., & Golden-Meadow, S. (2015). A tale of two hands: Children's early gesture use in narrative production. *Journal of Child Language, 42*(03), 662–681.
- de Ruiter, J. P. (1998). *Gesture and speech production* [Doctoral dissertation, MPI Series of Psycholinguistics].
- de Ruiter, J. P. (2000). The production of gesture and speech. In D. McNeill (Ed.), *Language and gesture*. Cambridge University Press.
- Ely, R., & McCabe, A. (1993). Remembered voices. *Journal of Child Language, 20*(3), 671–696.
- Emmorey, K. (2002). *Language, cognition, and the brain: Insights from sign language research*. Lawrence Erlbaum Associates.
- Emmorey, K. (2014). Iconicity as structure mapping. *Philosophical Transactions of the Royal Society of London: Series B, Biological Sciences, 369*.
- K., Emmorey (Ed.). (2003). *Perspectives on classifier constructions in sign languages*. Lawrence Erlbaum Associates.
- Engberg-Pedersen, E. (2003). How composite is a fall? Adults' 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.
- Ertaş, S., Koşukulu-Sancar, S., Ger, E., Liszkowski, U., & Küntay, A. C. (2023). Relation of infants' and mothers' pointing to infants' vocabulary measured directly and with parental reports. *Infancy, 28*(6), 1007–1029.
- Esteve-Gibert, N., Lævenbruck, H., Dohen, M., & d'Imperio, M. (2022). Pre-schoolers use head gestures rather than prosodic cues to highlight important information in speech. *Developmental Science, 25*(1), e13154.
- Furman, R., Küntay, A. C., & Özyürek, A. (2014). Early language-specificity of children's event encoding in speech and gesture: Evidence from caused motion in Turkish. *Language, Cognition and Neuroscience, 29*(5), 620–634.
- Furman, R., Özyürek, A., & Allen, S. (2006). Learning to express causal events across languages. What do speech and gesture patterns reveal? In *Proceedings of the 30th Annual Boston University Conference on Language Development* (pp. 190–201). Cascadilla Press.
- Furman, R., Özyürek, A., & Küntay, A. C. (2010). Early language-specificity in Turkish children's caused motion event expressions in speech and gesture. In *The 34th Boston University Conference on Language Development* (pp. 126–137). Cascadilla Press.
- Ger, E., Küntay, A. C., Ertaş, S., Koşukulu-Sancar, S., & Liszkowski, U. (2023). Correlates of infant pointing frequency in the first year. *Infancy, 28*(6), 986–1006. <https://doi.org/10.1111/inf.12560>
- German, A. (2023). Abrupt grammatical reorganization of an emergent sign language: The expression of motion in Zinacantan Family Homesign. *Diachronica*. <https://doi.org/10.1075/dia.22039.ger>
- Göksun, T., Hirsh-Pasek, K., & Golinkoff, R. M. (2010). How do preschoolers express cause in gesture and speech? *Cognitive Development, 25*(1), 56–68.
- Goldin-Meadow, S. (2007). Pointing sets the stage for learning language – And creating language. *Child Development, 78*, 741–745.
- Goldin-Meadow, S. (2014). Widening the lens: What the manual modality reveals about language, learning and cognition. *Philosophical Transactions of the Royal Society, Series B, 369*(1651), 20130295. <https://doi.org/10.1098/rstb.2013.0295>
- Goldin-Meadow, S., & Brentari, D. (2017). Gesture, sign, and language: The coming of age of sign language and gesture studies. *Behavioral and Brain Sciences, e4046*.
- Goldin-Meadow, S., & Butcher, C. (2003). Pointing toward two-word speech in young children. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 85–107). Lawrence Erlbaum Associates.

- Goldin-Meadow, S., Nusbaum, H., Kelly, S., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science, 12*(6), 516–522.
- Greenfield, P., & Smith, J. (1976). *The structure of communication in early language development*. Academic Press.
- Gu, S., Chen Pichler, D., Kozak, L. V., & Lillo-Martin, D. (2022). Phonological development in American Sign Language-signing children: Insights from pseudosign repetition task. *Frontiers in Psychology, 13*, 921047.
- Guevara, I., & Rodríguez, C. (2023). Developing communication through objects: Ostensive gestures as the first gestures in children's development. *Developmental Review, 68*, 101076.
- Gullberg, M., Kita, S., & Boster, C. (2008). Gesture and speech in interaction: An overview. In S. Duncan, J. Cassell, & E. T. Levy (Eds.), *Gesture and the Dynamic Dimension of Language: Essays in honor of David McNeill* (pp. 259–283). John Benjamins.
- Hagoort, P., & Özyürek, A. (2024). Extending the architecture of language from a multimodal perspective. *Topics in Cognitive Science*. Advance online publication. doi:10.1111/tops.12728.
- Hatzopoulou, M. (2011). The emergence of pronominal pointing in Greek Sign Language. *Proceedings of the 9th International Gesture Workshop*, Athens, Greece, 117–120.
- Hickmann, M. (2005). Children's discourse: Person, space, and time across languages. *Journal of Child Language, 32*(2), 463–469.
- Hickmann, M., Hendriks, H., Roland, F., & Liang, J. (1996). The marking of new information in children's narratives: A comparison of English, French, German, and Mandarin Chinese. *Journal of Child Language, 3*, 591–610.
- Hoiting, N. (2009). *The myth of simplicity: Sign language acquisition by Dutch deaf toddlers* [Unpublished doctoral dissertation, Rijksuniversiteit].
- Hoiting, N., & Slobin, D. (2007). From gestures to signs in the acquisition of sign language. In S. Duncan, J. Cassell, & E. Levy (Eds.), *Gesture and the Dynamic Dimension of Language: Essays in Honour of David McNeill* (pp. 51–65). John Benjamins.
- Holler, J., & Levinson, S. (2019). Multimodal language processing in human communication. *Trends in Cognitive Sciences, 23*(8), 639–652.
- Hostetter, A. B., & Alibali, M. W. (2019). Gesture as simulated action: Revisiting the framework. *Psychonomic Bulletin & Review, 26*(3), 721–752.
- Iverson, J. M., & Goldin-Meadow, S. (1997). What's communication got to do with it? Gesture in children blind from birth. *Developmental Psychology, 33*(3), 453–467.
- Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science, 16*(5), 367–371.
- Johnston, J. R., & Slobin, D. I. (1979). The development of locative expressions in English, Italian, Serbo-Croatian and Turkish. *Journal of Child Language, 6*(3), 529–545.
- Kanto, L., Laakso, M., & Huttunen, K. (2015). Differentiation in language and gesture use during early bilingual development of hearing children of deaf parents. *Bilingualism: Language and Cognition, 18*(4), 769–788.
- Kanto, L., Laakso, M., & Huttunen, K. (2024). Use of pointing in parent-child interactions by hearing children of deaf and hearing parents: A follow-up from 1- to 3-years of age. *Journal of Child Language, 51*, 411–433.
- Kantor, R. (1980). The acquisition of classifiers in American Sign Language. *Sign Language Studies, 28*, 198–208.
- Karadöller, D. Z., Peeters, D., Manhardt, F., Özyürek, A., & Ortega, G. (2024). Iconicity and gesture jointly facilitate learning of second language signs at first exposure in hearing nonsigners. *Language Learning*. <https://doi.org/10.1111/lang.12636>
- Karadöller, D. Z., Sümer, B., & Özyürek, A. (2021). Effects and non-effects of late language exposure on spatial language development: Evidence from deaf adults and children. *Language Learning and Development, 17*(1), 1–25.

- Karadöller, Z. D., Sümer, B., Ünal, E., & Özyürek, A. (2021). Spatial language use predicts spatial memory of children: Evidence from sign, speech, and speech-plus-gesture. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (pp. 672–678). The Cognitive Science Society.
- Karadöller, Z. D., Sümer, B., Ünal, E., & Özyürek, A. (2022). Sign advantage: Both children and adults' spatial expressions in sign are more informative than those in speech and gestures combined. *Journal of Child Language*, 1–27.
- Karadöller, Z. D., Ünal, E., Sümer, B., Göksun, T., Özer, D., & Özyürek, A. (2019). Children but not adults use both speech and gesture to produce informative expressions of left-right relations. In the 44th Annual Boston University Conference on Language Development (BUCLD 44).
- Karmiloff-Smith, A. (1981). The grammatical marking of thematic structure in the development of language production. In W. Deutsch (Ed.), *The child's construction of language*. Academic Press.
- Karnopp, L. B. (2008). Sign phonology acquisition in Brazilian Sign Language. In *Sign Languages: Spinning and Unraveling the Past, Present and Future*, ed. R. M. deQuadros (RJ. Brazil: Editora Arara Azul. Petrópolis), 204–218.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge University Press.
- Kirk, E., Donnelly, S., Furman, R., Warmington, M., Glanville, J., & Eggleston, A. (2022). The relationship between infant pointing and language development: A meta-analytic review. *Developmental Review*, 64, 101023.
- Kita, S. (2003). *Pointing: Where language, culture, and cognition meet*. Psychology Press.
- Kita, S. (2009). Cross-cultural variation of speech-accompanying gesture: A review. *Language and Cognitive Processes*, 24(2), 145–167. DOI: 10.1080/01690960802586188
- Kita, S., Alibali, M. W., & Chu, M. (2017). How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychological Review*, 124(3), 245.
- Kita, S., & Emmorey, K. (2023). Gesture links language and cognition for spoken and signed languages. *Nature Reviews Psychology*, 2(7), 407–420.
- Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, 48(1), 16–32.
- Koşkulu, S., Küntay, A. C., & Uzundag, B. A. (2021). Maternal behaviors mediate the relationship between socioeconomic status and joint attention. *Journal of Applied Developmental Psychology*, 75, 101291.
- Krauss, R. M., Chen, Y., & Chawla, P. (1996). Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us? In M. P. Zanna (Ed.), *Advances in experimental social psychology*. Academic Press.
- Krauss, R. M., Chen, Y., & Gotfexnum, R. F. (2000). Lexical gestures and lexical access: A process model. In D. McNeill (Ed.), *Language and Gesture* (pp. 261–283). Cambridge University Press
- Krauss, R. M., Dushay, R. A., Chen, Y., & Rauscher, F. (1995). The communicative value of conversational hand gestures. *Journal of Experimental Social Psychology*, 31, 533–552.
- Küntay, A. (2002). Development of the expression of indefiniteness: Presenting new referents in Turkish picture-series stories. *Discourse Processes*, 33, 77–101.
- Lillo-Martin, D. (2012). Utterance reports and constructed action. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign language: An international handbook* (pp. 365–387). Mouton DeGruyter.
- Lillo-Martin, D., & Henner, J. (2021). Acquisition of sign languages. *Annual Review of Linguistics*, 7, 395–419.

- Lillo-Martin, D., & Quadros, R. (2011). Acquisition of the syntax–discourse interface: The expression of point of view. *Lingua*, 121(4), 623–636.
- Liszkowski, U., Brown, P., Callaghan, T., Takada, A., & De Vos, C. (2012). A prelinguistic gestural universal of human communication. *Cognitive Science*, 36(4), 698–713.
- Liszkowski, U., & R  ther, J. (2021). Ontogenetic origins of infant pointing. In N. Gontier, A. Lock, & C. Sinha (Eds.), *Oxford handbook of human symbolic evolution*. Oxford University Press.
- Liszkowski, U., & Tomasello, M. (2011). Individual differences in social, cognitive, and morphological aspects of pointing at 12 months of age. *Cognitive Development*, 26, 16–29.
- L  ke, C., Grimminger, A., Rohlfing, K. J., Liszkowski, U., & Ritterfeld, U. (2017). In infants' hands: Identification of preverbal infants at risk for primary language delay. *Child Development*, 88(2), 484–492.
- Lutzenberger, H., Fikkert, P., de Vos, C., & Crasborn, O. (2023). Development of sign phonology in Kata Kolok. *Journal of Child Language*, 1–34.
- Mamus, E., Speed, L. J., Rissman, L., Majid, A., &   zy  rek, A. (2023). Lack of visual experience affects multimodal language production: Evidence from congenitally blind and sighted people. *Cognitive Science*, 47(1), e13228.
- Matthews, D., Behne, T., Lieven, E., & Tomasello, M. (2012). Origins of the human pointing gesture: A training study. *Developmental Science*, 15(6), 817–829.
- McGillion, M., Herbert, J. S., Pine, J., Vihman, M., DePaolis, R., Keren-Portnoy, T., & Matthews, D. (2017). What paves the way to conventional language? The predictive value of babble, pointing, and socioeconomic status. *Child Development*, 88(1), 156–166.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.
- McNeill, D. (2005). *Gesture and thought*. University of Chicago Press.
- Meier, R. P., & Newport, E. L. (1990). Out of the hands of babes: On a possible sign advantage in language acquisition. *Language*, 66, 1–23.
- Melinger, A., & Levelt, W. J. (2004). Gesture and the communicative intention of the speaker. *Gesture*, 4(2), 119–141.
- Mitchell, R. E., & Karchmer, M. (2004). Chasing the mythical ten percent: Parental hearing status of deaf and hard of hearing students in the United States. *Sign Language Studies*, 4(2), 138–163.
- Morgan, G. (2000). Discourse cohesion in sign and speech. *International Journal of Bilingualism*, 4(3), 279–300.
- Morgan, G. (2002). The encoding of simultaneity in children's BSL narratives. *Journal of Sign Language and Linguistics*, 5(2), 127–161.
- Morgan, G., Herman, R., Barriere, I., & Woll, B. (2008). The onset and the mastery of spatial language in children acquiring British Sign Language. *Cognitive Development*, 23, 1–19.
- Morgan, G., & Woll, B. (2003). The development of reference switching encoded through body classifiers in British Sign Language. In K. Emmorey (Ed.), *Perspectives on Classifier Constructions in Sign Languages* (pp. 297–310). Lawrence Erlbaum Associates.
- Morgan, G., & Woll, B. (Eds.). (2002). *Directions in sign language acquisition*. John Benjamins.
- Morgenstern, A. (2024). Gesture and first language development. In A. Cienki (Ed.), *The Cambridge Handbook of Gesture Studies* (pp. 368–397). Cambridge University Press.
- Morgenstern, A., Ca  t, S., Collombel-Leroy, M., Limousin, F., & Blondel, M. (2010). From gesture to sign and from gesture to word: pointing in deaf and hearing children. *Gesture*, 10(2–3), 172–202.

- Newport, E. (1981). Constraints on structure: Evidence from American Sign Language and language learning. In W. A. Collins (Ed.), *Minnesota Symposium on Child Psychology, 14*, (pp. 93-124). Lawrence Erlbaum.
- Newport, E., & Meier, R. (1985). Acquisition of American Sign Language. In D. Slobin (Ed.), *The crosslinguistic study of language acquisition* (pp. 881-938). Erlbaum.
- Newport, E., & 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. 81-114). Verlag Chemie.
- Niederberger, N. (2004). *Capacités langagières en langue des signes française et en français écrit chez l'enfant sourd bilingue* [Unpublished doctoral dissertation, University of Geneva].
- Novogrodsky, R., & Meir, N. (2020). Age, frequency, and iconicity in early sign language acquisition: Evidence from the Israeli Sign Language MacArthur-Bates Communicative Development Inventory. *Applied Psycholinguistics, 41*, 817-845.
- Orlansky, M., & Bonvillian, J. (1984). The role of iconicity in early sign language acquisition. *Journal of Speech & Hearing Disorders, 49*, 287-292.
- Ortega, G., Sümer, B., & Özyürek, A. (2017). Type of iconicity matters in the vocabulary development of signing children. *Developmental Psychology, 53*, 89-99.
- Özçalışkan, Ş., Gentner, D., & Goldin-Meadow, S. (2014). Do iconic gestures pave the way for children's early verbs? *Applied Psycholinguistics, 35*(6), 1143-1162.
- Özçalışkan, Ş., & Goldin-Meadow, S. (2005). Gesture is at the cutting edge of early language development. *Cognition, 96*(3), B101-B113.
- Özçalışkan, Ş., & Goldin-Meadow, S. (2009). When gesture-speech combinations do and do not index linguistic change. *Language and Cognitive Processes, 24*(2), 190-217.
- Özçalışkan, Ş., Lucero, C., & Goldin-Meadow, S. (2016). Is seeing gesture necessary to gesture like a native speaker? *Psychological Science, 27*(5), 737-747.
- 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*.
- Özer, D., & Göksun, T. (2020). Gesture use and processing: A review on individual differences in cognitive resources. *Frontiers in Psychology, 11*, 573555.
- Özyürek, A. (2018). Role of gesture in language processing: Toward a unified account for production and comprehension. In S. Rueschemeyer, & M. G. Gaskell (Eds.), *The Oxford Handbook of Psycholinguistics, Oxford Library of Psychology* (pp. 592-560).
- 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.
- Özyürek, A., & Woll, B. (2019). Language in the visual modality: Cospeech gesture and sign language. In P. Hagoort (Ed.), *Human language: From genes and brain to behavior*. MIT Press.
- Padden, C., Hwang, S., Lopic, R., & Seegers, S. (2014). Tools for language: Patterned iconicity in sign language nouns and verbs. *Topics in Cognitive Science, 7*, 81-94.
- Peeters, D., & Özyürek, A. (2016). This and that revisited: A social and multimodal approach to spatial demonstratives. *Frontiers in Psychology, 7*, 222.
- Perniss, P. (2007). *Space and iconicity in German Sign Language (DGS)* [Unpublished doctoral dissertation]. Max Planck Institute for Psycholinguistics.
- Perniss, P. (2018). Why we should study multimodal language. *Frontiers in Psychology, 9*(1109).
- Perniss, P., Özyürek, A., & Morgan, G. (2015). The influence of the visual modality on language structure and conventionalization: insights from sign language and gesture. *Topics in Cognitive Science, 7*(1), 2-11.

- Perniss, P., Thompson, R., & Vigliocco, G. (2010). Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in Psychology, 1*(227).
- Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: From a world of experience to the experience of language. *Philosophical Transactions of the Royal Society B, 369*.
- Perry, M., Church, R. B., & Goldin-Meadow, S. (1992). Is gesture-speech mismatch a general index of transitional knowledge? *Cognitive Development, 7*(1), 109–122.
- Peterson, C. (1990). The who, when, and where of early narratives. *Journal of Child Language, 17*, 433–456.
- Peterson, C., & McCabe, A. (1983). *Developmental psycholinguistics: Three ways of looking at a child's narrative*. Plenum.
- Petitto, L. A. (1987). On the autonomy of language and gesture: evidence from the acquisition of personal pronouns in American Sign Language. *Cognition, 27*(1), 1–52.
- Pizzuto, E. (1990). The early development of deixis in American Sign Language: what is the point?. In V. Volterra, & C. J. Erting (Eds.), *From gesture to language in hearing and deaf children* (pp. 142–152). Springer.
- Pizzuto, E., & Volterra, V. (2000). Iconicity and transparency in sign languages: a crosslinguistic cross-cultural view. In K. Emmorey, & H. Lane (Eds.), *The signs of language revisited: An anthology to honor Ursula Bellugi and Edward Klima* (pp. 261–286). Lawrence Erlbaum Associates.
- Reggin, L. D., Franco, L. E. G., Horchak, O. V., Labrecque, D., Lana, N., Rio, L., & Vigliocco, G. (2023). Consensus Paper: Situated and Embodied Language Acquisition. *Journal of Cognition, 6*(1), 1–16.
- Rohrer, P. L., Florit-Pons, J., Vilà-Giménez, I., & Prieto, P. (2022). Children use non-referential gestures in narrative speech to mark discourse elements which update common ground. *Frontiers in Psychology, 12*, 661339.
- Rowe, M. L., & Goldin-Meadow, S. (2009a). Differences in early gesture explain SES disparities in child vocabulary size at school entry. *Science, 323*(5916), 951–953.
- Rowe, M. L., & Goldin-Meadow, S. (2009b). Early gesture selectively predicts later language learning. *Developmental Science, 12*(1), 182–187.
- Rowe, M. L., Özçaliskan, S., & Goldin-Meadow, S. (2006). The added value of gesture in predicting vocabulary growth. In *Proceedings of the 30th Annual Boston University Conference on Language Development* (pp. 501–512). Cascadilla Press.
- Russo, T. (2004). Iconicity and productivity in sign language discourse: An analysis of three LIS discourse registers. *Sign Language Studies, 4*(2), 164–197.
- Rüther, J., & Liszkowski, U. (2023). Ontogeny of index-finger pointing. *Journal of Child Language, 1–17*.
- 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.
- Salo, V. C., Rowe, M. L., & Reeb-Sutherland, B. C. (2018). Exploring infant gesture and joint attention as related constructs and as predictors of later language. *Infancy, 23*(3), 432–452.
- Sandler, W., & Lillo-Martin, D. (2006). *Sign language and linguistic universals*. Cambridge University Press.
- Sansavini, A., Guarini, A., Zuccarini, M., Lee, J. Z., Faldella, G., & Iverson, J. M. (2019). Low Rates of Pointing in 18-Month-Olds at Risk for Autism Spectrum Disorder and Extremely Preterm Infants: A Common Index of Language Delay? *Frontiers in Psychology, 10*, 2131.

- Sauter, M., Uttal, D. H., Alman, A. S., Goldin-Meadow, S., & Levine, S. C. (2012). Learning what children know about space from looking at their hands: The added value of gesture in spatial communication. *Journal of Experimental Child Psychology*, *111*(4), 587–606.
- Schembri, A. (2003). Rethinking classifiers in signed languages. In K. Emmorey (Ed.), *Perspectives on classifier constructions in sign languages* (pp. 3-34). Lawrence Erlbaum Associates.
- Schick, B. (1990). The effects of morphosyntactic structures on the acquisition of classifier predicates in ASL. In C. Lucas (Ed.), *Sign language research: Theoretical issues* (pp. 358-374). Gallaudet University Press.
- Sevcikova Sehyr, Z., & Emmorey, K. (2019). The perceived mapping between form and meaning in American Sign Language depends on linguistic knowledge and task: Evidence from iconicity and transparency judgments. *Language and Cognition*, *11*(2), 208–234.
- Slobin, D., Hoiting, N., Kuntze, M., & Lindert, R. (2003). A cognitive / functional perspective on the acquisition of classifiers. In K. Emmorey (Ed.), *Perspectives on classifier constructions in sign languages* (pp. 271-296). Lawrence Erlbaum Associates.
- Slonimska, A. (2022). The role of iconicity and simultaneity in efficient communication in the visual modality: Evidence from LIS (Italian Sign Language). PhD Thesis, Radboud University, Nijmegen.
- 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.
- Slonimska, A., Özyürek, A., & Capirci, O. (2022). Simultaneity as an emergent property of efficient communication in language: A comparison of silent gesture and sign language. *Wiley Interdisciplinary Reviews: Cognitive Science*, *46*(5), 13133.
- Slonimska, A., & Özyürek, A. (in press). Methods to study evolution of iconicity in sign languages. In L. Raviv, & C. Boeckx (Eds.), *The Oxford Handbook of Approaches to Language Evolution*. Oxford University Press.
- Stites, L. J., & Özçalışkan, Ş. (2017). Who did what to whom? Children track story referents first in gesture. *Journal of Psycholinguistic Research*, *46*, 1019–1032.
- 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.
- Sümer, B. (2016). Scene–setting and reference introduction in sign and spoken languages: What does modality tell us?. In N. Ketrez, & B. Haznedar (Eds.), *Trends in language acquisition research: The acquisition of Turkish in childhood* (pp. 193-220). John Benjamins.
- Sümer, B., Grabitz, C., & Küntay, A. (2017). Early produced signs are iconic: Evidence from Turkish Sign Language. In the *39th Annual Conference of the Cognitive Science Society (CogSci 2017)* (pp. 3273–78). Cognitive Science Society.
- Sümer, B., & Özyürek, A. (2020). No effects of modality in development of locative expressions of space in signing and speaking children. *Journal of Child Language*, *47*(6), 1101–1131.
- Sümer, B., & Özyürek, A. (2022). Cross–modal investigation of event component omissions in language development: A comparison between signing and speaking children. *Language, Cognition and Neuroscience*, *37*(8), 1023–1039.
- Sümer, B., Perniss, P., Zwitserlood, I., & Özyürek, A. (2014). Learning to express ‘left–right’ & ‘front–behind’ in sign versus spoken language. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Meeting of the Cognitive Science Society* (pp. 1550-1555). Cognitive Science Society.
- Supalla, T. (1982). *Structure and acquisition of verbs of motion and location in American Sign Language* [Unpublished doctoral dissertation, University of California, San Diego].
- Talmy, L. (2000). *Toward a cognitive semantics*. MIT Press.

- 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. (2001). *Language from the body: Iconicity and metaphor in American Sign Language*. Cambridge University Press.
- Thompson, R. (2011). Iconicity in language processing and acquisition: What signed languages reveal. *Language and Linguistics Compass*, 5(9), 603–616.
- Thompson, R., Vinson, D., Woll, B., & Vigliocco, G. (2012). The road to language learning is iconic: Evidence from British Sign Language. *Psychological Science*, 23(12), 1443–1448.
- Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development*, 78, 705–722.
- Tomaszewski, P., Krzysztofciak, P., Morford, J. P., & Ezlakowski, W. (2022). Effects of age-of-acquisition on proficiency in Polish Sign Language: Insights to the critical period hypothesis. *Frontiers in Psychology*, 13, 866339.
- Umiker-Sebeok, D. J. (1979). Preschool children's intra-conversational narratives. *Journal of Child Language*, 6, 91–109.
- Ünal, E., Mamus, E., & Özyürek, A. (2023). Multimodal encoding of motion events in speech, gesture and cognition. *Language and Cognition* Advance online publication. doi:10.1017/langcog.2023.61.
- Vercaingne-Menard, A., Godard, L., & Labelle, M. (2001). The emergence of narrative discourse in two young children. In V. Dively (Ed.), *Sign Languages: Discoveries from international research* (pp. 120-132). Gallaudet University Press.
- Vigliocco, G., Perniss, P., & Vinson, D. (2014). Language as a multimodal phenomenon: implications for language learning, processing and evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1651), 20130292.
- Volterra, V., Capirci, O., Caselli, M. C., Rinaldi, P., & Sparaci, L. (2017). Developmental evidence for continuity from action to gesture to sign/word. *Language, Interaction and Acquisition*, 8(1), 13–41.
- Wagner, S. M., Nusbaum, H., & Goldin-Meadow, S. (2004). Probing the mental representation of gesture: Is handwaving spatial? *Journal of Memory and Language*, 50, 395–407.
- Wellsby, M., & Pexman, P. M. (2019). Learning labels for objects: Does degree of sensorimotor experience matter? *Languages*, 4(1), 3.
- Wille, B., Mouvet, K., Vermeerbergen, M., & Van Herreweghe, M. (2018). Flemish Sign Language development: A case study on Deaf mother – deaf child interactions. *Functions of Language*, 252, 289–322.