# 6 Language in the Visual Modality: Co-speech Gesture and Sign Language

ASLI ÖZYÜREK AND BENCIE WOLL

As humans, our ability to communicate and use language is instantiated not only in the vocal modality but also in the visual modality. The main examples of this are sign languages and (co-speech) gestures used in spoken languages. Sign languages, the natural languages of Deaf[1] communities, use systematic and conventionalized movements of the hands, face, and body for linguistic expression (Brentari, 2010; Emmorey, 2002; Klima & Bellugi, 1979; Stokoe, 1960). Co-speech gestures, though nonlinguistic, are produced and perceived in tight semantic and temporal integration with speech (Kendon, 2004; McNeill, 1992, 2005). Thus, language—in its primary face-to-face context (as is the case both phylogenetically and ontogenetically)—is a multimodal phenomenon (Kendon, 2014; Vigliocco, Perniss, & Vinson, 2014). Expression in the visual modality appears to be an intrinsic feature of human communication. As such, our models of language need to take these visual modes of communication into account and provide a unified framework for how the semiotic and expressive resources of the visual modality are recruited in both spoken and sign languages and what the consequences of this recruitment are for cognitive architecture and processing of language. Most research on language, however, has focused on spoken or written language and has rarely considered the visual context in which it is embedded as a means of understanding our linguistic capacity.

The aim of the current chapter is to outline what the expressive resources of language look like both in spoken and sign languages, and what their roles are in communication, in cognition, and specifically in language processing. This will be set in the context of an exploration of the cognitive and neural architecture of language, addressing such issues as commonalities and contrasts in the brain's network for auditory and visual components of human communication, taking gestures in spoken languages as well as sign languages into account.

Even though historically, co-speech gestures and signs have been studied separately and belong to different traditions of literature, recently there have been attempts to achieve a joint perspective that seeks to understand the role of the visual modality in language in general by studying both gesture and sign (e.g., Goldin-Meadow & Brentari, 2015; Perniss, Özyürek, & Morgan, 2015). Studies comparing hearing speakers' gestures with systems used in emerging sign languages and *homesign* systems (e.g., Goldin-Meadow, 2003; Senghas, Kita, & Özyürek, 2004) have shown that as gestures move toward sign language, idiosyncratic gestures used with speech are replaced by conventionalized expressions, and linguistic properties increase (McNeill, 1992). Unlike co-speech gestures, sign languages are complete linguistic systems exhibiting linguistic structure and language-specific constraints at the phonological, morphological, lexical, grammatical, and discourse levels.

Even though sign languages differ from gestures in significant ways, there is now a clearer understanding that the pragmatic, semantic, and cognitive functions employed by co-speech gestures during the use of spoken languages are also evident in the use of sign languages. Furthermore there are similarities arising from the iconic and indexical properties afforded by the visual modality and these properties may be difficult to express within the auditory affordances of the speech channel. The joint perspective that we will adopt here may shed new light on how communication in the visual modality reflects modality-specific as well as modality-independent aspects of our language capacity and on the extent to which a common cognitive and neural architecture underpins linguistic and nonlinguistic communication across modalities.

## *1. Visual Modality in Spoken Language: Co-speech Gestures*

Kendon (2004) defined *gestures* as visible actions of the hand, body, and face that are intentionally used to communicate and are expressed together with the verbal utterance.

Gestures are universal in the sense that all speaking communities around the world are known to produce gestures, even though the communicative and social value of gesturing and thus the frequency of its use may differ among cultures or show variation across individuals (Chu, Meyer, Foulkes, & Kita, 2014; Kita, 2009). Whenever individuals speak, gesture—a unique human ability starting around the first year of life—is involved. The link between gesturing and speaking also appears innate in the sense that it does not require people to see other people gesturing—congenitally blind people also gesture while speaking (Iverson & Goldin-Meadow, 1998; Özçalışkan, Lucero, & Goldin-Meadow, 2016).

In spite of the close links of gesture to language use, most grammatical theories and linguistic descriptions omit gestural specifications. Recognition of the relation of gesture to language has been even more recent (Kendon, 1986; McNeill, 1992) than the recognition that sign languages are on a par with spoken languages (e.g., Klima & Bellugi, 1979; Stokoe, 1960). One of the reasons for this is that linguistic theories have generally taken what can be spoken or written as their main domain of investigation and have been mostly occupied with aspects of language that denote things arbitrarily and categorically (e.g., words, phrases, sentences).

Less attention has been allocated to expressions of other aspects of communication such as *indicating* and *depicting*. Gestures are frequently found to serve the latter functions. For *indicating*, speakers use pointing gestures to place and locate referents in the shared communicative context and for *depicting*, they use so-called representational or iconic gestures to represent virtual objects and events in the gesture space around them (Clark, 2016). In many face-to-face contexts, without *indicating* and *depicting*, successful communication may be hard to achieve.

1.1. Forms and Functions of Gestures in Language   Co-speech gestures manifest themselves in different form and meaning pairings as well as in different semiotic types and functions during communication (Clark, 1996; Kendon, 2004; McNeill, 1992, 2005). While some gestures, such as representational gestures, abstract points, and beats, occur as accompaniments to speech, other gestures, such as emblems or interactional gestures, can replace or complement speech in an utterance or can be used without speech, as will be discussed further.[2] Gestures allow speakers to be co-expressive and create *composite signals* (Clark, 1996) by making use of the specific representational capacities of each modality (visual and auditory).

Different forms of gestures fulfill different semantic and communicative functions when used with speech. For example, in so-called *emblems* there is an arbitrary relationship between their form and the meaning they convey, and they serve very similar functions to lexicalized words. On the other hand, *representational* gestures (also referred to as *iconic* gestures) bear a more visually motivated relation between their form and the referent, action, or event they represent. For example, a stirring hand movement accompanying a spoken utterance about cooking bears a resemblance in form to the actual act of stirring. Even though such gestures are visually motivated, the meaning they convey relies heavily on the speech they accompany. Experimental studies have shown that in the absence of speech, the meaning of these gestures is highly ambiguous and not at all transparent from their form (Krauss, Dushay, Chen, & Rauscher, 1995). When these gestures occur, they almost always overlap with semantically relevant speech—which supports the disambiguation of their meaning: speech and such gestures form a co-expressive ensemble.

Representational gestures vary in terms of their semiotic characteristics—that is, in the way they can represent objects, actions, or events—revealing modality-specific ways to convey or depict information such as the different visual perspectives of speakers to events, size, three-dimensional characteristics, shapes, relative spatial relations among objects (Debreslioska, Özyürek, Gullberg, & Perniss, 2013; McNeill, 1992; Müller, 2009; Tversky, 2011). Their meaning comes from the holistic representation of the image they represent rather than from a combinatorial representation of individual meaning units such as those we see in spoken languages. As such they mostly fulfill the *depictive* aspects of communication—reenacting objects and events talked about in a "visible" way in the shared space between the speaker and the interlocutor.

*Points* are co-expressive accompaniments of demonstrative forms and pronouns in discourse, specifying referents, places, and locations (for a review, see Peeters & Özyürek, 2016). Their form is not informative, but the direction of the point links the referent to the object/space, fulfilling indexical aspects of reference. Such pointing gestures can either be oriented toward concrete objects (when targeting objects or places in the here-and-now of the participants' discourse) or point to meaningful abstract locations in the gesture space in front of the speaker. Pointing to abstract gesture space allows speakers to express coherent relationships among the referents that figure in their discourse by locating them in the gesture space (e.g., McNeill, Cassell, & Levy, 1993; Perniss & Özyürek, 2015).

Finally, *beats* (meaningless repetitive hand movements) can be used to emphasize parts of speech at the information structure level to express focus on certain parts of speech. The so-called *interactional* gestures (e.g., a gesture for *I don't know*) are used to regulate different aspects of dialogic interaction (e.g., expressing stance, turn taking) between the interlocutors (Bavelas, Chovil, Lawrie, & Wade, 1992).

Thus when all types of gestures are considered, we see that they have very similar functions to lexical, semantic, pragmatic, discourse, and interactional features of spoken languages—albeit conveyed in a different format and thus allowing speakers to also convey aspects of messages (e.g., iconic, indexical) that cannot be conveyed through the affordances and the linguistic structures of spoken language (arbitrary, categorical). In this chapter we will mostly focus on representational and pointing gestures, discuss whether and how they are integrated with the spoken language in their use, and consider their roles in language processing and their neural correlates.

1.2. Role of Gesture in Language Processing The production of co-speech gestures is closely linked to the production of the linguistic message conveyed in speech. This is evident especially when we consider the close timing between gestures and speech during production and the co-expressive meaning they convey. With respect to timing, a co-speech gesture is produced along with the relevant part of speech and together they express a communicative act. People do not gesture the entire time they are speaking. Nor is it the case that each and every gesture is accompanied by speech. The important point, rather, is that when people produce co-speech gestures, those gestures are almost always temporally aligned in some meaningful way with a spoken utterance. With respect to meaning, gesture and speech have been argued to share an underlying conceptual message (Bernardis & Gentilucci, 2006; McNeill, 1992). In this sense, gesture and speech are considered to be co-expressive, although the contributions of these communicative channels may be supplementary to, or redundant with, one another, and the representational formats of speech and gesture differ (de Ruiter, Bangerter, & Dings, 2012; McNeill, 1992), in other words, speech is categorical and discrete, whereas gesture is continuous and analog.[3]

More evidence that speech and gesture are tightly linked comes from studies demonstrating that the timing and co-expressive meaning alignment between speech and gesture vary systematically between typologically different languages (Defina, 2016; Floyd, 2016; Gu, Mol, Hoetjes, & Swerts, 2017; Kita & Özyürek, 2003). One demonstration of this phenomenon is how people speak and gesture about motion events across languages. Languages vary in how they linguistically encode the path and manner of motion events (Talmy, 1985), and the gestures that speakers of a given language use have been shown to reflect this variation. Speakers of Japanese and Turkish (*contra* English) are unlikely to tightly package information about both path and manner within a linguistic clause (e.g., rather than *she runs down the stairs* they will say *she runs and goes down the stairs*). Likewise, when gesturing about motion events, speakers of Japanese and Turkish (*contra* English) are unlikely to encode both path and manner within a single gesture and prefer either to represent only one element or split the two elements into separate expressive gestures (Kita & Özyürek, 2003). In English, however, since it is grammatically possible to express manner and path within one linguistic clause (*she ran down the stairs*), speakers are more likely to express both components in a single gesture. That is, gestures in each language seem to be shaped by the syntactic and semantic packaging of information at the clause level. This is confirmed by recent evidence from blind speakers of Turkish and English, who show similar patterns of speech and gesture to those of sighted speakers (Özçalışkan et al., 2016). Recently, Defina (2016) showed that speakers of Avatime (an indigenous spoken language of Australia), which has serial verb constructions, also package two or more semantic elements into a single gesture.

A second area where gesture may vary in relation to linguistic structure is in the expression of spatial frames of reference. Speakers of languages (e.g., Guugu Yimidhirr) that preferentially express spatial relationships using cardinal directions (e.g., east, west, north, and south) rather than egocentric ones (e.g., left, right), also tend to express cardinal relationships in their gestures (Haviland, 1993). Gu et al. (2017) also showed that Chinese speakers' time metaphors based along the vertical dimension are also reflected in their gestures, unlike English speakers' speech and gestures that reflect time along a horizontal axis. Finally, in some languages, gestures may consistently express semantic information not expressed in speech. In a recent study, Floyd (2016) showed that speakers of the Brazilian indigenous language Nheengatú use pointing gestures "adverbially" to indicate time. While speech in Nheengatú gives information about time in general terms (e.g., morning,

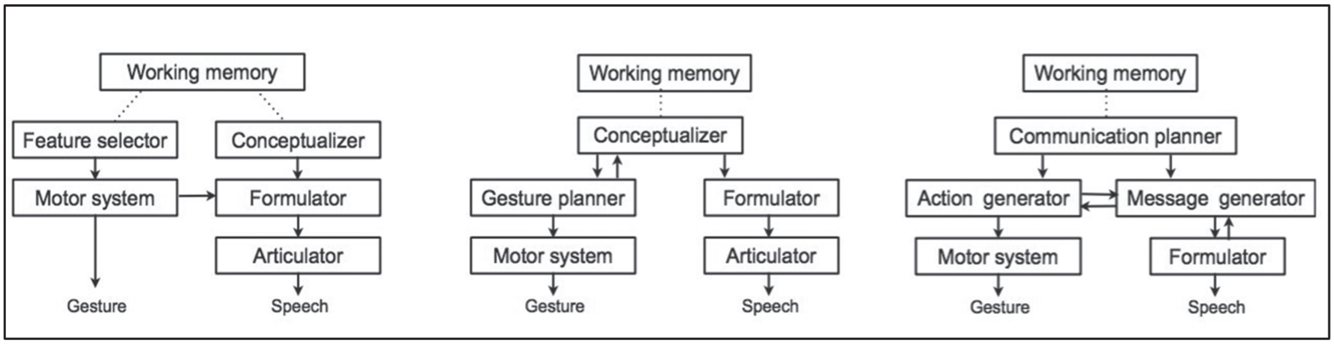


FIGURE 6.1 Adapted schematic overview of different models in relation to speech and gesture production (from left to right: Morrel-Samuels & Krauss, 1992; de Ruiter, 2000; Kita & Özyürek, 2003) (taken from Wagner, Malisz, & Kopp [2014], reprinted with permission).

noon, evening), celestial pointing gestures indicate more specific times of the day (e.g., 10 in the morning). Speakers' judgments about these composite utterances have also shown that they rely on gestures for further specification of time. Floyd argued that there is no *a priori* reason for linguistic properties not to develop in the visual practices (i.e., gestures) that accompany spoken language.

The studies presented show that speech and gesture (at least for representational and pointing gestures) are orchestrated in different ways in different languages, and that gestures are informed by the lexical, syntactic, and pragmatic possibilities of different languages. This poses interesting questions about how speakers coordinate speech and gesture during production to achieve the semantic and temporal alignment they co-express—also considering crosslinguistic variation in this respect. This brings us to language production models proposed so far that take into account the close links between speech and gesture.

The speech and gesture production models suggested to date have either viewed gestures as representing an independent but parallel expressive channel (de Ruiter, 2000; Krauss, Chen, & Gottesman, 2000) or have assumed that speech and gesture interact during the formulation of the linguistic message at different levels (Kita & Özyürek, 2003; McNeill, 1992). These models also differ in terms of whether they consider gesture to be part of the communicatively intended message (de Ruiter, 2000; Kita & Özyürek, 2003) or to be produced independently (Krauss et al., 2000).[4]

According to Krauss and colleagues, gestures are generated from images in working memory, which might help to prime lexical items cross-modally. They are not necessarily assumed to be communicative. In de Ruiter's (2000) model, gestures are generated from conceptualizations that are intended to be part of the communicative message, but during production the information to be conveyed in both modalities is split and expressed through independent channels. In contrast, according to more interactionist models, for example that of McNeill (1992, 2005), gesture and speech are derived from an initial single unit, which McNeill refers to as a *growth point*, composed of both types of representations, imagistic and linguistic. Both gesture and speech are manifestations of this combined unit of representation. According to another interactionist view—the interface hypothesis—proposed by Kita and Özyürek (2003) with a slightly revised version proposed by Chu and Kita (2016), representational gestures and speech are best characterized as originating from different representations: gestures from imagistic/action representations and language from propositional representations. During the language production process, however, both representations interact at the level of linguistic formulation. Figure 6.1 shows different models proposed for speech and gesture production.

As in the case of production, there is growing evidence that gestures are integrated with speech comprehension. It has been a long-standing finding that addressees pick up information from gestures that accompany speech. Even though most models of speech and gesture have focused on production, recent research has also provided ample evidence that addressees integrate the information coming from both modalities during comprehension, in other words, they are perceived as communicative. Listeners/viewers pay attention to iconic gestures and pick up the information that they encode. For example, Kelly, Barr, Church, and Lynch (1999) showed participants video stimuli where gestures conveyed information additional to that conveyed in speech (producing a gesture pantomiming drinking, while the speech is "I stayed up all night") and asked them to write down what they heard. In addition to the speech they heard, participants included information that was conveyed only in gesture and not in speech (i.e., "I stayed up drinking all night"). In another study, Beattie and Shovelton (1999) showed that listeners/viewers answered questions about the size and relative position of objects in a speaker's message more accurately when gestures were part of the description and conveyed information additional to speech than when they heard speech only.

In a priming study by Kelly, Özyürek, and Maris (2010), participants were presented with action primes (e.g., someone chopping vegetables) followed by targets comprising speech accompanied by gesture. They were asked to press a button if what they heard in speech or saw in gesture depicted the action prime. Participants related primes to targets more quickly and accurately when they contained congruent information (speech: "chop"; gesture: chop) than when they contained incongruent information (speech: "chop"; gesture: twist). Moreover, the degree of incongruence between overlapping speech and gesture affected processing, with fewer errors for weak incongruities (speech: "chop"; gesture: cut) than for strong incongruities (speech: "chop"; gesture: open). This indicates that in comprehension, the semantic relations between the two channels are taken into account, providing evidence against independent processing of the two channels. Furthermore and crucially, this effect was bidirectional and was found to be similar when either speech or gesture targets matched or mismatched the action primes. That is, gesture influences processing of speech and speech influences processing of gesture. Further research has shown that gestures also show semantic priming effects. For example, Yap, So, Yap, and Tan (2010) showed that iconic gestures shown without speech (highly conventionalized gestures such as flapping both hands at the side of the body to mean *bird*) prime the sequentially presented spoken target words.

Finally, evidence for semantic integration between representational gestures and speech is also found in many neurocognitive studies. Several studies have shown that comprehension of iconic gestures activates brain processes known to be involved in semantic processing of speech. First of all, gestures modulate the electrophysiological component N400 (e.g., Özyürek, Willems, Kita, & Hagoort, 2007), which has previously been found to be sensitive to the ease of semantic comprehension of words in relation to a previous context. Second, viewing iconic gestures in the context of speech (matched or mismatched) recruits the left-lateralized frontal–posterior temporal network (left inferior frontal gyrus [LIFG], medial temporal gyrus [MTG], and superior temporal gyrus/sulcus [STG/S]) known to be involved in semantic integration of words in sentences (Hagoort & van Berkum, 2007; Özyürek, 2014; see Emmorey & Özyürek, 2014, for a broader overview). Recent studies by Peeters and colleagues (Peeters, Hagoort, & Özyürek, 2015; Peeters & Özyürek, 2016; Peeters, Snijders, Hagoort, & Özyürek, 2017) have also shown that match/mismatch between an expressed element in speech (e.g., "apple") and pointing to a referent (e.g., an apple) evokes semantic integration, as indexed by modulation of N400, and also recruits LIFG and MTG and STG/S—as found for iconic gestures in the above-mentioned studies.

1.3. Conclusions: Gesture These findings show that gestures are an integral part of language at the level of semantics, syntax, pragmatics, and discourse. Gestures, because of the affordances of the visual modality, can subserve *indicating* and *depicting* aspects of communication—albeit in a different representational format from that found in speech. In visible ways, they allow the grounding of concepts conveyed to the visual here-and-now by the speech component, either by linking speech to objects visibly through pointing, or by reenacting them in a virtual space created among the conversational participants, to convey analog representations of events. In doing so, they play a role in the conceptualization, formulation, and comprehension of utterances. They also recruit language networks in the brain during processing. Thus, although their visual and semiotic properties differ from the linguistic units of spoken language, in relation to the cognitive and neural architecture of language, gestures are an integral part of our language capacity. They are integrated into the language-specific semantic and structural aspects of spoken language and interact with spoken language during the production and processing of utterances. Their role also underscores the claim that language is not a fully modular system at the level of production, comprehension, or in its neural architecture. Many of the features of language—semantic, pragmatic, or even syntactic, can be subserved in orchestration with gestures.

Now we turn to how the visual modality is recruited in sign languages: languages created by Deaf communities, who communicate entirely in the visual modality, and where visible bodily articulators alone express all functions of language and communication.

## 2. Visual Modality in Sign Language

Following the groundbreaking work by linguists and cognitive scientists over the last 50 years, it is now generally recognized that sign languages of Deaf communities, such as ASL (American Sign Language), BSL

(British Sign Language),[5] and Sign Language of the Netherlands are not idiosyncratic compilations of silent gestures/pantomimes used by deaf people but are structured and processed in a similar manner to spoken languages. The one striking difference is that they operate in a wholly nonauditory, visual-spatial modality. In this section we first summarize which aspects of sign language structure and processing are similar to those found in spoken language regardless of the differences in modality. Second, we illustrate which aspects of sign languages reveal modality-specific features and how these compare not only to speech but also to the gestural properties that we see in spoken languages as discussed in section 1. Sign languages thus offer unique insights about the modality-independent versus modality-specific (e.g., iconic, embodied) aspects of our language capacity and its cognitive and neural architecture.

Before we begin it is important to say a few words about the social and linguistic contexts in which sign languages are most likely to emerge. Just as all languages need a community of users, sign languages need a Deaf community, which can only come into existence where deaf people are in contact with one another. Although there are descriptions of deaf people's signing going back hundreds of years, the establishment of schools for deaf children, starting in the late 18th century in Europe, triggered the creation of Deaf communities and sign languages as we know them today. At these schools, communication between children and teachers resulted in the conventionalization of what were previously widely varying *home signs*, gestures used by isolated deaf individuals. In many countries, education for deaf children has only recently begun and this can provide an environment in which new sign languages can emerge. Kegl, Senghas, and Coppola (1999) described how the establishment of the first school for deaf children in Nicaragua in the 1980s led to the beginnings of a national sign language. There are also communities where an unusually high incidence of deafness in a small village results in a sign language used by both deaf and hearing people, even in the absence of schools (Sandler, Meir, Padden, & Aronoff, 2005; see Woll & Ladd, 2010, for a review of a number of these).

Where deaf children are *not* exposed to a sign language, they invariably develop systematic gestural communication within their families. Such communication is called *homesigning* (Coppola & Newport, 2005; Goldin-Meadow, 2003). Homesigning systems do not have a full linguistic structure, unlike sign language. They may have rudimentary features of linguistic structure such as a lexicon, simple morphology, segmentation, and consistent word order. However, there is evidence that homesigners cannot master a sign language fully if they are exposed to one only after childhood (e.g., later than six to seven years of age). In other words, homesigning does not serve as a "first" language when homesigners are later exposed to a conventionalized sign language (Mayberry, 2010).

Although the emergence and sociolinguistic context of sign languages differs from spoken languages, which are—with rare exceptions—transmitted by native speakers and in which the only "new" languages are creoles derived from contact between two different spoken languages, the sign languages used by Deaf communities display many of the linguistic structures we see in spoken languages. It is clear that when deaf individuals are able to communicate with each other, in a very short amount of time the communication system they use shows substantial divergence from the gestures used by speaking people. For example, Senghas, Özyürek, and Kita (2004) investigated co-speech gestures used by Spanish speakers in expressing simultaneously occurring manner and path (for example, *run-downhill*) and compared them to signs used by three cohorts of Nicaraguan signers (cohort 1: deaf homesigners brought together as adults; cohort 2: deaf homesigners brought together as children and exposed to the communication of cohort 1; cohort 3: deaf children exposed to cohorts 1 and 2). While gestures produced by Spanish speakers expressed manner and path elements in one gesture, signer cohorts 2 and 3 started to segment them—akin to the separate verbs for manner and path that we see in some spoken languages—and to combine them sequentially to express the simultaneity of manner and path. A recent study by Özyürek, Furman, and Goldin-Meadow (2014) also observed similar developmental changes in Turkish deaf children's homesigning systems compared to their hearing Turkish caregivers' gestures and the gestures of Turkish-speaking children. Even though Turkish homesigners also started segmenting elements of manner and path into separate units, they were more likely to produce gestures that expressed both elements in one unit, and their patterns resembled those of the first cohort of Nicaraguan signers rather than the second and third cohort signers (Özyürek et al., 2014). These studies indicate that in the context of deafness, visual communication goes beyond the expressive possibilities of the gestures used by speakers. However, as we will discuss, some similarities between the two systems (e.g., the role of iconicity) exist because of the shared affordances of modality between gesture and sign and because sign languages may make use of gestures, as spoken languages do.

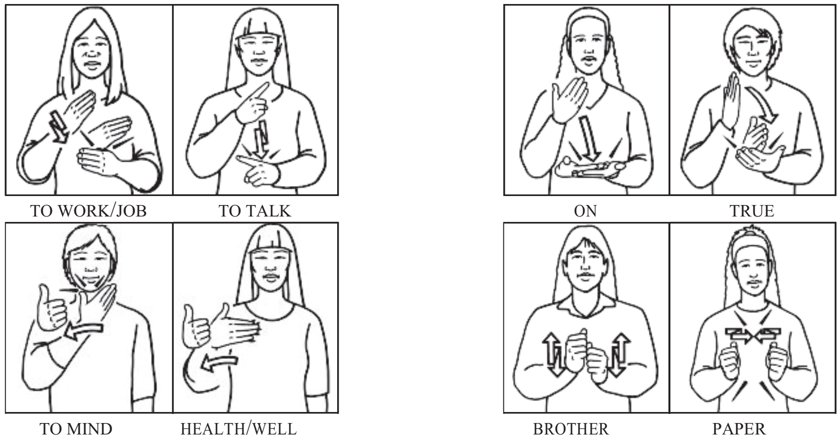


Figure 6.2 Minimal pairs in BSL.

In section 2.1, a brief description will be provided of both modality-independent and modality-specific characteristics of languages, including phonology, lexicon, and the exploitation of space for grammatical purposes. Although structuring at different linguistic levels is similar across signed and spoken languages, the ways in which these structures can be expressed show modality-specific patterning—this is discussed in sections 2.1.1 and 2.1.2. We will also indicate areas of sign language structure for which there has been a debate about whether they should be analyzed solely in terms of abstract, categorical, grammatical structures or whether their analysis needs to take into account instantiations of gestural form-meaning mappings and/or iconic correspondences. Finally, the processing and neural organization of sign languages will be discussed in relation to those structures that are modality-independent or modality-specific.

2.1. Modality-Independent Aspects of Sign Languages As a result of research undertaken in the past half century on the longer-established sign languages of Europe and North America, it is now recognized that the sign languages used by Deaf communities are complex natural human languages and that they are not derived from the spoken languages of the surrounding hearing communities—even though, due to contact, some features from spoken languages can influence sign language structures.

Meier (2002) listed a number of the noneffects of modality (i.e., the shared properties of spoken and signed languages).

a. Conventional vocabularies: learned pairing of form and meaning
b. Duality of patterning: meaningful units built of meaningless sublexical units, whether orally or manually produced
c. Productivity: new vocabulary may be added to signed and spoken languages
d. Syntactic structure:
   i. Same word classes in spoken and signed languages: nouns, verbs, and adjectives
   ii. Embedding of clauses and recursion
   iii. Trade-offs between word order and morphological complexity in how grammatical relations are marked
e. Acquisition: similar timetables for acquisition of signed and spoken language.

Next we describe some of the features of sign languages that exhibit similar linguistic structures to those found in spoken languages, albeit expressed differently due to different affordances of the visual modality. In many cases, however, especially where visual motivation is evident (i.e., iconicity), these features may also resemble those found in co-speech gesturing.

*2.1.1. Phonology* Since Stokoe (1960), linguists have seen the phonological structure of signs as consisting of simultaneous combinations of configuration(s) of the hand(s), a location where the sign is articulated, and movement—either a path through signing space or an internal movement of the joints in the hand. Each is understood to be a part of the phonology, because changing one of these parameters can create a minimal pair (see figure 6.2). There have been considerable modifications to Stokoe's framework since 1960, but this model has remained the basic description of sign language phonology.

*2.1.2. Morphology* Sign language morphology tends to manifest itself in simultaneous combinations of meaningful handshapes, locations, and movements, rather than in affixation. In derivational morphology, for example, handshape can change to

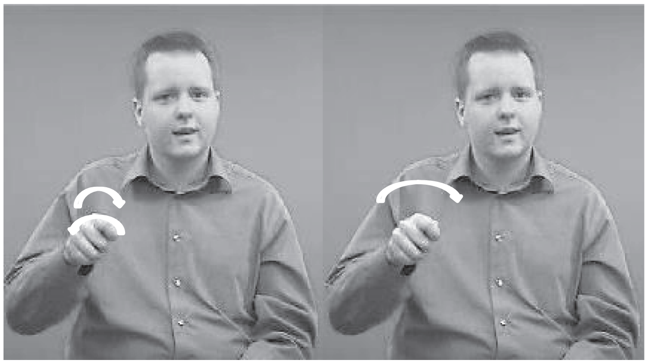


Figure 6.3 Movement contrast between derivationally related BSL signs KEY (A) and LOCK (B).


Figure 6.4 (A) Example of syntactic space. The referent "film" is located in the upper right of signing space by means of an index, but this does not map onto any real-world location. (B) Example of topographic space (spatialized grammar). In the predicate, the referents "book" and "pen" are replaced with classifiers for "flat object" and "long thin object," respectively, and these handshapes are located adjacent to each other and at the height in signing space of the sign "table."

reflect numbers. For example in BSL,[6] n-WEEKS, n-O'CLOCK, and n-YEARS-OLD are articulated with conventionalized location and movement, while the handshapes (e.g., of *three* or *five*) incorporated into the time signs indicate the number (e.g., 3-WEEKS, 5-YEARS-OLD).

Signs referring to objects and actions may also differ only in movement, so the verbs LOCK, READ-NEWSPAPER, and EAT are made with long movements, compared to the derivationally related nouns KEY, NEWSPAPER, and FOOD, which have short, repeated movements (figure 6.3).

Other morphological features are also shown by changes in movement and location. Thus, degree is shown through size, speed, onset speed, and length of hold in a movement, with, for example, LUCKY having a smaller, faster, and smoother movement than VERY-LUCKY. Movement changes conveying temporal aspect are frequently visually motivated, so that repeated actions or events are shown through repetition of the sign; duration of an event is paralleled by duration of the sign (signs for shorter events being articulated for less time than signs for longer events) and when an event is interrupted suddenly, the movement of the sign is also interrupted, as shown for example in encoding of telicity (i.e., whether the event expressed is abrupt or continuous) in signs (Strickland et al., 2015). Signs can also change handshape to indicate how a referent is handled. So (I) HAND-OVER-A-FLOWER-TO-EACH-OF-YOU has the same movement as (I) HAND-OVER-AN-ICE-CREAM-TO-EACH-OF-YOU—with the hand moving away from the signer to virtual or real recipients—but with different handshapes that incorporate the shape of a hand as if holding a flower or an ice cream.

Both spoken and signed languages articulate lexical items sequentially. In sign languages, however, the availability of two hands enables the extensive use of simultaneously articulated structures (Vermeerbergen, Leeson, & Crasborn, 2007). Two hands can be used to represent the relative locations and movements of two referents in space and their topic–comment relations. Thus simultaneity is an aspect of sign languages that allows the expression of distinctions found in spoken languages but in a different, visual, format.[7] We will see that this feature also allows the depiction of some iconic structures similar to those found in gestures accompanying spoken languages.

Sign languages exploit the use of space for grammatical purposes, preferring dimensionality (the analog representation of size and shape) and simultaneity in syntax, while spoken languages prefer linearization and affixation. In earlier literature, on ASL in particular (e.g., Poizner, Klima, & Bellugi, 1987; Padden, 1988), two uses of space for linguistic purposes were contrasted. Topographic space was described as being used to depict spatial relationships and to map referents onto a representation of real space (spatialized grammar), while syntactic space was conceived of as an exploitation of space for purely grammatical purposes, without any mapping to real-world spatial relationships (see figure 6.4 for examples of sentences illustrating these different uses of space). These models were very closely related to the linguistic descriptions used for spoken languages, such as pronouns and agreement in person and number. We will see that recently some of these analyses have been questioned and many researchers now see them as more closely involving iconic gestural features.

2.2. Modality-Specific Aspects of Sign Language: Iconicity and Gestural Elements While significant progress has been made by treating sign languages as having many features in common with those described for spoken languages, the ways in which sign languages differ from spoken languages may have implications for how they are processed and understood by language users.

Perhaps the most obvious of these differences is the way that sign languages exploit the visual modality through iconicity. Iconicity refers to the resemblance between an object or action and the word or sign used to represent that object or action. Early studies emphasized that iconicity might not play a major role in sign language structure or processing. Klima and Bellugi (1979) provided a detailed discussion of what is meant by a sign being *iconic*, pointing out that (a) many signs in ASL (and other sign languages) are noniconic; (b) iconic signs vary from one sign language to another, since different visual motivations for a sign form may be selected (OLD represented by wrinkles in BSL and by a beard in ASL); and (c) they are conventionalized forms, subject to regular processes of phonological change. There have been contrasting findings in relation to the role iconicity plays in sign language processing at the lexical level. Poizner, Bellugi, and Tweney (1981) showed that highly iconic signs were not more easily remembered than signs that are highly opaque; Atkinson, Marshall, Woll, and Thacker (2005) reported that signers with word-finding difficulties following stroke found iconic signs no easier to retrieve than noniconic signs (also see section 2.3 on sign language aphasia and gesture); and Meier, Mauk, Cheek, and Moreland (2008) suggested that iconicity is not a factor in the early sign language acquisition of deaf children.

More recent studies have suggested that iconicity does have a role in the structure of the lexicon and grammar of sign language as well as in processing and learning (Emmorey, 2014; Perniss, Thompson, & Vigliocco, 2010; Strickland et al., 2015; Taub, 2001). For example, Thompson, Vinson, Woll, and Vigliocco (2012) reported that iconic signs are learned earlier than noniconic signs are. Thompson, Vinson, and Vigliocco (2010) also reported effects of iconicity on phonological decision tasks. Studies of ASL at the narrative and discourse level have suggested that in order to understand ASL, the addressee must process *surrogates* (Liddell, 2003) or *depictions* (Dudis, 2004) produced by the signer. Both Liddell and Dudis argued that the signer creates a visual scene and "paints a picture" for the addressee (close to Clark's [2016] notion of *depiction* and *indication* discussed in section 1.1 in relation to gesture), utilizing the visual medium and signing space to convey meaning in ways that are difficult, if not impossible, in spoken languages—except when conveyed through gesture. Thus these authors suggested that sign languages are produced and understood in ways that are very different from spoken languages. However, recent research shows that iconic depiction and indication are not sign language–specific constructions but rather modality-specific forms, as revealed both in signing (ASL) and speaking, when we take into account the gestures speakers use (Quinto-Pozos & Parrill, 2015).

Researchers working on a number of sign languages (Engberg-Pedersen, 1993 [Danish Sign Language]; Liddell, 1990 [ASL]) have argued that the sign language linguistic structures that use sign space to express space and syntactic features are not based on abstract linguistic properties but on inherent analog locative relationships among people, objects, and places. Liddell's later work (2003) has gone further, proposing that verbs such as GIVE are a mix of gestural and linguistic elements and analyzing such signs as being composed of a linguistic part expressed by the handshape and a gestural part linking the referent to a locus.

An increasing number of sign language researchers now assume that mixed forms (i.e., structures involving both linguistic and nonlinguistic components) exist, particularly in classifier constructions. Classifiers are handshapes that provide information about the class by which a noun can be described (see figure 6.4). For example, in the BSL verb VEHICLE-MOVE, the handshape varies according to the class of vehicle (e.g., bicycle, ship, car/bus). Constructions involving classifiers, while originally described in purely linguistic terms as polymorphemic (Supalla, 1986), or semantically multicomponential (Slobin et al., 2003; Morgan & Woll, 2007), have also been seen as blends of nonlinguistic gesture and linguistic structures (Liddell, 2003).

Schembri, Jones, and Burnham (2005) compared the representation of motion events by hearing nonsigners using gesture without accompanying speech and by native signers of three different sign languages. They found that constructions in the three sign languages were strikingly similar, but also that the descriptions of motion events produced by the hearing gesturers showed significant points of correspondence with the signed constructions. In both cases, the location and movement parameters were similar and the handshape component showed the greatest differences. They argue therefore that these data are consistent with the claim that verbs of motion and location in sign languages are blends of gestural and linguistic elements.

Recently, Wilbur (2008) and Strickland et al. (2015) also showed in comparisons of a number of different sign languages that the phonetic realization of the telicity of verbs also contains iconic features (e.g., telic verbs such as *decide*, whose meaning has an intrinsic culmination point, are marked with rapid deceleration to an abrupt stop; whereas atelic verbs such as *ponder*, with no intrinsic culmination, are not). These differences can also be detected by hearing nonsigners and are expressed similarly in gestures.

Thus all this research makes evident the point that sign languages are composed of gestural and linguistic elements just as spoken languages are. However more research is needed to understand the extent to which gestural elements look similar across signed and spoken languages (Özyürek, 2012).

2.3. Processing of Sign Language The main focus in this section is on neuroscience studies of sign language processing. However, it should be noted that substantial numbers of behavioral studies have also explored the cognitive mechanisms underlying sign language perception and production and shown similarities between spoken and sign languages. These include studies of lexical segmentation (Orfanidou, Adam, Morgan, & McQueen, 2010); lexical access—exploring lexicality effects in tasks using signs and nonsigns (Carreiras, Gutiérrez-Sigut, Baquero, & Corina, 2008; Corina & Emmorey 1993); and priming effects at all linguistic levels. Examples of the latter include studies of priming effects related to phonology (Dye and Shih, 2006), morphology (Emmorey, 1991), syntax (Hall, Ferreira, & Mayberry, 2015), and semantics (Bosworth & Emmorey, 2010). There has been relatively little research on production but studies of tip-of-the-finger phenomena (Thompson, Emmorey, & Gollan, 2005) and slips of the hand (Hohenberger, Happ, & Leuninger, 2002) reveal phonological and semantic effects in sign retrieval. For comprehensive reviews of sign language processing research, we refer to Emmorey (2002).

It has been suggested that the right hemisphere (RH) is known to be dominant for a number of visuospatial processing abilities, suggesting that there is a right hemisphere advantage for simultaneous processing (Hellige, 1993). One might therefore infer that spoken language, being more linearized, is left-lateralized, while sign language, which is perceived visually and uses space for grammatical purposes, might be either right-lateralized or show more mixed lateralization. However, case studies of Deaf individuals with acquired brain damage and neuroimaging studies of healthy Deaf subjects have provided evidence for left hemisphere dominance in processing sign languages as well as spoken languages. Deaf signers, like hearing speakers, exhibit language disturbances when left hemisphere (LH) cortical regions are damaged (e.g., Hickok, Love-Geffen, & Klima, 2002; Marshall, Atkinson, Smulovitch, Thacker, & Woll, 2004; Poizner, Klima, & Bellugi 1987; for a review, see MacSweeney, Capek, Campbell, & Woll, 2008). Right hemisphere damage, although it can disrupt visual-spatial abilities (including some involved in sign language processing of spatial language), does not produce sign aphasia (Atkinson et al., 2005).

More evidence for the claim that similar brain structures are involved in signed and spoken languages in healthy subjects comes from a study by MacSweeney et al. (2002), which compared BSL presented to deaf and hearing native signers with audiovisual English presented to hearing monolingual speakers. The perception of BSL and audiovisual English sentences recruited very similar neural systems in native users of those languages. Both languages recruited the perisylvian cortex in the LH. However, there was also RH recruitment by both languages and no differences in the extent of recruitment of the RH by BSL and English (see figure 6.5, columns 1 and 3; see also Bedny & MacSweeney, chapter 37 of this volume). Presumably, this reflects the contribution of both hemispheres to comprehension of the sentences presented, whether the language was spoken or signed. Figure 6.5 also provides evidence that cerebral organization of language in Deaf signers utilizes a left-lateralized frontal/temporal network as in speakers.

At first glance, the robust nature of the left-lateralized sign production system does not seem to be modulated by the iconic forms in sign languages. As we have mentioned, many signs bear some iconic relationship to their real-world referents. For example, a sign may trace the outline of a referent (e.g., HOUSE, which traces the outline of the roof and walls of a house) or may refer to a particular visual characteristic of the referent (e.g., CAT, which traces a cat's whiskers, but means *cat*, rather than whiskers). Despite this, sign-aphasic patients are often unable to produce iconic signs in response to prompts such as "Show me the sign for 'toothbrush,'" although they can produce the same actions elicited as pantomimed gesture—"How do you brush your teeth?" (see Corina et al., 1992; Marshall et al., 2004). That is, they show a dissociation between sign language (impaired) and gesture (unimpaired). Imaging studies, too, suggest that iconicity fails to influence the cortical regions activated in the production of sign language (see Emmorey et al., 2004, and San Jose-Robertson, Corina, Ackerman, Guillemin, & Braun, 2004 for further

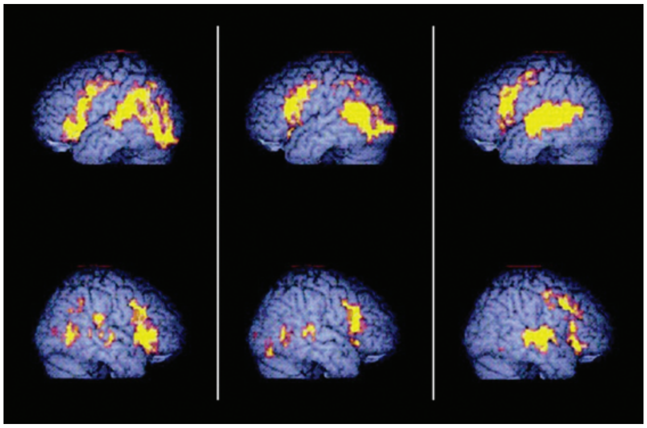

Figure 6.5 Color-rendered images of the brain depicting (group) functional MRI activation. Regions activated by BSL perception in Deaf and hearing native signers (first and second columns, respectively) and by audiovisual speech perception in hearing nonsigners (third column). For language in both modalities, and across all three groups, activation is greater in the left than the right hemisphere and perisylvian regions are engaged (from MacSweeney, Woll, Campbell, McGuire, et al. [2002], reprinted with permission).

discussion of the role of iconicity in signed language production).

On the other hand, some modality-specific influences on the brain's processing of sign language have been observed. Neville et al. (1998) suggested that the bilateral activation they observed for ASL could be related to ASL's use of space for linguistic purposes. However, MacSweeney et al.'s (2002) study of BSL sentence processing (figure 6.5), which included sentences with spatial grammar, showed left lateralization comparable to that of spoken languages. These discrepancies underline the need for further studies, using a variety of different signed languages and also a variety of signers with different language experiences and backgrounds. We will see, however, that there is more evidence for the recruitment of the right hemisphere for sign language during processing of spatial language (Emmorey, McCullough, Mehta, Ponto, & Grabowski, 2013) and from lesion studies.

The next question one might ask is the extent to which sign language processing is similar to or different from gesture processing. One way to answer this question has been to compare and contrast the processing of linguistically well-formed material with material that may be superficially similar but which cannot be analyzed linguistically. This type of contrast addresses the question of whether the brain bases for sign language processing are the same as those for the processing of other visual manual actions. In one study, MacSweeney et al. (2004) contrasted BSL utterances with gestures derived from TicTac, the gestural code used by racetrack bookmakers to signal betting odds to each other. The stimuli were modeled by a deaf native signer who constructed "sentences" using hand gestures derived from TicTac codes, adding nonmanual markers (facial gestures) to make these sequences more similar to BSL. Both types of input caused extensive activation throughout the left and right superior temporal lobe when compared to watching the model at rest. That is, much of the upper part of the temporal lobe is involved in attending to gestural displays whether these are linguistically structured or not. However, the brains of the signers who viewed the displays discriminated between the inputs: BSL activated a left-sided region located at the junction of the left posterior superior temporal gyrus and the supramarginal gyrus in the parietal lobe (see figure 6.6) much more than TicTac did. This difference was not due to perceptual differences in the visual quality of the stimuli, because hearing people with no BSL knowledge showed no differences in activation between BSL and TicTac in this region. This region thus appears to be a language-specific region that is not sensitive to the modality of the language it encounters.

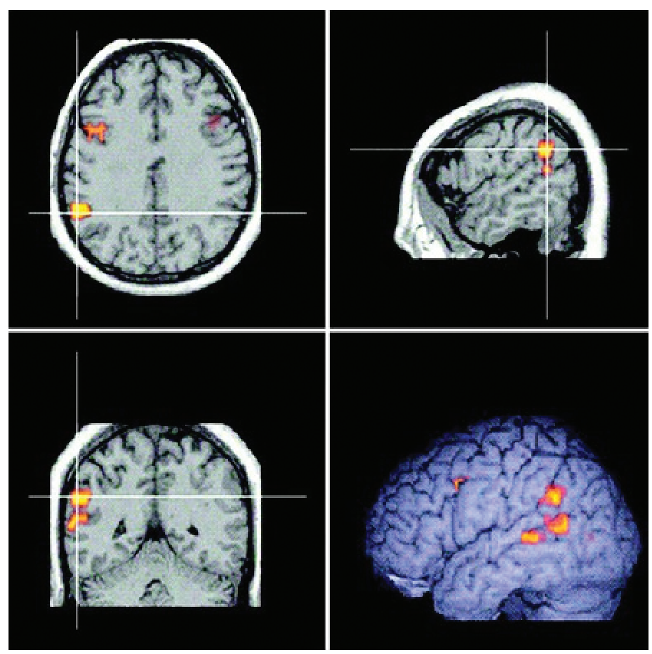

Figure 6.6 Colored regions are those recruited to a greater extent by BSL perception than TicTac (nonsense gestures) in Deaf native signers. We interpret these regions as being particularly interested in language processing. Activation up to 5 mm below the cortical surface is displayed. Crosshairs are positioned at Talairach coordinates: $x=-58$, $y=-48$, $z=31$. This is the junction of the inferior parietal lobule and the superior temporal gyrus (from MacSweeney et al. [2004], reprinted with permission).

Even though the comparison of sign and meaningless gesture has activated different cerebral networks, as in MacSweeney et al.'s (2002) study, meaningful gestures (pantomimes) and spoken language are found to recruit overlapping areas. Xu, Gannon, Emmorey, Smith, and Braun (2009) found that comprehending pantomimes (e.g., opening a jar) and their spoken language equivalents both engaged LIFG and left posterior MTG. The authors suggest these are part of a general semantic network for human communication when it comes to comprehension. Emmorey, Xu, Gannon, Goldin-Meadow, and Braun (2010) also found very similar patterns of activation within bilateral posterior temporal cortex when deaf signers passively viewed pantomimed actions and ASL signs, but with evidence for greater activation in the LIFG when viewing ASL signs.

Another domain where we see modality-dependent differences between signed and spoken languages is in the use of spatialized structures to express locative relationships. Liddell (2003) argued that such constructions are partially gestural in nature, and their grammatical status has been debated. Several studies have found differential disruptions in the use and comprehension of sign language sentences that involve spatialized grammar compared to other grammatical constructions. For example, Atkinson et al. (2005) conducted a group study of left and right hemisphere–damaged signers of BSL. They devised tests that included comprehension of single sign and single predicate-verb constructions (e.g., THROW-DART), simple and complex sentences that varied in argument structure and semantic reversibility, locative constructions encoding spatial relationships, constructions involving lexical prepositions, and a final test of classifier placement, orientation, and rotation. Their findings indicated that left hemisphere–damaged (LHD) BSL signers, relative to elderly signing control subjects, exhibited deficits on all comprehension tests, but were better at classifier constructions than lexical prepositions. Right hemisphere–damaged signers (RHD) did not differ from controls on single sign and single predicate-verb construction, or on sentences with a variety of argument structures and semantic reversibility. RHD signers, however, were impaired equally on tests of locative relationships expressed via classifier constructions and prepositions, and on a test of placement, orientation, and rotation of objects. Hickok, Pickell, Klima, and Bellugi (2009) also found the same patterns for classifier production in RHD and LHD patients.

One interpretation of these findings is that the comprehension of spatial language constructions requires not only intact left hemisphere resources, but intact right hemisphere visual-spatial processing mechanisms as well. That is, while both LHD and RHD signers show comprehension deficits, the right hemisphere damaged signers' difficulties stem from more general visual spatial deficits rather than linguistic malfunction per se. The question of whether these visual spatial deficits are nonlinguistic lies at the heart of the debate. In a more recent functional MRI study, Emmorey et al. (2013) also found that viewing both location and motion expressions involving classifier constructions engaged bilateral superior parietal cortex.

Studies of sign language impairments and neural processing thus demonstrate on the one hand, modality-independent aspects of the brain's processing of language; on the other hand, there are also indications that the modality and/or form (i.e., iconic) of the linguistic system may place specific demands on the neural mediation and implementation of language.

2.4. Conclusions: Sign Language   Research on sign languages has shown that when the auditory channel is not available, language can exist within the visual modality alone and reveal many of the linguistic structures identified for spoken languages. Unlike spoken languages, however, where new languages are always derived from interactions of existing languages and language users, new sign languages can emerge when deaf people communicate with each other, even if they have no access to previously existing signed or spoken languages, and are then able to transfer that language system to new generations. This is possible because modern humans can scaffold a new sign language on gesture, even when no language is accessible to them. Thus the cognitive architecture of language can be instantiated anew—out of gesture—even in the absence of full conventionalized language input. Thus the human capacity for language structure is not modality specific—to some extent.

However, although signed languages and spoken/written languages share many features in terms of structure, processing, and neural structure, it has become recently more evident that sign languages make use of iconic structures specifically available to the visual modality. Furthermore there is also evidence that modality-specific brain regions might subserve such modality-specific structures.

## *3. General Conclusions*

Use of the visual modality for linguistic expressions and communication is pervasive both in spoken and sign languages and is inevitable when we think of how languages evolved, emerge anew, and are acquired in a face-to-face context. Understanding the role the visual modality plays in language through sign and gesture (in signed or spoken language) is necessary for our understanding of linguistic structure as well as the cognitive and neural architecture of language. By reviewing the role of gesture and sign language in language structure, processing, and neural correlates, we have aimed to offer a joint perspective toward understanding the role that visual modality plays in both spoken and sign language.

The first insight we get when we look at both co-speech gestures and sign languages is that the visual modality can subserve similar linguistic functions and structures to those found in spoken languages—regardless of whether such visual structures accompany spoken language or whether the visual modality takes the whole burden of language and communication. In spoken languages, gestures are semantically, pragmatically, and syntactically, as well as at the discourse level, integrated into the linguistic structures of the spoken languages. In sign languages, visual modality alone can subserve all levels of linguistic structure such as phonology, morphology, syntax, and the lexicon. Thus the visual modality can pattern and function linguistically in similar ways to those we see expressed through speech.

Cognitive and neural processing of visual expressions, as in gesture or sign language, also bears similarities to those of spoken/written structures. Gestures in production and comprehension are influenced by the processing stages of spoken languages (and vice versa). Comprehension of gestures recruits similar brain areas to those used by spoken languages (i.e., the left lateralized frontotemporal network). Similar cognitive and neural processes are involved in processing both sign and spoken languages. There are also similarities between sign and gesture in terms of their cognitive and neural underpinnings (see for reviews Emmorey & Özyürek, 2014; Özyürek, 2014). These findings suggest, first, that gesture should not be excluded from spoken language research, as this misses a great deal of the structures and processes involved in formulating a linguistic message. Second, communication in the visual modality can be supported by cognitive/neural processes that are not specific to any modality.

However the additional insight we get by looking at the visual modality is that both sign language and gesture can also reveal modality-specific characteristics (i.e., those which cannot be expressed through the speech channel) and in similar ways to each other. The visual modality allows visible, analog, gradient, imagistic, and indexical representations to be expressed together with categorical and descriptive ones (in both spoken and sign languages). As such it allows descriptive and categorical aspects of language to be grounded in the here-and-now to convey the Clarkian *depictive* and *indicating* functions (Clark, 2016). This observation indicates that we need to widen our notion of language and its processing possibilities to modality-specific structures visible in both gesture and sign (Goldin-Meadow & Brentari, 2015; Perniss et al., 2010).

Finally, such iconic and analog, gradient representations also recruit modality-specific brain areas—shared by both gesture and sign—outside of what we described as the classical language network—as can be seen in the greater recruitment of parietal areas and right hemisphere in spatial language processing in sign languages. Sign language perception also recruits areas overlapping with those involved in pantomime perception. These show that the cognitive and neural architecture of language is broader than is traditionally assumed and encompasses modality-specific and "embodied" structures and representations—not only the abstract, categorical and arbitrary ones.

All in all, the review provided in this chapter suggests that our understanding of the structures, processes, and neural architecture of language based on data from spoken languages alone needs to be updated if we wish to fully characterize our linguistic capacity and its cognitive and neural architecture. Recent findings have challenged the views that sign language and spoken language are structured and processed identically and that sign languages do not share similarities to gestures or iconic representations. In fact, there is growing research on spoken languages showing how iconicity (motivated form-meaning mappings) may also be a unique characterizing feature of language when non-European languages are considered, such as Japanese or several African languages that contain many more iconic words in their lexicons than European languages (e.g., Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015). Future research is likely to characterize in more detail the modality-specific and modality-independent aspects of sign language, its cognitive underpinnings and neural correlates. Research on sign languages not of European origin is also needed to generalize the current findings, which are mostly based on ASL and BSL, to other sign languages. Finally we hope that comparisons of spoken and sign languages will also increasingly include audiovisual speech and gesture, rather than only comparing sign languages to heard speech and writing (Goldin-Meadow & Brentari, 2015; Perniss, Özyürek, & Morgan et al., 2015), so that we can more fully understand the modality-specific as well as the modality-independent nature of our language capacity.

## NOTES

1. *Deaf* with an uppercase *D* refers to linguistic communities characterized by the use of sign languages. Lowercase *deaf* refers to an individual's audiological status.
2. When talking about types of gestures, it is important to keep in mind that different scholars have proposed different—and sometimes overlapping—categories and semiotic types of gestures used by speakers. Thus the list given here does not include all of the categories proposed so far (see Müller, 2009).
3. Note that gradient expressions can be also expressed and thus depicted using the speech channel, for example, in *loooong* (the extended duration of the vowel meaning *very long* (Okrent, 2002).
4. Even though there is still debate on this topic, gestures are now considered to have double functions: for the self (e.g., to reduce cognitive load, help verbalization, assist lexical access, and form new thought processes) as well as for the addressee. Many studies show that gestures relating to the same event are produced differently (in terms of size and shape or frequency) in relation to changing needs of the addressee (e.g., knowledge, common ground, age, social space) (Alibali, Heath, & Myers, 2001; Campisi & Özyürek, 2013; Özyürek, 2002).
5. ASL and BSL are historically unrelated and mutually unintelligible.
6. All examples are from BSL unless otherwise stated.
7. It should be noted, however, that simultaneity with two hands is an option exercised differently by different sign languages (Perniss, Zwitserlood, & Özyürek, 2015; Saeed, Sutton-Spence and Leeson, 2000).

## REFERENCES


Alibali, M. W., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language*, *44*, 169–188.

Atkinson, J. A., Marshall, J., Woll, B., & Thacker, A. (2005). Testing comprehension abilities in users of British Sign Language following CVA. *Brain and Language*, *94*(2), 233–248.

Bavelas, J. B., Chovil, N., Lawrie, D. A., & Wade, A. (1992). Interactive gestures. *Discourse Processes*, *15*, 469–489.

Beattie, G., & Shovelton, H. (1999). Do iconic hand gestures really contribute anything to the semantic information conveyed by speech? An experimental investigation. *Semiotica*, *123*, 1–30.

Bernardis, P., & Gentilucci, M. (2006). Speech and gesture share the same communication system. *Neuropsychologica*, *44*(2), 178–190.

Bosworth, R. G., & Emmorey, K. (2010). Effects of iconicity and semantic relatedness on lexical access in American Sign Language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(6), 1573–1581.

Brentari, D. (Ed.) (2010). *Sign languages. Cambridge Language Surveys*. Cambridge: Cambridge University Press.

Campisi, E., & Özyürek, A. (2013). Iconicity as a communicative strategy: Recipient design in multimodal demonstrations for adults and children. *Journal of Pragmatics*, *47*, 14–27.

Carreiras, M., Gutiérrez-Sigut, E., Baquero, S., & Corina, D. (2008). Lexical processing in Spanish Sign Language (LSE). *Journal of Memory and Language*, *58*, 100–122.

Chu, M., & Kita, S. (2016). Co-thought and co-speech gestures are generated by the same action generation process. *Journal of Experimental Psychology: Learning Memory and Cognition*, *42*(2), 257–270.

Chu, M., Meyer, A., Foulkes, L., & Kita, S. (2014). Individual differences in frequency and saliency of speech-accompanying gestures: The role of cognitive abilities and empathy. *Journal of Experimental Psychology: General*, *143*(2), 694–709.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Clark, H. H. (2016). Depicting as a method of communication. *Psychological Review*, *123*, 324–347.

Coppola, M., & Newport E. L. (2005). Grammatical subjects in home sign: Abstract linguistic structure in adult primary gesture systems without linguistic input. *Proceedings of the National Academy of Sciences*, *102*(52), 19249–19253.

Corina, D. P., & Emmorey, K. (1993). *Lexical priming in American Sign Language*. Paper presented at the meeting of the Linguistic Society of American Conference, Philadelphia, PA.

Corina, D. P., Poizner, H., Bellugi, U., Feinberg, T., Dowd, D., & O'Grady-Batch, L. (1992). Dissociation between linguistic and non-linguistic gestural systems: A case for compositionality. *Brain and Language*, *43*, 414–447.

Debreslioska, S., Özyürek, A., Gullberg, M., & Perniss, P. M. (2013). Gestural viewpoint signals referent accessibility. *Discourse Processes*, *50*(7), 431–456.

Defina, R. (2016). Do serial verb constructions describe single events? A study of co-speech gestures in Avatime. *Language*, *92*(4), 890–910.

de Ruiter, J. (2000). The production of gesture and speech. In D. McNeill (Ed.), *Language and gesture* (pp. 248–311). Cambridge: Cambridge University Press.

de Ruiter, J., Bangerter, A., & Dings, P. (2012). The interplay between gesture and speech in the production of referring expressions: Investigating the tradeoff hypothesis. *Topics in Cognitive Science*, *4*(2), 232–248.

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity and systematicity in language. *Trends in Cognitive Sciences*, *19*(10), 603–615.

Dudis, P. (2004). Body partioning and real space blends. *Cognitive Linguistics*, *15*(2), 223–238.

Dye, M. W. G., & Shih, S. (2006). Phonological priming in British Sign Language. In L. M. Goldstein, D. H. Whalen, & C. T. Best (Eds.), *Papers in laboratory of phonology* (Vol. 8, pp. 243–263). Berlin: Mouton de Gruyter.

Emmorey, K. (1991). Repetition priming with aspect and agreement morphology in American Sign Language. *Journal of Psycholinguistic Research*, *20*(5), 365–388.

Emmorey, K. (2002). *Language, cognition, and the brain: Insights from sign language research*. Mahwah, NJ: Lawrence Erlbaum.

Emmorey, K. (2014). Iconicity as structure mapping. *Philosophical Transactions of the Royal Society B*, *369*(1651), 20130301.

Emmorey, K., Grabowski, T., McCullough, S., Damasio, H., Ponto, L., Hichwa, R., & Bellugi, U. (2004). Motor-iconicity of sign language does not alter the neural systems underlying tool and action naming. *Brain and Language*, *89*(1), 27–37.

Emmorey, K., McCullough, S., Mehta, S. H., Ponto, L. B., & Grabowski, T. J. (2013). The biology of linguistic expression impacts neural correlates for spatial language. *Journal of Cognitive Neuroscience*, *25*(4), 517–533.

Emmorey, K., & Özyürek, A. (2014). Language in our hands: Neural underpinnings of sign language and co-speech gesture. In M. S. Gazzaniga, & G. R. Mangun (Eds.), *The cognitive neurosciences* (5th ed., pp. 657–666). Cambridge, MA: MIT Press.

Emmorey, K., Xu, J., Gannon, P., Goldin-Meadow, S., & Braun, A. (2010). CNS activation and regional connectivity during pantomime observation: No engagement of the mirror neuron system for deaf signers. *Neuroimage*, *49*, 994–1005.

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

Floyd, S. (2016). Modally hybrid grammar? Celestial pointing for time-of-day reference in Nheengatú. *Language*, *92*(1), 31–64.

Goldin-Meadow, S. (2003). *The resilience of language: What gesture creation in deaf children can tell us about how all children learn language*. New York: Psychology Press.

Goldin-Meadow, S., & Brentari, D. (2015). Gesture, sign and language: The coming of age of sign language and gesture studies. *Behavioral and Brain Sciences*, *5*, 1–82.

Gu, Y., Mol, L., Hoetjes, M., & Swerts, M. (2017). Conceptual and lexical effects on gestures: The case of vertical spatial metaphors for time in Chinese. *Language, Cognition and Neuroscience*, *32*(8), 1048–1063.

Hall, M. L., Ferreira, V. S., & Mayberry, R. I. (2015). Syntactic priming in American Sign Language. *PLOS ONE*, *10*(3), e0119611.

Hagoort, P., & van Berkum, J. (2007). Beyond the sentence given. *Philosophical Transactions of the Royal Society. Series B: Biological Sciences*, *362*, 801–811.

Haviland, J. B. (1993). Anchoring, iconicity, and orientation in Guugu Yimidhirr pointing gestures. *Journal of Linguistic Anthropology*, *3*(1), 3–45.

Hellige, J. B. (1993). *Hemispheric asymmetry: What's right and what's left*. Cambridge, MA: Harvard University Press.

Hickok, G., Love-Geffen, T., & Klima, E. S. (2002). Role of the left hemisphere in sign language comprehension. *Brain and Language*, *82*(2), 167–178.

Hickok, G., Pickell, H., Klima E., & Bellugi, U. (2009). Neural dissociation in the production of lexical versus classifier signs in ASL: Distinct patterns of hemispheric asymmetry. *Neuropsychologia*, *47*, 382–387.

Hohenberger, A., Happ, D., & Leuninger, H. (2002). Modality dependent aspects of signed language production: evidence from slips of the hands and their repairs in German Sign Language. In R. P. Meier, K. Cormier, & D. Quinto-Pozos (Eds.), *Modality and structure in signed and spoken languages* (pp. 112–142). Cambridge: Cambridge University Press.

Iverson, J. M., & Goldin-Meadow, S. (1998). Why people gesture when they speak. *Nature*, *396*, 228.

Kegl, J., Senghas, A., & Coppola, M. (1999). Creation through contact: Sign language emergence and sign language change in Nicaragua. In M. Degraff, (Ed.), *Language creation and language change* (pp. 179–237). Cambridge, MA: MIT Press.

Kelly, S. D., Barr, D. J., Church, R. B., & Lynch, K. (1999). Offering a hand to pragmatic understanding: The role of speech and gesture in comprehension and memory. *Journal of Memory and Language*, *40*, 577–592.

Kelly, S. D., Özyürek, A., & Maris, E. (2010). Two sides of the same coin: Speech and gesture mutually interact to enhance comprehension. *Psychological Science*, *21*, 260–267.

Kendon, A. (1986). Some reasons for studying gesture. *Semiotica*, *62*, 1–28.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.

Kendon, A. (2014). Semiotic diversity in utterance production and the concept of 'language." *Philosophical Transactions of the Royal Society B*, *369*, 20130293.

Kita, S. (2009). Cross-cultural variation of speech-accompanying gesture: A review. *Language and Cognitive Processes*, *24*(2), 145–167.

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.

Klima, E., & Bellugi, U. (1979). *The signs of language*. Cambridge, MA: Harvard University Press.

Krauss, R. M., Chen, Y., & Gottesman, R. F. (2000). Lexical gestures and lexical access: A process model. In D. McNeill (Ed.), *Language and gesture* (pp. 261–283). New York: 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.

Liddell, S. K. (1990). Four functions of a locus: Reexamining the structure of space in ASL. In L. Ceil (Ed.), *Sign language research: Theoretical issues* (pp. 176–198). Washington, DC: Gallaudet University Press.

Liddell, S. K. (2003). *Grammar, gesture and meaning in American Sign Language*. Cambridge: Cambridge University Press.

MacSweeney, M. F., Campbell, R., Woll, B., Giampietro, V., David, A. S., McGuire, P. K., & Brammer, M. J. (2004). Dissociating linguistic and non-linguistic gestural communication in the brain. *Neuroimage*, *22*(4), 1605–1618.

MacSweeney, M., Capek, C. M., Campbell, R., & Woll, B. (2008). The signing brain: the neurobiology of sign language. *Trends in Cognitive Science*, *12*, 432–440.

MacSweeney, M., Woll, B., Campbell, R., McGuire, P., David, A. S., Williams, S., & Brammer, M. (2002). Neural systems underlying British Sign Language and audio-visual English processing in native users. *Brain*, *125*, 1583–1593.

Marshall, J., Atkinson, J., Smulovitch, E., Thacker, A., & Woll, B. (2004). Aphasia in a user of British Sign Language: Dissociation between gesture and sign. *Cognitive Neuropsychology*, *21*(5), 537–555.

Mayberry, R. I. (2010). Early language acquisition and adult language ability: What sign language reveals about the critical periods. In M. Marschark & P. E. Spencer (Eds.), *The Oxford handbook of deaf studies, language, and education* (2nd ed., pp. 281–291). Oxford: Oxford University Press.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.

McNeill, D. (2005). *Gesture and thought*. Chicago: University of Chicago Press.

McNeill, D., Cassell, J., & Levy, E. (1993). Abstract deixis. *Semiotica*, *95*, 5–19.

Meier, R. P. (2002). Why different, why the same? Explaining effects and non-effects of modality upon linguistic structure in sign and speech. In R. P. Meier, K. A. Cormier, & D. G. Quinto-Pozos (Eds.), *Modality and structure in signed and spoken languages* (pp. 1–25). Cambridge: Cambridge University Press.

Meier, R. P., Mauk, C.E., Cheek, A., & Moreland, C. J. (2008). The form of children's early signs: Iconic or motoric determinants? *Language Learning and Development*, *4*, 63–98.

Morgan, G., & Woll, B. (2007) Understanding sign language classifiers through a polycomponential approach. *Lingua*, *117*(7), 1159–1168.

Morrel-Samuels, P., & Krauss, R. M. (1992). Word familiarity predicts temporal asynchrony of hand gestures and speech. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *18*, 615–623.

Müller, C. (2009). Gesture and language. In K. Malmkjaer (Ed.), *The Routledge linguistics encyclopedia* (pp. 214–217). London: Routledge.

Neville, H. J., Bavelier, D., Corina, D., Rauschecker, J., Karni, A., Lalwani, A., … Turner, R. (1998). Cerebral organization for language in deaf and hearing subjects: Biological constraints and effects of experience. *Proceedings of the National Academy of Sciences USA*, *95*, 922–929.

Okrent, A. (2002) A modality-free notion of gesture and how it can help us with the morpheme vs. gesture question in sign language linguistics (or at least give us some criteria to work with). In R. P. Meier, K. A. Cormier, & D. G. Quinto-Pozos (Eds.), *Modality and structure in signed and spoken languages* (pp. 175–198). Cambridge: Cambridge University Press.

Orfanidou, E., Adam, R., Morgan, G., & McQueen, J. M. (2010). Recognition of signed and spoken language: Different sensory inputs, the same segmentation procedure. *Journal of Memory and Language*, *62*(3), 272–283.

Özçalışkan, S., Lucero, C., & Goldin-Meadow, S. (2016). Is seeing gesture necessary to gesture like a native speaker? *Psychological Science*, *27*(5), 737–747.

Özyürek, A. (2002). Do speakers design their co-speech gestures for their addressees? The effects of addressee location on representational gestures. *Journal of Memory and Language*, *46*(4), 688–704.

Özyürek, A. (2012). Gesture. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign language: An international handbook* (pp. 626–646). Berlin: Mouton.

Özyürek, A. (2014). Hearing and seeing meaning in speech and gesture: Insights from brain and behaviour. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1651), 20130296.

Özyürek, A., Furman, R., & Goldin-Meadow, S. (2014). On the way to language: Event segmentation in homesign and gesture. *Journal of Child Language*, *42*, 64–94.

Özyürek, A., Willems, R. M., Kita, S., & Hagoort, P. (2007). On-line integration of semantic information from speech and gesture: Insights from event-related brain potentials. *Journal of Cognitive Neuroscience*, *19*(4), 605–616.

Padden, C. (1988). *Interaction of morphology and syntax in American sign language*. Outstanding Dissertations in Linguistics, Series IV. New York: Garland Press.

Peeters, D., Hagoort, P., & Özyürek, A. (2015). Electrophysiological evidence for the role of shared space in online comprehension of spatial demonstratives. *Cognition*, *136*, 64–84.

Peeters, D., & Özyürek, A. (2016). This and that revisited: A social and multimodal approach to spatial demonstratives. *Frontiers in Psychology*, *7*, 222.

Peeters, D., Snijders, T. M., Hagoort, P., & Özyürek, A. (2017). Linking language to the visual world: Neural correlates of comprehending verbal reference to objects through pointing and visual cues. *Neuropsychologia*, *95*, 21–29.

Perniss, P. M., & Özyürek, A. (2015). Visible cohesion: A comparison of reference tracking in sign, speech, and co-speech gesture. *Topics in Cognitive Science*, *7*(1), 36–60.

Perniss, P. M., 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. L., & Vigliocco, G. (2010). Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in Psychology*, *1*, 1–15.

Perniss, P. M., Zwitserlood, I., & Özyürek, A. (2015). Does space structure spatial language? A comparison of spatial expression across sign languages. *Language*, *91*(3), 611–641.

Poizner, H., Bellugi, U., & Tweney, R. D. (1981). Processing of formational, semantic, and iconic information in American Sign Language. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(5), 1146–1159.

Poizner, H., Klima, E. S., & Bellugi, U. (1987). *What the hands reveal about the brain*. Cambridge, MA: MIT Press.

Quinto-Pozos, D., & Parrill, F. (2015) Signers and co-speech gesturers adopt similar strategies for portraying viewpoint in narratives. *Topics in Cognitive Science*, *7*, 12–35.

Saeed, J. I., Sutton-Spence, R., & Leeson, L. (2000). *Constituent order in Irish Sign Language and British Sign Language: A preliminary examination*. Poster presented at the meeting of the 7th International Conference on Theoretical Issues in Sign Language Research (TISLR7), Amsterdam, the Netherlands.

Sandler, W., Meir, I., Padden, C., & Aronoff, M. (2005). The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences*, *102*(7), 2661–2665.

San Jose-Robertson, L., Corina, D. P., Ackerman, D., Guillemin, A., & Braun, A. R. (2004). Neural systems for sign language production: Mechanisms supporting lexical selection, phonological encoding, and articulation. *Human Brain Mapping*, *23*, 156–167.

Schembri, A., Jones, C., & Burnham, D. (2005). Comparing action gestures and classifier verbs of motion: Evidence from Australian Sign Language, Taiwan Sign Language, and non-signers' gestures without speech. *Journal of Deaf Studies and Deaf Education*, *10*(3), 272–290.

Senghas, A., Kita, S., & Özyürek, A. (2004). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science*, *305*(5691), 1779–1782.

Slobin, D., Hoiting, N., Kuntze, K., Lindert, R., Weinberg, A., Pyers, J., … Thumann, H., 2003. A cognitive/functional perspective on the acquisition of "classifiers." In K. Emmorey (Ed.), *Perspectives on classifier constructions in sign languages* (pp. 297–310). Mahwah, NJ: Lawrence Erlbaum.

Stokoe, W. (1960). *Sign language structure: An outline of the visual communication systems of the American Deaf*. Silver Spring, MD: Linstok Press.

Strickland, B., Geraci, C., Chemla, E., Schlenker, P., Kelepir, M., & Pfau, R. (2015). *Proceedings of the National Academy of Sciences*, *112*(19), 5968–5973.

Supalla, T. (1986). The classifier system in American Sign Language. In C. Craig (Ed.), *Noun classes and categorization* (pp. 181–214). Amsterdam: John Benjamins.

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

Taub, S. F. (2001). *Language from the body: Iconicity and metaphor in American Sign Language*. Cambridge: Cambridge University Press.

Thompson, R. L., Emmorey, K., & Gollan, T. H. (2005). "Tip of the fingers" experiences by deaf signers insights into the organization of a sign-based lexicon. *Psychological Science*, *16*(11), 856–860.

Thompson, R. L., Vinson, D. P., & Vigliocco, G. (2010). The link between form and meaning in British sign language: Effects of iconicity on phonological decisions. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *36*, 1017–1027.

Thompson, R. L., Vinson, D. P., Woll, B., & Vigliocco, G. (2012). The road to language learning is iconic: Evidence from British Sign Language. *Psychological Science*, *23*(12), 1443–1448.

Tversky, B. (2011). Visualizations of thought. *Topics in Cognitive Science*, *3*, 499–535.

Vermeerbergen, M., Leeson, L., & Crasborn, O. A. (Eds.). (2007). *Simultaneity in signed languages: Form and function* (Vol. 281). Amsterdam: John Benjamins

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*, *369*, 20130292.

Wagner, P., Malisz, Z., & Kopp, S. (2014). Gesture and speech in interaction: An overview. *Speech Communication*, *57*, 209–232.

Wilbur, R. B. (2008). Complex predicates involving events, time, and aspect: Is this why sign languages look so similar? In J. Quer (Ed.), *Theoretical issues in sign language research* (pp. 217–250). Hamburg: Signum Press.

Woll, B., & Ladd, P. (2010). Deaf communities. In M. Marschark & P. Spencer (Eds.), *The handbook of deaf studies, language and education* (2nd ed., pp. 159–172). Oxford: Oxford University Press.

Xu, J., Gannon, P., Emmorey, K., Smith, J. F., & Braun A. R. (2009). Symbolic gestures and spoken language are processed by a common neural system. *Proceedings of the National Academy of Sciences*, *106*(49), 20664–20669.

Yap, D. F., So, W. C., Yap, M., & Tan, Y. Q. (2011). Iconic gestures prime words. *Cognitive Science*, *35*, 171–183.