

Referring strategies in American Sign Language and English (with co-speech gesture): The role of modality in referring to non-nameable objects

ZED SEVCIKOVA SEHYR

San Diego State University

BRENDA NICODEMUS

Gallaudet University

JENNIFER PETRICH and KAREN EMMOREY

San Diego State University

Received: November 22, 2016 Accepted for publication: January 5, 2018

ADDRESS FOR CORRESPONDENCE

Zed Sevcikova Sehyr, 6495 Alvarado Road, suite #200, San Diego, CA 92120.

E-mail: zsevcikova@sdsu.edu

ABSTRACT

American Sign Language (ASL) and English differ in linguistic resources available to express visual–spatial information. In a referential communication task, we examined the effect of language modality on the creation and mutual acceptance of reference to non-nameable figures. In both languages, description times reduced over iterations and references to the figures’ geometric properties (“shape-based reference”) declined over time in favor of expressions describing the figures’ resemblance to nameable objects (“analogy-based reference”). ASL signers maintained a preference for shape-based reference until the final (sixth) round, while English speakers transitioned toward analogy-based reference by Round 3. Analogy-based references were more time efficient (associated with shorter round description times). Round completion times were longer for ASL than for English, possibly due to gaze demands of the task and/or to more shape-based descriptions. Signers’ referring expressions remained unaffected by figure complexity while speakers preferred analogy-based expressions for complex figures and shape-based expressions for simple figures. Like speech, co-speech gestures decreased over iterations. Gestures primarily accompanied shape-based references, but listeners rarely looked at these gestures, suggesting that they were recruited to aid the speaker rather than the addressee. Overall, different linguistic resources (classifier constructions vs. geometric vocabulary) imposed distinct demands on referring strategies in ASL and English.

Keywords: Cognition; Language Production; Signed Languages

© Cambridge University Press 2018 0142-7164/18

Establishing a mutual reference to an object is an important function of communication; however, it is unclear whether the referring processes in spoken and signed communication converge on similar referring strategies. In referential communication, speakers and addressees collaborate together to minimize communicative effort to achieve a common goal (Brennan & Clark, 1996; Clark & Wilkes-Gibbs, 1986). Studies of referential communication have provided insights into the ways spoken referring expressions evolve over time as speakers align their representations and establish a shared understanding in order to mutually agree on a reference to novel objects or concepts (Brennan & Clark, 1996; Clark & Wilkes-Gibbs, 1986; Fowler, 1988; Galati & Brennan, 2010; Krauss & Weinheimer, 1966). Conversational partners use lexical choices relatively consistently once a mutual ground has been established. This process is termed “lexical entrainment” (Brennan & Clark, 1996) and is an example of “alignment” between speakers (Pickering & Garrod, 2006). Lexical entrainment and alignment processes provide a basis for successful communication.

Sign languages differ from spoken languages in the resources available to express visual–spatial information (Emmorey, 1996; Emmorey, Tversky, & Taylor, 2001; Supalla, 1982, 1986). The visual–manual modality of sign languages offers signers opportunities for iconic mapping between sign form and the meaning of a referent, providing signers with a unique set of tools for expressing visual or geometric properties of referents (e.g., via tracing or size-and-shape depicting constructions). Signed descriptions of spatial relationships or object shapes are likely to be more readily transparent to the addressee (at least prior to grounding) than spoken descriptions that have limited opportunities for transparent form–meaning mapping. Although both modalities are equally good at segmented and combinatorial encoding (Goldin-Meadow & McNeill, 1999), the visual–manual modality of sign languages is superior to the auditory–oral modality for capturing the gradient components of referents (e.g., gradient changes in the size or shape of an object; see Emmorey & Herzig, 2003; Sehr & Cormier, 2016). The visual–manual system of sign language also allows signers to simultaneously express many referent properties within a single construction; for example, signers can simultaneously describe the location and orientation of a figure object in relation to a ground object (and the orientation of that ground object) using a depicting (classifier) construction. In contrast, the auditory–oral system of spoken language is limited in this regard (Brentari, 2002; Meier, Cormier, & Quinto-Pozos, 2002).

The production rate for signs is slower than for words, although sign and speech do not differ in the rate of propositional content (Bellugi & Fisher, 1972; Klima & Bellugi, 1979). The slower signing rate encourages simultaneous layering of information and discourages sequential encoding, which tends to be prevalent in spoken language (Klima & Bellugi, 1979). Emmorey (1996) reported that American Sign Language (ASL) descriptions of spatial layouts were significantly shorter than spoken English descriptions, suggesting that the spatialization of ASL descriptions allows for relatively fast and efficient expression of information about locations and positions of objects in space. It remains unclear how the modality-specific resources contribute to communicative efficiency when speakers must agree on mutual reference for difficult-to-name figures, rather than spatial layouts.

However, the ability to express gradient spatial meanings is not limited to sign languages. Manual gestures that co-occur with speech can be used to convey gradient analog information about referents that is not expressed via speech (Goldin-Meadow & Brentari, 2017; McNeill, 1992). For example, in describing spatial layouts (e.g., landmarks in a town) speakers use tracing gestures to indicate shape and path information, and they depict spatial relations between referents by how they move or position their hands while speaking (Emmorey et al., 2001). Thus, we will also examine the role of co-speech gesture in referring to difficult-to-name objects.

In a referential communication task based on Clark and Wilkes-Gibbs (1986), we examined how expressions referring to abstract non-nameable figures (Atneave, 1957) were created and evolved over time in ASL and spoken English. The iterative nature of the task allowed us to investigate how the different linguistic resources for expressing information about shape (including co-speech gesture) impact communication efficiency, what the linguistic and cognitive strategies for reference are, and whether the geometric complexity of the shapes influences referring expression choices.

The conventionalization of reference among interacting partners has been investigated in experimental contexts (for a review, see Galantucci, Garrod, & Roberts, 2012) and computational contexts (Barr, 2004; Steels & Loetzsch, 2012). In a standard iterative referential communication task, interacting partners must collaborate to achieve a common goal in interaction. Interacting partners describe and identify objects in a set over several rounds of description, repeating words and phrases until they converge on a perspective and mutually agree on lexical choices (Brennan & Clark, 1996; Clark & Wilkes-Gibbs, 1986; Garrod & Anderson, 1987). Clark and Wilkes-Gibbs (1986) pioneered the referential communication task to examine how partners interact to support reference over an extended exchange. Pairs of English speakers described a set of Tangram figures (flat shapes put together to create humanlike figures), with the goal to identify and rearrange the figures in a numbered order that matched their partner's ordered layout of figures. One participant, the Director, described his or her set of figures to a partner, the Matcher, who identified each figure from his or her set and rearranged their layout based on the Director's descriptions. Each round was repeated six times, and the Director's numbered arrangement changed each round. In order to complete the task successfully, interacting partners had to converge on a set of linguistic conventions to establish understanding and reach mutual acceptance of each other's expressions. Changes in referring expressions arose from collaborations to minimize effort and establish common ground (Brennan & Clark, 1996; Clark & Wilkes-Gibbs, 1986).

Speakers minimized conversational effort by shortening and simplifying utterances to help their addressee identify the referent in context. Initial descriptions were lengthy and elaborate but became more succinct over time while remaining just as understandable to the addressee (Clark & Wilkes-Gibbs, 1986; Fussell & Krauss, 1992; Galati & Brennan, 2010). Clark and Wilkes-Gibbs (1986) reported a decrease in word count from 41 words per figure on average in the first round to 8 words in Round 6. For example, in Round 1, the description "Okay, the-number 7 looks like, sort of like an angel flying away or something. It's got two arms,"

became “Sixth one’s the angel” in Round 6. Further, other studies show that repeated references to a target item become reduced in duration and acoustic prominence compared to the initial reference as speakers eliminate redundancy and conserve effort (Aylett & Turk, 2004; Bard et al., 2000; Fowler, 1988; Lam & Watson, 2014).

Similar reduction and refinement of reference has been observed in sign languages. In a referential communication task conducted in Sign Language of the Netherlands, repeated references to people and furniture became shorter and contained fewer and shorter signs than initial references (Hoetjes, Krahmer, & Swerts, 2014). Repeated references to objects tended to be judged as less precise by an independent group of perceivers compared to the initial references (Bard et al., 2000; Hoetjes et al., 2014), suggesting that common ground is crucial for communicative success. Without interaction, referring expressions might remain lengthy and complex (Garrod, Fay, Lee, Oberlander, & MacLeod, 2007; Hupet & Chantraine, 1992; Krauss & Weinheimer, 1966). Thus, the shortening and refinement of reference that occurs with the establishment of mutual common ground is independent of language modality.

Here, we examined whether repeated references to non-nameable figures in ASL and spoken English become reduced over successive rounds of descriptions. We expected to find a similar process of reduction in both signed and spoken referential communication and hypothesized that referring expressions to non-nameable shapes become shorter over time in both ASL and English. However, because ASL descriptions might take advantage of spatial (iconic) constructions and simultaneous layering of information, shape descriptions in ASL might be shorter overall than English descriptions, which are limited to sequential descriptions (see Emmorey, 1996).

An important question is how do interacting partners in spoken and signed dialogue employ linguistic resources to reach mutually grounded reference and achieve communicative success? Communicative success can be measured as the length of time it takes to successfully converge on a referring expression. However, the type of referring expression that conveys a particular perspective on how objects should be viewed will also contribute to communicative success. As Clark and Brennan (1991) pointed out, the medium of communication (modality), in addition to a host of other constraints (e.g., visibility, simultaneity, or audibility), influences referring techniques necessary to converge on a perspective and reach mutual grounding.

For spoken English, Clark and Wilkes-Gibbs (1986) demonstrated that descriptions of Tangram figures generally conveyed two main perspectives on how the figures can be viewed or described, either as a whole (e.g., an angel) or as individual parts or features of objects (e.g., wings, feathers, or halo). Expressions referring to the figure as a whole took the form of an analogy (e.g., “It looks like a person, sitting with his legs under him”; Clark & Wilkes-Gibbs, 1986, p. 31; authors use the term *analogical perspective*). Expressions focusing on the figure’s geometric features, relations, or individual segments were direct shape-based descriptions (e.g., “It’s a hexagonal shape, and then on the bottom right side it has this diamond”; Clark & Wilkes-Gibbs, 1986, p. 31; authors use the terms *literal perspective*).

Analogy-based expression allows speakers to refer holistically to the figure by describing its resemblance to a naturally occurring object or concept, typically by naming (e.g., using a noun or noun phrase), while a shape-based expression allows speakers to focus on the permanent visual properties and figure parts. In the Clark and Wilkes-Gibbs (1986) study, English speakers exhibited a robust preference for analogy-based reference across several repetitions. In Round 1, 42% of Tangram figures were described using a combination of analogy and shape-based referring expressions; however, analogy-based expressions (“It looks like a ...”) were almost always produced first, with shape-based expressions produced as secondary elaborations. As communication between the two partners unfolded, by Round 6, 77% of references were analogy-based alone and only 19% contained shape-based geometric expressions, suggesting a much stronger tendency for analogy-based reference. Further, when speakers used an analogy to refer to an object, they tied concepts together (e.g., “person meditating”), in comparison with shape-based expressions where the descriptions of geometric properties of the object were more literal, juxtaposed, and segmented (e.g., “It’s got just a diamond sticking up at the top and then one long column that has something sticking out to the left”; p. 31). Accepting a view of the object as a whole establishes perspectives on each part of the object, but not vice versa (Clark & Wilkes-Gibbs, 1986). Thus, English descriptions of difficult-to-name objects tended to recycle conventional labels to establish a shared perspective and achieve communication efficiency.

The linguistic repertoire of ASL for expressing information about referents allows signers to refer to novel objects that are not easily named in several ways. The hand(s) can represent the object itself statically, by standing in for either the object as a whole or as its parts (e.g., the extended index finger might represent a sticklike object or a sticklike part of an object). The hand(s) can also dynamically represent an object by depicting its shape or size (e.g., two curved handshapes tracing a semicircular path together to outline the shape of a round object or object part). The hand(s) can also trace the object shape or size (e.g., using the index finger to trace the outline of a shape or use of a pinching motion to outline spikes on an object). In these examples, it is primarily the movement of the hand(s), not the hand shape, that conveys information about the referent via outlining or tracing (Zwitserslood, 1996). The range of shape-depicting strategies in ASL provides productive tools for reference. Such referring strategies focus mainly on the salient geometric characteristics of an object to convey a shape-based reference. In addition, lexical signs can provide the signer with ready-to-use labels that express the referent’s perceived resemblance to another naturally occurring object by means of analogy, that is, analogy-based reference (e.g., signing ANGEL for a shape that resembles an angel).

Highly abstract figures with no preexisting names prevent speakers and signers from relying on conventional labels and push them toward a greater degree of linguistic innovation. Previous referential communication studies with signed languages have used photographs of realistic objects as stimuli (e.g., faces, people, furniture, and cars; Hoetjes et al., 2014; Jordan & Battison, 1987). Here we asked how do the different linguistic resources for creating and establishing reference impact the choice of referring expressions in ASL and English for difficult-to-name

objects (“Attneave shapes”) and how do the signers and speakers’ choices impact communicative efficiency? We hypothesized that in referring to such difficult-to-name objects, ASL signers would take advantage of the iconic mapping between aspects of a depicting sign and the object and display a preference for shape-based strategies that express visual–geometric information via depicting or tracing over analogy-based strategies via lexical naming. The means of expressing shape information in spoken English are more limited, and therefore, we predicted that speakers would show a preference for analogy-based referring strategies over shape-based referring strategies, as has been previously demonstrated for English speakers in referential communication tasks. We examined whether the particular choices of speakers and signers impacted communicative efficiency by investigating whether and how the type of referring expression affected description length over time.

In addition, the choice of referring expression might vary depending on the nature or complexity of the figures to be described. Clark and Wilkes-Gibbs (1986) found that more complex Tangram figures resulted in longer descriptions (i.e., number of words) than less complex shapes. The extent to which non-nameable shapes can be associated with existing nameable objects might play an important role in perception or memory and is also likely to influence referring strategies (Vanderplas & Garvin, 1959). An inverse relationship between figure complexity (i.e., the number of points in the Attneave shapes) and association value (i.e., percent of associative responses for the shape) indicates that as shapes decrease in complexity they evoke a greater number and range of lexical associations than higher complexity shapes (Vanderplas & Garvin, 1959). Vanderplas and Garvin (1959) found that highly complex shapes with many points and angles were less likely to resemble real objects than simpler shapes. In this study, we specifically examined how speakers and signers recruit linguistic resources to create and mutually agree on reference to abstract Attneave shapes that vary in shape complexity. We compared referring expressions in English and ASL for shapes varying in the number of points (4, 6, 8, and 12 points) to examine whether signers and speakers are differently sensitive to the complexity level of the shape and whether the type of referring expression used varies as a function of shape complexity. We hypothesized that English descriptions of less complex shapes would contain a greater number of analogy-based references than higher complexity shapes. However, shape attributes and complexity might matter less in ASL due to iconicity and the variety of shape-depicting strategies available in the language.

Finally, references to objects may be multimodal. In referential communication, speakers might employ both speech and manual co-speech gesture to convey information about referents. Representational (iconic) gestures can express analog information not easily encoded in speech (McNeill, 1992). Gestures have been argued to help organize complex information for speaking (Kita, 2000), reduce communicative load (Goldin-Meadow, 2003), or facilitate lexical retrieval (Krauss & Hadar, 1999; Rauscher, Krauss, & Chen, 1996). Although there is a consensus that a tight, co-expressive relationship exists between speech and gesture (Kendon, 1980, 2004; McNeill, 1992), it remains unclear how co-speech gesture contributes to establishing mutual reference in successive referring and whether gesture becomes quantitatively reduced over time, similarly to sign and speech.

Hoetjes, Koolen, Goudbeek, Krahmer, and Swerts (2015) found that repeated spoken references to novel figures often contained manual gestures and that these co-speech gestures also became reduced over time in terms of number, size, and the percentage of two-handed gestures, following the emergence of common ground. The decrease in the number of words and gestures was proportionally the same. Thus, these findings pointed to parallel reduction processes between gesture and speech. Similarly to speech, common ground and conceptual pacts between participants affected gestural referring. Further, when partners converse about mutually known referents or scenes, gestures tend to be reduced in complexity and precision or are judged less informative (Gerwing & Bavelas, 2004; Holler & Stevens, 2007). However, Hoetjes, Krahmer, and Swerts (2015) found that when communication was unsuccessful, gesture rate increased and gestures that followed negative feedback were judged as more precise. These findings suggest that speakers might call upon gesture when communication becomes effortful. Moreover, the relationship between speech and gesture might not be entirely interdependent or linear. Hoetjes, Krahmer, et al. (2015) reported that when communication was unsuccessful, speech became proportionately more reduced than gesture, suggesting that speech and gesture may be, to some extent, separate processes.

We examined co-speech gestures produced with spoken English in order to establish whether the number of co-speech gestures reduces as a function of repetition. We hypothesized that co-speech gestures would decline in later rounds of description as partners establish mutual reference. In addition, we examined whether gesture use varies as a function of referring expression type and shape complexity. We hypothesized that representational (iconic) gestures would co-occur more frequently with shape-based reference rather than analogy-based reference because spatial language appears to promote the use of co-speech gesture (Hostetter, 2011). We also predicted that increasing shape complexity would promote the use of gesture because we hypothesized that as communication becomes more challenging, speakers' descriptions would contain more gestures. Analyzing co-speech gesture patterns can also shed light on modality-specific versus modality-general preferences for creating novel referring expressions.

METHOD

Participants

Ten pairs of deaf ASL signers (Directors mean age = 31.9 years, $SD = 9.6$; 5 female; Matchers mean age = 31.9 years, $SD = 7.1$, 9 female) and 10 pairs of hearing speakers of English (Directors mean age = 24.3 years, $SD = 5.0$, 5 female; Matchers mean age = 25.6 years, $SD = 5.8$, 6 female) participated in the study. The hearing participants were all native speakers of English and had no knowledge of ASL, except 1 participant who had completed one semester of ASL. All deaf participants were congenitally deaf and acquired ASL before age 7. All deaf participants were fluent ASL signers and indicated ASL as their main and preferred language of communication. All deaf and hearing pairs had known each other prior to the study for a minimum of at least 6 months. All participants received payment for their participation.

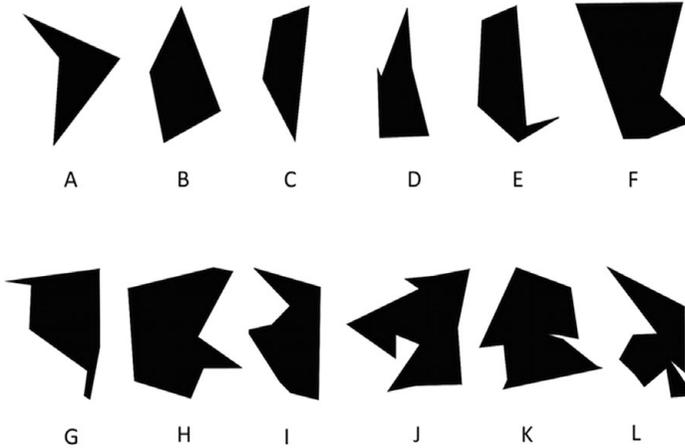


Figure 1. The set of Attneave shapes used in the experiment. The letters were not shown to the participants.

Materials

The stimuli were 12 Attneave shapes (Attneave, 1957; Attneave & Arnoult, 1956) selected from Vanderplas and Garvin (1959), with three shapes at four complexity levels (4, 6, 8, and 12 points; see Figure 1). Each shape was printed in black ink on white on 3- × 4-inch cards and laminated. Two sets of shapes were created, one for the Matcher and one for the Director.

Procedure

Participants sat beside each other at a table with the identical 12 cards laid out in front of them in a randomized order. Participants were seated side by side in order to avoid the viewpoint differences that occur for signed languages with respect to spatial descriptions (see Pyers, Perniss, & Emmorey, 2015). Each card contained one Attneave shape, and each shape was assigned an identifying letter for coding purposes (the identifying letters in Figure 1 did not appear on the stimulus cards). A low divider allowed participants to view each other but not their partner's cards. For one participant, designated the Director, the cards were prearranged in two rows of six shapes. For the other participant, designated the Matcher, the same cards were also arranged in two rows of shapes but in a different order. The task required the Matcher to rearrange the cards to match the Director's layout based on the Director's descriptions. Participants performed the task six times, and each time was considered one round, with the same cards reordered for each round by the experimenter. The deaf participants were instructed in ASL by a deaf experimenter who was a native signer, and the hearing English speakers were instructed in spoken English by a hearing experimenter. The Director was

instructed to describe the shapes as quickly and accurately as possible, starting in the top left corner and continuing left to right for two rows without skipping any shapes. The participants could refer back to the shapes already described if the Matcher requested clarification. Round completion time was recorded by the experimenter using a stopwatch. Participants compared their orderings after each round and determined whether any errors were made. The shapes were then rearranged, and the procedure was repeated in the next round.

Coding and analysis. ASL and English productions were filmed, coded, and analyzed using the software package ELAN (<http://www.lat-mpi.eu/tools/elan>). The Directors' responses were identified as either primary references or secondary elaborations. A primary reference was identified as the Director's first description of a shape in a round. Secondary elaborations were coded as all other referring expressions after the primary reference was provided in each round for a given shape and included reiteration, clarification, or extension of the primary reference. We analyzed secondary elaborations separately in order to determine whether such elaborations pattern differently from initial referring expressions (e.g., shape-based elaborations might accompany analogy-based primary references). In English, primary references and secondary elaborations were separated by a prosodic break, such as a pause or falling intonation. In ASL, primary and secondary expressions were separated by a lowering of the signer's hands to a rest or neutral position. Matchers occasionally produced referential expressions that mainly served as clarification. These productions by the Matcher were relatively rare and were excluded from the analyses.

Referring expressions were coded as *shape based* if the description consisted of words or signs referring to the geometry of the shape, such as "four-sided object" or "triangle." ASL depicting constructions (also referred to as classifiers or classifier constructions) that expressed the size or shape of the figures were also coded as shape based. Referring expressions were coded as *analogy based* if the description referred to an object that is not a geometric shape, such as "house." We coded expressions as *mixed* when participants used both reference types within the same phrase (e.g., "big triangular house"). We coded expressions as *other* if the participants used another type of referring strategy (e.g., "the next one is the last one from the last board"). ASL and English examples are provided in [Figure 2](#). The agreement between two independent judges coding all referring expressions using these categories was 87%. The judges discussed and resolved all discrepancies.

RESULTS

The overall error rate was low and similar for both signers and speakers: 5% for the English pairs and 6% for the ASL pairs. We first conducted a 2 (Language: English, ASL) \times 6 (Rounds: 1–6) repeated-measures analysis of variance with round completion time (in seconds) as the dependent variable. Note that only generalizations over subjects (F1) are reported in this paper; generalizations over items (F2) are not appropriate (Raaijmakers, Schrijnemakers, & Gremmen, 1999)

 K	English	ASL
Shape-based	"A four-sided object on top another four-sided object on the bottom."	 CL:square.entity [above-LOC] CL:square.entity
Analogy-based	"It looks like a house."	 TWO HOUSE
Mixed	"It's flat on the bottom then it goes up and there's a big like box-shaped head."	 BELL CL:square.entity TRIANGLE
Other	"The next one is the last one from the last board."	

Figure 2. Examples of referring expressions in English and ASL. The target shape described is shown in the upper left-hand corner. By convention, signs are glossed with their English translation in capital letters (CL = Classifier).

as stimulus items in this study were not randomly sampled and vary systematically in their physical properties and complexity.

There was a linear decline in round completion times from Round 1 ($M = 364.9$ s, $SE = 55.06$) to Round 6 ($M = 83.15$ s, $SE = 18.60$); main effect of round, $F(5, 90) = 21.57, p < .001, \eta_p^2 = 0.55$. Round completion times were shorter for English speakers than for ASL signers across all rounds (English: $M = 115.13$ s, $SE = 14.66$, ASL: $M = 248.43$ s, $SE = 29.22$); main effect of Language, $F(1, 18) = 5.73, p = .028; \eta_p^2 = 0.24$. There was no interaction between language and round, $F(5, 90) < 1, p = .46$, indicating a similar linear decline in round completion times for ASL signers and English speakers (see Figure 3a).

In addition to round completion times, we analyzed active language duration as the dependent variable (see Figure 3b), that is, the actual time Directors spent describing the shapes in either ASL or English. Short pauses and hesitations within a reference were included, as they could have served an important discourse function providing additional cues to the Matchers. Active language duration also significantly decreased over time from Round 1 ($M = 9.3$ s, $SE = 1.7$) to Round 6 ($M = 3.3$ s, $SE = 0.6$); main effect of round, $F(5, 90) = 22.3, p < .001, \eta_p^2 = 0.55$. English descriptions ($M = 4.5$ s, $SE = 0.9$) were shorter than ASL descriptions ($M = 7.1$ s, $SE = 0.9$), but this difference was only marginally significant,

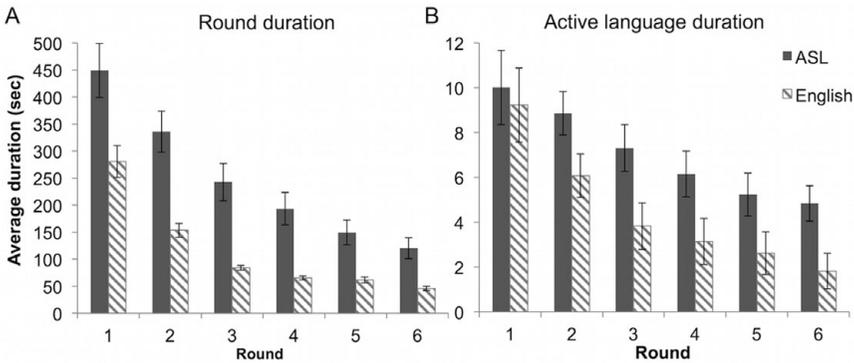


Figure 3. (A) Average round duration (seconds) per each round of ASL and English descriptions, and (B) Average active language duration per each round of ASL and English descriptions.

Table 1. Total number of primary references and secondary elaborations elicited across six rounds of description in ASL and English (% given in parentheses)

	ASL		English	
	Primary	Secondary	Primary	Secondary
Shape-based	479 (67%)	278 (89%)	321 (45%)	100 (46%)
Analogy	179 (25%)	13 (4%)	349 (48%)	63 (29%)
Mixed	54 (8%)	19 (6%)	32 (4%)	38 (17%)
Other	8 (1%)	4 (1%)	18 (3%)	17 (8%)
Total observations	720	314	720	218

$F(1, 18) = 4.3, p = .054, \eta_p^2 = 0.19$. As with round completion times, there was no interaction between language and round for description durations, $F(5, 90) < 1, p = .502, \eta_p^2 = 0.03$.

Referring expressions in English and ASL

There were 1,440 primary referring expressions (720 in each language) and 531 secondary elaborations (313 in ASL and 218 in English). The Director in each participant pair produced 12 primary referring expressions in each round, but the number of secondary elaborations differed across pairs. The number and percentages of primary and secondary references are provided in Table 1, and the distribution of primary and secondary references across rounds is shown in Figure 4.

A binomial logistic regression was conducted to examine the main effects and interactions between language (ASL, English) and round (Rounds 1–6) on the likelihood of a shape-based referring expression as a primary reference and (separately) as a secondary elaboration. For each shape description (i.e., trial) a score of 1 was given when a shape-based description was used and 0 when an

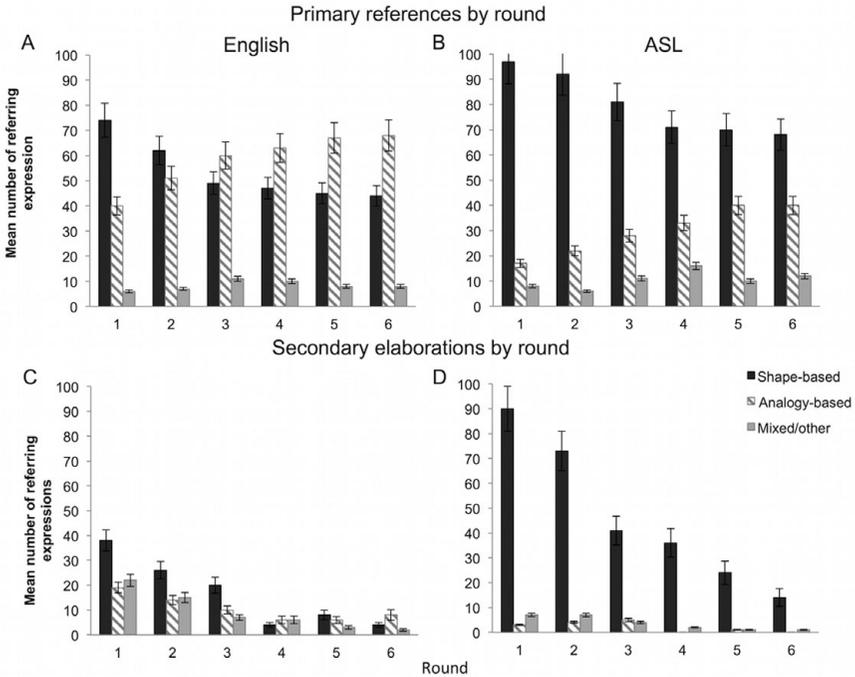


Figure 4. The number of primary references in (A) English and (B) ASL and secondary elaborations in English (C) and ASL (D) across six rounds of descriptions.

analogy-based (or other/mixed) reference was used. Categorical predictor variables were specified in the model using a method of forced entry.

Primary reference. The distribution of primary references is plotted in Figures 4a and 4b. The full model with language and round as categorical variables significantly predicted the likelihood of shape-based referring expression in the data, $\chi^2(11) = 121.3, p < .001$ (see Table 2), correctly classifying 63% of data. There was a main effect of language, $B = 0.88, SE = 0.29, \text{Wald } \chi^2(1) = 9.06, p = .003$; the log odds ratio indicated that ASL signers were 2.4 times more likely to use a shape-based expression as a primary reference than English speakers. The likelihood of shape-based references significantly decreased over time, as we also found a main effect of round, $\text{Wald } \chi^2(5) = 23.6, p < .001$. Finally, there was a parallel decline in shape-based reference in both groups, as we found no Language \times Round interaction, $\text{Wald } \chi^2 = 1.4, p = .925$ (see Table 2). Nonetheless, analogy-based reference in English was dominant by Round 2, while ASL signers maintained preference for shape-based reference throughout the rounds.

Secondary elaboration. The distribution of secondary elaborations is plotted in Figures 4c and 4d. ASL signers produced a total of 314 secondary referring descriptions. Of these, 89% were shape-based and only 4% were analogy-based

Table 2. Binary logistic regression model showing the log odds ratio of shape-based expression in primary reference and secondary elaboration as a function of language and round

		<i>B</i>	<i>SE</i>	Wald	<i>df</i>	<i>p</i>	95% CI for Exp(<i>B</i>)		
							Exp(<i>B</i>) odds ratio	lower	upper
Primary	Constant	0.22	0.05	17.2	1	<.001	1.25	—	—
	Language	0.88	0.29	9.1	1	.003	2.41	1.36	4.28
	Round	—	—	23.6	5	<.001	—	—	—
	Language × Round	—	—	1.4	5	.925	—	—	—
	Round 1 vs. 2	−0.41	0.26	2.4	1	.119	0.66	0.40	1.11
	vs. 3	−0.85	0.26	10.3	1	.001	0.43	0.26	0.72
	vs. 4	−0.92	0.27	11.9	1	.001	0.40	0.24	0.67
	vs. 5	−0.99	0.27	13.7	1	<.001	0.37	0.22	0.63
vs. 6	−1.02	0.27	14.7	1	<.001	0.36	0.21	0.61	
Secondary	Constant	0.90	0.10	89.1	1	<.001	2.47	—	—
	Language	2.27	0.40	31.9	1	<.001	9.71	4.41	21.4
	Round	—	—	5.4	5	.368	—	—	—
	Language × Round	—	—	8.4	5	.136	—	—	—

references. English speakers produced 218 secondary elaborations. Of these, 46% were shape-based and 29% were analogy-based references (see Table 1).

The model significantly predicted the likelihood of shape-based expression in secondary elaborations, $\chi^2(11) = 128.4, p < .001$, with a total prediction success of 75%. However, in this model, only language significantly predicted the outcome as we found a main effect of language, $B = 2.27, SE = 0.40, \text{Wald } \chi^2(1) = 31.9, p < .001$, with the log odds ratio indicating that ASL signers were 9.7 times more likely to produce shape-based expressions as secondary elaborations compared to English speakers. We found no effect of round, $\text{Wald } \chi^2(5) = 5.4, p = .368$, and no Language \times Round interaction, $\text{Wald } \chi^2(5) = 8.4, p = .136$ (see Table 2).

Finally, round completion times were strongly and positively correlated with the number of shape-based expressions overall (primary and secondary combined; $r = .873, p < .001$), and these correlations sustained separately in English ($r = .654, p < .001$) and in ASL ($r = .901, p < .001$). Active language duration was also highly correlated with the number of shape-based expressions overall ($r = .646, p < .001$; English $r = .555, p < .001$; ASL $r = .662, p < .001$). Thus, in both modalities, rounds took longer to complete and descriptions were longer when shape-based expressions were used.

The effect of shape complexity on referring expressions in ASL and English

We performed a logistic regression to examine whether language (ASL, English) and shape complexity (4, 6, 8, and 12 points) influence the likelihood of shape-based referring expression in primary references and secondary elaborations. The models are presented in Table 3. The distribution of primary references across four levels of shape complexity is shown in Figures 5a and 5b.

Primary reference. The model significantly predicted the likelihood of shape-based reference with an overall prediction success of 64.4%, $\chi^2(6) = 130, p < .001$. We found a main effect of shape complexity, $\text{Wald } \chi^2(3) = 53.3, p < .001$, but no main effect of language, $B = 0.28, SE = 0.23, \text{Wald } \chi^2(1) = 1.5, p = .218$, although the log odds ratio indicated that ASL signers were 1.3 times more likely to use shape-based expressions. Further, the Language \times Shape Complexity interaction was significant, $\text{Wald } \chi^2(3) = 21.2, p < .001$. A visual inspection of the data suggests that shape-based expressions decreased with increasing shape complexity. Thus, the interaction term was followed up by a separate logistic regression for each group.

In English, the increasing shape complexity was associated with a significant decline in shape-based referring expressions, $\chi^2(3) = 58.5, p < .001$; however in ASL, shape-based reference remained unaffected by shape complexity, $\chi^2(3) = 1.7, p = .631$. Post hoc contrasts (Bonferroni) confirmed that English speakers used significantly more shape-based reference to describe simpler 4-point shapes in contrast with medium (6-point shapes, $p = .001$) and high (12-point, $p < .001$) complexity shapes, and no difference was found between medium complexity 6-point and 8-point shapes ($p = 1$). Thus, English and ASL appear to differ in the primary referring strategies that are used to describe complex shapes.

Table 3. Binary logistic regression model showing the log odds ratio of shape-based expression in primary reference and secondary elaboration as a function of language and shape complexity

		<i>B</i>	<i>SE</i>	Wald	<i>df</i>	<i>p</i>	95% CI for Exp(<i>B</i>)		
							Exp(<i>B</i>) odds ratio	lower	upper
Primary	Constant	0.22	0.05	17.2	1	<.001	1.25	—	—
	Language	0.28	0.23	1.5	1	.218	1.32	0.85	2.05
	Complexity	—	—	53.3	3	<.001	—	—	—
	Language × Complexity	—	—	21.2	3	<.001	—	—	—
	Complexity 1 vs. 2	−0.77	0.22	12.8	1	<.001	0.46	0.30	0.71
	vs. 3	−0.77	0.22	12.8	1	<.001	0.46	0.30	0.71
	vs. 4	−1.70	0.23	53.3	1	<.001	0.18	0.12	0.29
Secondary	Constant	0.90	0.10	89.1	1	<.001	2.47	—	—
	Language	0.81	0.46	3.1	1	.078	2.24	0.91	5.49
	Complexity	—	—	14.4	3	.002	—	—	—
	Language × Complexity	—	—	11.9	3	.008	—	—	—
	Complexity 1 vs. 2	1.74	0.66	7	1	.008	5.67	1.56	20.57
	vs. 3	1.67	0.63	7.1	1	.008	5.30	1.55	18.12
	vs. 4	2.07	0.69	9	1	.003	7.90	2.05	30.44

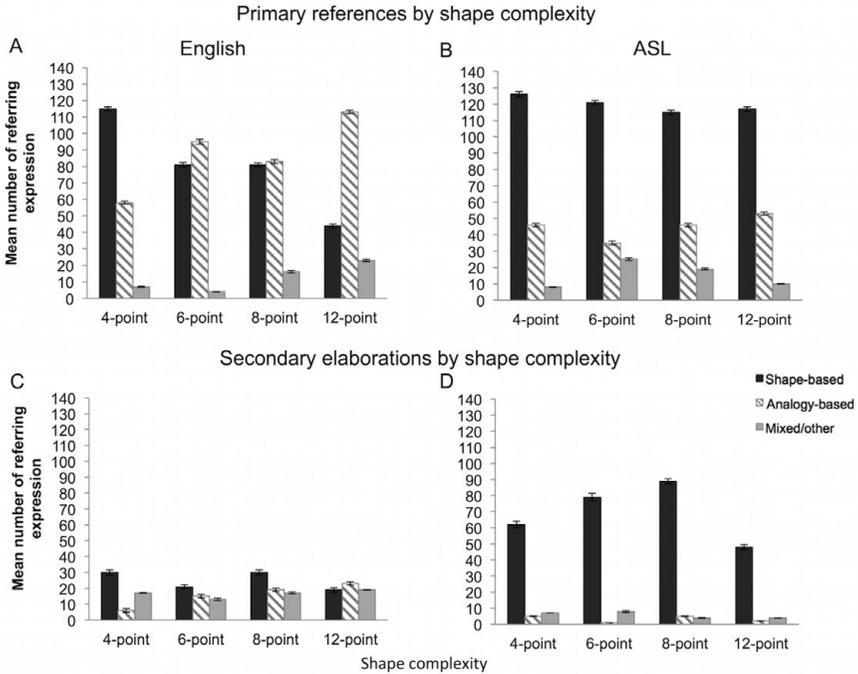


Figure 5. The number of primary references in (A) English and (B) ASL and secondary elaborations in English (C) and ASL (D) per four levels of shape complexity.

Secondary elaboration. The distribution of secondary elaborations across four levels of shape complexity is shown in Figures 5c and 5d. The model with language and shape complexity was statistically significant, correctly predicting 77.8% of outcome, $\chi^2(7) = 136, p < .001$. ASL signers were 2.2 times more likely to use shape-based expressions as secondary elaboration; however, the main effect of language was only marginally significant, $B = 0.81, SE = 0.46, \text{Wald } \chi^2(1) = 3.1, p = .078$. There was a main effect of shape complexity, $\text{Wald } \chi^2(3) = 14.4, p = .002$, and the Language \times Shape Complexity interaction was significant, $\text{Wald } \chi^2(3) = 11.9, p = .008$. A follow-up analysis for each group confirmed a similar pattern to primary references, that is, in English, higher shape complexity was associated with a decrease in shape-based expressions, $\chi^2(3) = 15.6, p = .001$, but in ASL, shape complexity did not influence the shape-based expression patterns, $\chi^2(3) = 2.5, p = .481$. However, the results for secondary elaborations in English must be interpreted with caution because the number of secondary elaborations produced by the English speakers was low.

Co-speech gesture in referential communication

Gestures were coded as representational (iconic) if they bore some resemblance to the figure being described in accordance with McNeill’s (1992) characterization

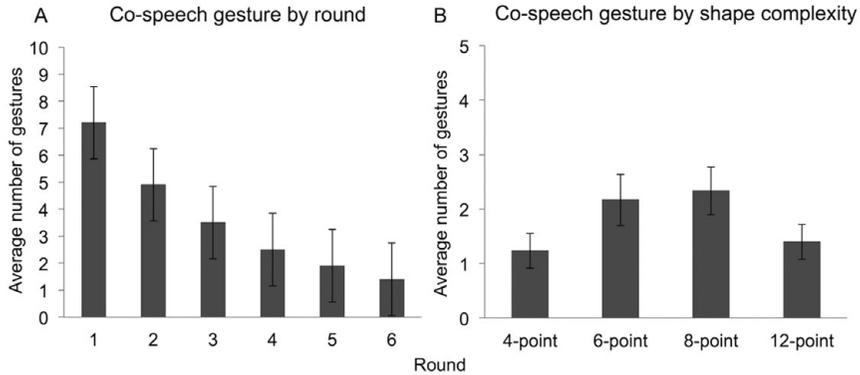


Figure 6. The number of co-speech gestures produced by English speakers per (A) six rounds of English descriptions, and (B) four levels of shape complexity.

of representational gestures. Self-touching movements, nonrepresentational beat gestures (e.g., moving the hand up and down, left or right), and pointing gestures were excluded from the analysis. Representational gestures could depict just a part of an object where the speaker used the hand or fingers to illustrate a feature or part of the shape. Speakers also produced gestures that depicted the object as a whole (e.g., the hand(s) referred to the shape itself) and “tracing” gestures where the speaker used the hand or fingers to trace the outline of the shape. Directors produced 214 representational co-speech gestures in total, 131 of which were produced with primary references. Gestures occurring in primary references and secondary elaborations were collapsed here to increase power. Gesture count was subjected to a repeated-measures analysis of variance with Round (1–6) as the independent variable.

The number of co-speech gestures declined over time from an average of 7.2 gestures in Round 1 ($SE = 1.3$) to 1.4 gestures in Round 6 ($SE = 1.3$), $F(5, 60) = 2.6$, $p = .034$, $\eta_p^2 = 0.20$ (see Figure 6a). In addition, the production of co-speech gestures positively correlated with shape-based references ($r = .12$, $p < .001$), and negatively correlated with analogy-based references ($r = -.21$, $p < .001$). Thus, co-speech gestures tended to co-occur with shape-based expressions and dropped off when speakers landed on a lexical label for a particular shape. Neither average round completion times ($r = .21$, $p = .1$) nor active language duration correlated with the number of co-speech gestures ($r = .04$, $p = .753$), suggesting that the representational gestures produced during this task may be related to a particular construction type (i.e., shape-based expressions), rather than to the amount of speech produced.

In addition, the average count of gestures per Attneave figure revealed a nonlinear relationship between shape complexity and co-speech gesture, linear term: $F(3, 119) = 0.143$, $p = .706$; quadratic term: $F(3, 119) = 5.6$, $p = .02$; gestures were more prevalent with medium-complexity shapes (6-point: $M = 2.2$, $SE = 0.47$; 8-point: $M = 2.3$, $SE = 0.44$) in contrast with low-complexity (4-point: $M = 1.2$, $SE = 0.32$) or high-complexity shapes (12-point: $M = 1.4$, $SE = 0.32$; Figure 6b).

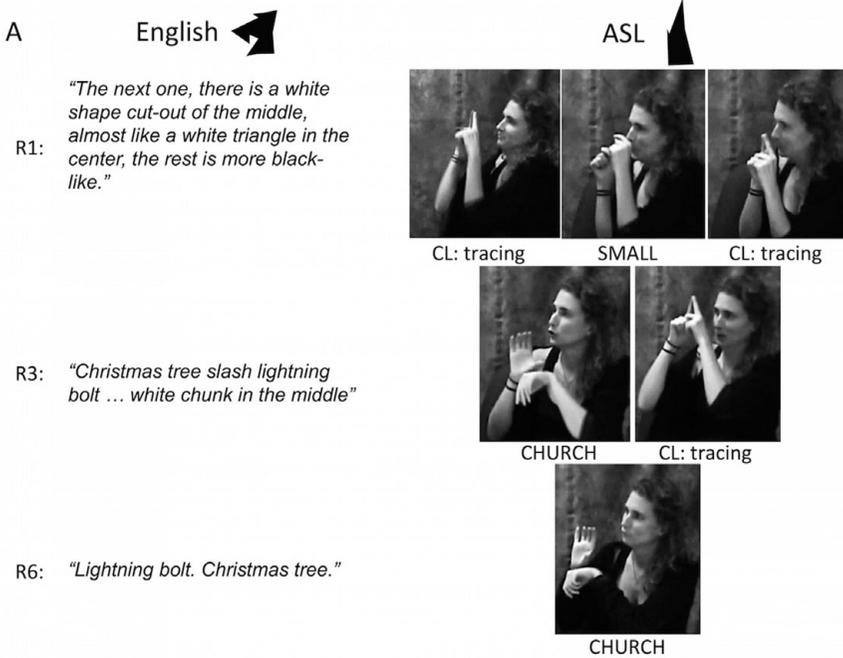


Figure 7a. Examples of a Director's successive descriptions of the illustrated shapes in Rounds 1, 3 and 6 in English and ASL: A shape-based reference describing the geometrical features of the shape in Round 1 becomes an analogy-based reference describing the shape's resemblance to a real object in Rounds 3-6; (CL = Classifier).

This pattern confirms that although gestures co-occurred with shape-based descriptions, more co-speech gestures were produced when shape descriptions became more challenging. The medium-complexity shapes do not easily lend themselves to simple geometric terms or to associative lexical labels, thus presenting a greater descriptive challenge for the speaker.

Finally, we examined how often a Matcher actually looked at the Director when he or she was producing a gesture rather than looking at the array of shapes in front of them. Matchers viewed only 20% of the Directors' gestures, and the Directors rarely produced demonstrative gestures (e.g., "It looks like this [gesture]"). Therefore, gestures in this task might have served primarily to organize information for the speaker (i.e., the Director), rather than to explicitly convey information to their interlocutor (the Matcher).

DISCUSSION

As predicted, we found that referring expressions for non-nameable Attneave shapes became reduced over repeated rounds of interaction, as evidenced by a dramatic decrease in round completion times from Round 1 to Round 6 for both English speakers and ASL signers (Figure 3). Referring expressions in early rounds contained many more signs/words than in later rounds. Figures 7a and 7b provide

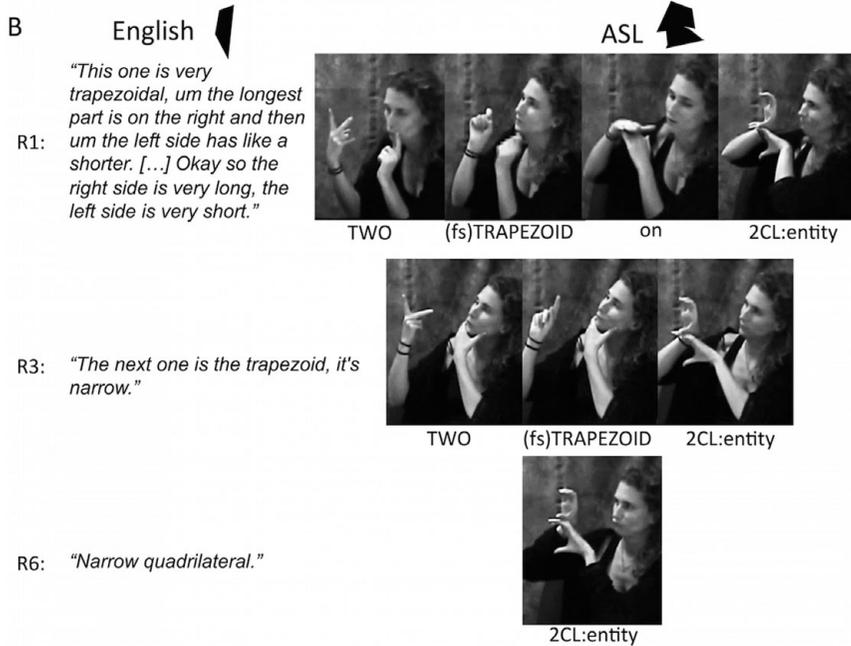


Figure 7b. Examples of a Director's successive descriptions of the illustrated shapes in Rounds 1, 3 and 6 in English and ASL: A shape-based reference in Round 1 remains shape-based for all rounds of description; (fs) = fingerspelling.

examples of shape-based reference reduction in English and ASL between the initial, third, and final round. Consistent with previous research on collaborative referring in spoken language (Brennan & Clark, 1996; Clark & Wilkes-Gibbs, 1986) and sign language (Hoetjes et al., 2014), our results indicate that regardless of language modality, repeated referential expressions become reduced over time as participants establish mutual ground.

Contrary to our expectations, round completion times in ASL were significantly longer than in English, and the active description duration for ASL was marginally longer than for English. This finding contrasts with previous results reported by Emmorey (1996), who found that ASL descriptions of spatial layouts were shorter than spoken English descriptions. The different findings could be due to differences in the type of depicting constructions used to describe the shape of novel figures compared to the spatial arrangement of several objects (furniture in a doll house). To describe the spatial arrangement of objects in a scene, signers can utilize iconic constructions in which each hand represents an object (an "entity" classifier) and the orientation and placement of the hands in space depict the spatial arrangement of the referent objects. This system allows for the simultaneous expression of the location and orientation of two objects and offers a more time-efficient referential strategy than sequential strings of prepositions required to describe the same scene in spoken English. In contrast, descriptions of the novel Attneave shapes required sequential expressions in both ASL and English. The ASL descriptions typically

involved the combination of tracing and entity constructions depicting parts of the referent object as illustrated in Figures 7a and 7b. We suggest that the communication efficiency observed by Emmorey (1996) for ASL compared to English may pertain primarily to the description of spatial, three-dimensional scenes with multiple objects and may not extend to the description of the shape of a single, two-dimensional object.

Longer round completion times in ASL than in English might be associated with distinct demands on gaze behavior for signing versus speaking partners during the task, pointing to an important modality difference related to the task. Hearing speakers were able to view the target shapes while either hearing or producing shape descriptions, but deaf signers could not easily look at the shape display in front of them and simultaneously at their signing partner, who was seated to the side. Rather, signers had to shift their gaze between the display and their partner. Our analysis of co-speech gestures revealed that speakers rarely looked at each other during the task, whereas signers frequently waited until the Director finished describing a shape before turning to look over the display. Nonetheless, this greater demand on memory for the ASL signers did not impair performance because the groups did not differ in error rate, which was low overall. Thus, we can eliminate the possibility that longer completion times for ASL were due to a speed–accuracy trade-off.

In both languages, shape-based references were more prevalent than analogy-based references in the initial rounds and then declined as analogy-based references emerged over time. We hypothesized that signers would make more references to the geometric properties of the shapes than speakers and display a preference for shape-based referring expressions because the shape of a referent can be iconically encoded in ASL depicting constructions. The signers (like the speakers) predominantly used shape-based over analogy-based reference in the initial rounds, and as expected, signers retained this preference until the final round. In comparison, English speakers moved toward analogy-based reference as the preferred strategy by Round 3 (see Figure 4).

Secondary elaborations followed a similar pattern to primary references. ASL signers displayed a robust preference for a shape-based strategy throughout the rounds and rarely used analogy-based elaborations (3.6%), suggesting that ASL signers maintained a shape-based referential strategy between initiating and re-fashioning the referential expression. While ASL signers rarely used mixed/other expressions in secondary elaborations (7%), 25% of secondary elaborations produced by English speakers were mixed/other, suggesting that speakers combined or alternated between referring strategies for the purpose of re-fashioning a referential expression. Further, similar to Clark and Wilkes-Gibbs (1986), we found that in Round 1, 46% of English primary references and secondary elaborations alternated between analogy-based and shape-based reference (cf. 42% in Clark and Wilkes-Gibbs, p. 32), but by Round 6, only 14% alternated between these perspectives (cf. 19% in Clark & Wilkes-Gibbs, p. 32). In ASL, however, only 23% of primary references and secondary elaborations alternated between analogy-based and shape-based reference in Round 1, and by Round 6, there were no descriptions that combined shape-based and analogy-based perspectives. The pattern of secondary elaborations even more clearly reflects the preferential choices for

referring strategies in English and ASL. While English speakers alternated between perspectives for how the Attneave shapes could be viewed, ASL signers consistently chose a shape-based over analogy-based perspective. Why should speakers and signers display different perspective choices?

Clark and Wilkes-Gibbs (1986) argued that English-speaking participants preferred a holistic perspective (e.g., using analogy-based reference) to a segmental perspective (e.g., referring to the parts and properties of a figure) in order to minimize collaborative effort (p. 30). For example, describing an object as a whole will take less cognitive effort on the addressee's part than describing the object in terms of its components in a segmental fashion (e.g., describing an object as "a bed" rather than describing its parts, e.g., "pillow, mattress, headboard"). Based on this argument, English speakers should prefer using an established name as a label for a shape that resembles a bed than describing the shape in terms of its parts. Similar to previous studies of referential communication (Garrod et al., 2007; Pickering & Garrod, 2004), the speakers in the present study tended to negotiate different perspectives, particularly in the initial rounds, but over time they aligned with a perspective that tended to be holistic. In comparison, ASL signers negotiated different perspectives less often and instead preferentially viewed the shapes in a segmental fashion, focusing on the geometric properties. Although the pressure to reduce collaborative effort in establishing mutual reference is similar for both speakers and signers, there may be modality-specific demands on conceptualization of difficult-to-name objects that is reflected in distinct linguistic strategies. For example, due to one-to-one correspondences between a shape description and the actual shape, more cognitive effort might be needed for signers to shift to an analogy-based reference that may not exhibit an iconic mapping between form and meaning. The patterns observed in our study suggest that the use of shape-based depicting strategies for reference to novel objects is preferred in signed communication despite the availability of other linguistic resources (i.e., lexical labels). The ability to perform shape-to-shape mappings between linguistic articulations and geometric forms may underlie this shape-based preference, in contrast to speakers for whom linguistic mappings are arbitrary for both shape-based and analogy-based references.

To further support this point, we found an interaction between language and shape complexity. Specifically, signers' referring expression choices remained unaffected by the complexity of the Attneave shapes, whereas English speakers shifted from shape-based to analogy-based descriptions as complexity increased. We suggest that shape complexity did not influence the choice of referential strategy for signers because of the variety of shape-depicting strategies available in ASL. We had hypothesized that English speakers would be more likely to employ analogy-based strategies to describe simple shapes because these shapes are more likely to resemble real objects (see Vanderplas & Garvin, 1959), thus enabling an analogy-based reference. However, English speakers actually preferred shape-based references to identify simple figures, possibly because it was relatively easy to describe such figures using geometric terms (e.g., *lopsided rectangle*). As figure complexity increased, however, such geometric descriptions became too lengthy and difficult, and speakers shifted to more efficient analogy-based descriptions. This result appears to contrast with Vanderplas and Garvin's (1959) finding of

fewer lexical associations for higher complexity shapes. However, providing possible names for Attneave shapes and providing linguistic descriptions that can distinguish among those shapes are different tasks that have different linguistic demands. For example, participants in the Vanderplas and Garvin study could indicate that a given complex Attneave shape did not remind them of anything (i.e., there was no lexical association), whereas the participants in our study were required to create a referring expression. Thus, when forced to create a reference for complex shapes, the English speakers in our study used an analogy-based strategy that entails a lexical label, while speakers in the Vanderplas and Garvin study simply indicated that these shapes did not evoke a lexical name (unlike the simpler shapes).

Shape-based reference coincided with longer round completion times and figure description times in both languages; put another way, the use of analogy-based reference was associated with shorter completion times for both speakers and signers. Why might this be so? As pointed out by Clark and Wilkes-Gibbs (1986), analogy-based references can be expressed with a single noun phrase that can pick out a variety of differently shaped objects (e.g., *a Christmas tree*, *a bell*, *a mountain*, etc.). In contrast, shape-based references expressed by a noun phrase are limited by the geometric vocabulary of the language (e.g., *three triangles* or *a lopsided rectangle*). To identify more complex Attneave figures using shape-based descriptions generally required segmented references to parts of the figure in order to clearly identify the figure as a whole, and this segmented reference leads to more lengthy descriptions. Thus, the greater use of shape-based references by ASL signers may have contributed to the longer description times and to longer round completion times.

The differential use of shape-based references between the signers and speakers was most prominent for initial descriptions in Round 1 (see Figure 4). The gesture results showed that the Directors' co-speech gestures were most frequent in Round 1 (Figure 6a) and that the gestures often depicted some geometrical or spatial properties of the shapes. Although the primary aim of this paper was to compare spoken vs. signed referring expressions, the early spoken references were often multimodal in nature. It therefore remains a possibility that if co-speech gestures were included in the primary or secondary references as a type of shape-based description, then the initial difference between speakers and signers' reference types would decrease.

ASL is a relatively young language compared to English (e.g., Sandler & Lillo-Martin, 2006) and has a smaller lexicon than English (Emmorey, 2002). It is possible that these factors contributed to the reduced use of analogy-based referring expressions for signers compared to speakers. That is, signers may have used more shape-based expressions because ASL is a young language with a smaller lexicon than English, rather than due to a language-specific bias related to modality. One way to address this question would be to conduct a referential communication task with a younger spoken language, such as a recent creole. A bias toward shape-based descriptions in ASL but not in the creole would confirm that the results of the current study are related to a language-specific bias and not to age of the language or size of the lexicon.¹ Another way to approach this question is to consider other types of stimuli, such as Tangram figures or Greebles (Clark & Wilkes-Gibbs, 1986;

Hoetjes, Koolen, Goudbeek, Krahmer, & Swerts, 2011; Hoetjes, Koolen, et al., 2015). If an early preference for shape-based description was again observed for ASL compared to English for these stimuli, it would confirm a language-specific influence on how referential expressions are initially created.

Previous laboratory studies of referential communication showed that nonlinguistic (graphical) referring expressions tend to evolve from highly iconic representations to more abstract, arbitrary symbolic representations through a grounding process and refinement similar to that proposed for interactive spoken and signed communication (Garrod et al., 2007). For example, Garrod et al. (2007) found that graphical descriptions of concepts (e.g., places, objects, and people) became simpler (e.g., the amount of ink used decreased) and more schematic (less iconic) over time. The authors suggested the distinction between iconic and noniconic symbols is graded and that iconicity can be measured by complexity of expression, with complex expressions being more iconic. Graphical referring expressions were found to maintain their complexity and iconicity in the absence of feedback from a communication partner (Garrod et al., 2007; Hupet & Chantraine, 1992). We suggest that the maintenance of iconic mappings in sign language referring processes is not merely residual or an accident of the visual modality and its affordances. Rather, the preference for shape-based references containing a range of depicting constructions indicates that iconic mapping is strategically deployed for establishing reference in ASL. Signers draw from an inventory of productive morphology dedicated for expressing visual-spatial information that is unavailable to gesturers or speakers. The resulting morphologically complex ASL forms preserve iconicity and become “entrained” over several reiterations, suggesting that unlike the graphical communication systems, novel sign language forms do not necessarily evolve toward an arbitrary system (see Morford, 1996, for a review).

We found that speakers sometimes gestured when describing a shape (particularly during the first round). Iconic gesture production dwindled over time, but it did not disappear completely, and this finding is in line with previous studies that reported that gesture rate declines in repeated referring in successful or unproblematic interactions (de Ruiter, Bangerter, & Dings, 2012; Galati & Brennan, 2014; Hoetjes, Koolen, et al., 2015; Jacobs & Garnham, 2007). The fact that Directors continued to gesture despite their partners rarely looking at their gestures implies that co-speech gestures served primarily to organize information for the Director. This argument is validated by the finding that spoken shape-based expressions declined as Attneave shapes became more complex, while gesture did not. Thus, gestures seem to be recruited to aid the speaker when descriptions of shapes were difficult (i.e., for shapes that could not be readily described using simple geometrical terms or an analogy-based label). The facilitatory role of co-speech gesture to the speaker has previously been recognized, for example, during lexical retrieval (Krauss, Chen, & Gottesman, 2000; Rauscher et al., 1996), when speakers organize complex information for speaking (Kita, 2000), and when children learn and process difficult concepts (Gunderson, Spaepen, Gibson, Goldin-Meadow, & Levine, 2015; Trofatter, Kontra, Beilock, & Goldin-Meadow, 2015).

Finally, because gesture co-occurred more with shape-based than analogy-based reference and declined over time as speakers agreed on labels for the shapes, we propose that spatial, geometrically oriented language promotes the use of gesture

(see also Hostetter, 2011). Further, the lack of relationship between the number of gestures produced and round completion times or actual description times suggests that the content of speech, rather than the length of the rounds or the amount of speech, is what determines the frequency of co-speech gesturing. This proposal is in line with previous studies reporting semantic integration between co-speech gestures and verbal meanings (see, e.g., Özyürek, Willems, Kita, & Hagoort, 2007). However, the precise role of co-speech gesture in creating and establishing a mutually agreed upon reference requires further research (see Hoetjes et al., 2015, for some evidence that lack of partner visibility impacts gesture but not speech in a referential communication task).

In summary, we examined referring strategies to non-nameable Attneave shapes as a function of repetition, language modality, and stimulus complexity. The iterative and conversational nature of the referential communication task provided a valid method to experimentally investigate how interacting partners create and establish referring expressions over a short time span. The reduction patterns were consistent with previous research on collaborative referring in both spoken and signed modalities, including gesture, suggesting that interacting partners minimize collaborative effort independently of communication modality. We demonstrated that sign language differs from speech (without gesture) in the resources available to express visual-spatial information about difficult-to-name shapes, such as iconicity and simultaneous expression, and that language modality had a specific impact on the descriptive choices of speakers compared with signers, placing distinct demands on referential communication. In addition, language modality interacted with stimulus properties because shape complexity distinctly impacted referring strategies in English but not in ASL.

Finally, we found that co-speech gestures reduced in number similarly to previous findings with speech and sign, and that gestures tended to accompany shape-based references. A nonlinear relationship between the use of gesture and stimulus complexity indicated that gestures might be recruited when speaking becomes difficult. The fact that listeners (Matchers) did not look at the speakers' gestures supports the hypothesis that co-speech gestures served primarily for the benefit of the speaker, rather than to explicitly convey information to the addressee. The study offered important insights into how language modality shapes referring strategies for identifying novel objects and contributes to studies of conventionalization in signed and spoken languages in general.

ACKNOWLEDGMENTS

[removed for peer review]

NOTE

1. We thank an anonymous reviewer for this suggestion.

REFERENCES

- Attneave, F. (1957). Physical determinants of the judged complexity of shapes. *Journal of Experimental Psychology*, 55, 221–227. doi:10.1037/h0043921

- Attneave, F., & Arnoult, M. D. (1956). The quantitative study of shape and pattern perception. *Psychological Bulletin*, 53, 452–471. doi:[10.1037/h0044049](https://doi.org/10.1037/h0044049)
- Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*, 47, 31–56. doi:[10.1177/00238309040470010201](https://doi.org/10.1177/00238309040470010201)
- Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language*, 42, 1–22. doi:[10.1006/jmla.1999.2667](https://doi.org/10.1006/jmla.1999.2667)
- Barr, D. J. (2004). Establishing conventional communication systems: Is common knowledge necessary? *Cognitive Science*, 28, 937–962. doi:[10.1016/j.cogsci.2004.07.002](https://doi.org/10.1016/j.cogsci.2004.07.002)
- Bellugi, U., & Fisher, S. (1972). A comparison of sign language and spoken language. *Cognition*, 1, 173–200.
- Brennan, S. E., & Clark, H. H. (1996). Lexical choice and conceptual pacts in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1482–1493. doi:[10.1037//0278-7393.22.6.1482](https://doi.org/10.1037//0278-7393.22.6.1482)
- Brentari, D. (2002). Modality differences in sign language phonology and morphophonemics. In R. P. Meier, K. Cormier, & D. Quinto-Pozos (Eds.), *Modality and structure in signed and spoken languages* (pp. 27–64). Cambridge: Cambridge University Press.
- Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognitions* (pp. 127–149). Washington, DC: American Psychological Association.
- Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22, 1–39. doi:[10.1016/0010-0277\(86\)90010-7](https://doi.org/10.1016/0010-0277(86)90010-7)
- de Ruiter, J. P., Bangertter, A., & Dings, P. (2012). The interplay between gesture and speech in the production of referring expressions: Investigating the trade-off hypothesis. *Topics in Cognitive Science*, 4, 232–248. doi:[10.1111/j.1756-8765.2012.01183.x](https://doi.org/10.1111/j.1756-8765.2012.01183.x)
- Emmorey, K. (1996). The confluence of space and language in signed languages. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and space*. Cambridge, MA: MIT Press.
- Emmorey, K., & Herzig, M. (2003). Categorical versus gradient properties of classifier constructions in ASL. In K. Emmorey (Ed.), *Perspectives on classifier constructions in sign languages* (1st ed., pp. 221–246). Mahwah, NJ: Erlbaum.
- Emmorey, K., Tversky, B., & Taylor, H. A. (2001). Using space to describe space: Perspective in speech, sign, and gesture. *Spatial Cognition and Computation*, 2, 157–180.
- Fowler, C. A. (1988). Differential shortening of repeated content words produced in various communicative contexts. *Language and Speech*, 31, 307–319. doi:[10.1177/002383098803100401](https://doi.org/10.1177/002383098803100401)
- Fussell, S. R., & Krauss, R. M. (1992). Coordination of knowledge in communication: Effects of speakers' assumptions about what others know. *Journal of Personality and Social Psychology*, 62, 378–391. doi:[10.1037//0022-3514.62.3.378](https://doi.org/10.1037//0022-3514.62.3.378)
- Galantucci, B., Garrod, S., & Roberts, G. (2012). Experimental semiotics. *Language and Linguistics Compass*, 6, 477–493. doi:[10.1002/lnc3.351](https://doi.org/10.1002/lnc3.351)
- Galati, A., & Brennan, S. E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee? *Journal of Memory and Language*, 62, 35–51. doi:[10.1016/j.jml.2009.09.002](https://doi.org/10.1016/j.jml.2009.09.002)
- Galati, A., & Brennan, S. E. (2014). Speakers adapt gestures to addressees' knowledge: Implications for models of co-speech gesture. *Language, Cognition and Neuroscience*, 29, 435–451. doi:[10.1080/01690965.2013.796397](https://doi.org/10.1080/01690965.2013.796397)
- Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, 27, 181–218. doi:[10.1016/0010-0277\(87\)90018-7](https://doi.org/10.1016/0010-0277(87)90018-7)
- Garrod, S., Fay, N., Lee, J., Oberlander, J., & MacLeod, T. (2007). Foundations of representation: Where might graphical symbol systems come from? *Cognitive Science*, 31, 961–987. doi:[10.1080/03640210701703659](https://doi.org/10.1080/03640210701703659)

- Gerwing, J., & Bavelas, J. (2004). Linguistic influences on gesture's form. *Gesture*, 4, 157–195. doi:10.1075/gest.4.2.04ger
- Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge, MA: Harvard University Press.
- Goldin-Meadow, S., & Brentari, D. (2017). Gesture, sign, and language: The coming of age of sign language and gesture studies. *Behavioral and Brain Sciences*, 40, 1–60. doi:10.1017/S0140525X15001247
- Goldin-Meadow, S., & McNeill, D. (1999). The role of gesture and mimetic representation in making language the province of speech. In M. C. Corballis & S. Lea (Eds.), *The descent of mind* (pp. 155–172). Oxford: Oxford University Press.
- Gunderson, E. A., Spaepen, E., Gibson, D., Goldin-Meadow, S., & Levine, S. C. (2015). Gesture as a window onto children's number knowledge. *Cognition*, 114, 14–28. doi:10.1016/j.cognition.2015.07.008
- Hoetjes, M., Koolen, R. M. F., Goudbeek, M. B., Krahmer, E. J., & Swerts, M. G. J. (2011). *GREEBLES greeble greeb*. On reduction in speech and gesture in repeated references. Paper presented at the 33rd Annual Conference of the Cognitive Society, Austin, TX.
- Hoetjes, M., Koolen, R. M. F., Goudbeek, M. B., Krahmer, E. J., & Swerts, M. G. J. (2015). Reduction in gesture during the production of repeated references. *Journal of Memory and Language*, 79–80, 1–17. doi:10.1016/j.jml.2014.10.004
- Hoetjes, M., Krahmer, E., & Swerts, M. (2014). Do repeated references result in sign reduction? *Sign Language and Linguistics*, 17, 56–81. doi:10.1075/sll.17.1.03hoe
- Hoetjes, M., Krahmer, E. J., & Swerts, M. G. J. (2015). On what happens in gesture when communication is unsuccessful. *Speech Communication*, 72, 160–175. doi:10.1016/j.specom.2015.06.004
- Holler, J., & Stevens, R. (2007). The effect of common ground on how speakers use gesture and speech to represent size information. *Journal of Language and Social Psychology*, 26, 4–27. doi:10.1177/0261927X06296428
- Hostetter, A. (2011). When do gestures communicate? A meta analysis. *Psychological Bulletin*, 137, 297–315. doi:10.1037/a0022128
- Hupet, M., & Chantraine, Y. (1992). Changes in repeated references: Collaboration or repetition effects? *Journal of Psycholinguistic Research*, 21, 485–496. doi:10.1007/BF01067526
- Jacobs, N., & Garnham, A. (2007). The role of conversational hand gestures in a narrative task. *Journal of Memory and Language*, 56, 291–303. doi:10.1016/j.cognition.2015.07.008
- Jordan, K. I., & Battison, R. (1987). A referential communication experiment with foreign sign languages. *Sign Language Studies*, 1056, 275–287. doi:10.1353/sls.1987.0017
- Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (Ed.), *The relationship of verbal and nonverbal communication* (pp. 207–227). The Hague, Netherlands: Mouton.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 162–185). Cambridge: Cambridge University Press.
- 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 retrieval: A process model. In D. McNeill (Ed.), *Language and gesture* (pp. 261–283). Cambridge: Cambridge University Press.
- Krauss, R. M., & Hadar, U. (1999). The role of speech-related arm/hand gestures in word retrieval. In R. Campbell & L. Messing (Eds.), *Gesture, speech, and sign* (pp. 93–116). Oxford: Oxford University Press.
- Krauss, R. M., & Weinheimer, S. (1966). Concurrent feedback, confirmation, and the encoding of referents in verbal communication. *Journal of Personality and Social Psychology*, 4, 343–346. doi:10.1037/h0023705

- Lam, T. Q., & Watson, D. G. (2014). Repetition reduction: Lexical repetition in the absence of referent repetition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 829–843. doi:10.1037/a0035780
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.
- Meier, R. P., Cormier, K., & Quinto-Pozos, D. (2002). *Modality and structure in signed and spoken languages*. Cambridge: Cambridge University Press.
- Morford, J. (1996). Insights to language from the study of gesture: A review of research on the gestural communication of non-signing deaf people. *Language & Communication*, *16*, 165–178. doi:10.1016/0271-5309(96)00008-0
- 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*, 605–616. doi:10.1162/jocn.2007.19.4.605
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*. doi:10.1017/s0140525x04000056
- Pickering, M. J., & Garrod, S. (2006). Alignment as the basis for successful communication. *Research on Language and Computation*, *4*, 203–228. doi:0.1007/s11168-006-9004-0
- Pyers, J. E., Perniss, P., & Emmorey, K. (2015). Viewpoint in the visual-spatial modality: The coordination of spatial perspective. *Spatial Cognition and Computation*, *15*, 143–169. doi:10.1080/13875868.2014.1003933
- Raaijmakers, J. G. W., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How to deal with “The language-as-fixed-effect fallacy”: Common misconceptions and alternative solutions. *Journal of Memory and Language*, *41*, 416–426. doi:10.1006/jmla.1999.2650
- Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: The role of lexical movements in speech production. *Psychological Science*, *7*, 226.
- Sehyr, S. Z., & Cormier, K. (2016). Perceptual categorization of handling handshapes in British Sign Language. *Language and Cognition*, *8*, 501–532. doi:10.1017/langcog.2015.4
- Steels, L., & Loetzsch, M. (2012). The grounded naming game. In L. Steels (Ed.), *Experiments in cultural language evolution* (pp. 41–59). Amsterdam: John Benjamins.
- Supalla, T. (1982). *Structure and acquisition of verbs of motion and location in American Sign Language* (Unpublished doctoral dissertation, University of California at San Diego).
- Supalla, T. (1986). The classifier system in American Sign Language. In C. Craig (Ed.), *Noun classification and categorization* (pp. 181–214). Philadelphia: John Benjamins.
- Trofatter, C., Kontra, C., Beilock, S., & Goldin-Meadow, S. (2015). Gesturing has a larger impact on problem-solving than action, even when action is accompanied by words. *Language, Cognition and Neuroscience*, *30*, 251–260. doi:10.1080/23273798.2014.905692
- Vanderplas, J. M., & Garvin, E. A. (1959). The association value of random shapes. *Journal of Experimental Psychology*, *57*, 147–154. doi:10.1037/h0048723
- Zwitzerlood, I. (1996). *Who'll handle the object?: An investigation of the NGT classifier* (Unpublished master's dissertation, Utrecht University).